2023

THE EFFECTS OF SOCIOECONOMIC STATUS AND LANGUAGE DIFFERENCE ON SPECIFIC COGNITIVE TEST PERFORMANCE: A COMPARATIVE STUDY OF AFRICAN AMERICAN CHILDREN

Fernanda Carvalho

THE EFFECTS OF SOCIOECONOMIC STATUS AND LANGUAGE DIFFERENCE ON SPECIFIC COGNITIVE TEST PERFORMANCE: A COMPARATIVE STUDY OF AFRICAN AMERICAN CHILDREN

A dissertation submitted in partial fulfillment

of the requirements for the degree of

DOCTOR OF PSYCHOLOGY

to the faculty of the

DEPARTMENT OF PSYCHOLOGY

of

ST. JOHN'S COLLEGE OF LIBERAL ARTS AND SCIENCES

at

ST. JOHN'S UNIVERSITY

New York

by

Fernanda Carvalho

Date Submitted: 08/17/2023

Date Approved: 08/17/2023

Fernanda Carvalho

Samuel O. Ortiz, Ph.D

ABSTRACT

THE EFFECTS OF SOCIOECONOMIC STATUS AND LANGUAGE DIFFERENCE ON SPECIFIC COGNITIVE TEST PERFORMANCE: A COMPARATIVE STUDY OF AFRICAN AMERICAN CHILDREN

Fernanda Carvalho

Previous research has established a high correlation between socioeconomic status and race/ethnicity. Historically, indicators of SES are often patterned by race, with African American children generally being the highest percentage of children living in poverty. The implication of childhood SES to brain development lies in its evident relationship to cognitive ability as measured by cognitive assessments, particularly in the area of language. The present study investigated the influence of language on performance of African American children when compared to normative performance on the WISC-V, a common and widely used cognitive assessment tool. Test creators claim to have included SES as a stratified variable that is highly controlled for in normative samples of cognitive assessments. However, although low SES groups may be included in the norm sample, the norm sample is likely representative of middle-class SES, based on the average of all SES groups. Given that differences in language development in children of the same age leads to significant differences in performance (Cormier et al., 2022; Ortiz, 2018), the assumption that stratification of SES provides adequate representation may be questionable. Results from the present study indicated that both SES and language differences were statistically significant and showed a medium to large effect size in terms of performance. These findings highlight the explicit role of SES and language as variables that should not be so easily overlooked by test

developers. Considering what we know about the implications of SES and language on child development, using norms derived from current stratification procedures may well represent inappropriate practice. These findings are of considerable value for school psychologists when selecting assessment tools, such as the Ortiz PVAT or the KABC-II, that may be more appropriate for use with African American and other culturally, linguistically, and socio-economically diverse children.

DEDICATION

This dissertation is entirely dedicated to my parents whose courage and sacrifice immigrating to the United States taught me resilience, compassion, and endless ambition. All of my academic accomplishments and personal successes are reflections of their ongoing and unconditional love and support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Introduction

It has been established that socioeconomic status (SES), often measured as a combination of education, income and occupation (APA, 2021), has shown to have a significant impact on child development. High SES has been shown to strongly correlate with factors such as parental cognitive stimulation (quantity and quality of speech that children receive from caregivers), material wealth, occupation, access to education, high-quality neighborhoods, social networks, and reduced stress, amongst many others (Demir & Küntay, 2014). As SES inequality continuously widens across the United States, it is imperative to consider the effects of SES on individual differences, including language development. Given that early language skills are predictors of later academic achievement (Demir & Küntay, 2014), understanding the cognitive mechanisms underlying the relation between SES and language development is crucial.

Research has also established a significantly high correlation between socioeconomic status and race/ethnicity. Historically, indicators of SES are often patterned by race. At the same time, the categorization of socioeconomic status and social groups into races/ ethnicities has in fact mirrored oppression and social inequality (Pieterse, Chung, Khan, & Bissram, 2019). Racial differences in SES have also been established to be strong predictors of racial disparities in educational outcomes among African Americans and Whites (Williams, Priest, & Anderson, 2019). Moreover, African American children are of the highest percentage of children living in poverty and enrolled in special education and remedial programs in the United States (Harris & Schroeder, 2013). To this end, when considering the aspect of education, SES is one of the most extensively researched variables in the study of cognitive performance. Thus, test developers included SES as a stratified variable that is highly controlled for in normative

samples of cognitive assessments. Although, low SES groups are reportedly included in the norm sample, the norm sample is likely representative of middle-class SES, as it is based on the average of all SES groups. Therefore, this form of representation is likely to yield inequitable and possible discriminatory results, when assessing low SES students.

Research has also demonstrated mixed views on language differences between African American and White children. Given the assumption that stratification of language, above and beyond SES, has been deemed unnecessary in development of normative samples, test performance between the two groups should yield similar results. The present study therefore examined the extent to which potential language differences in African American students of low SES backgrounds, play a role in cognitive test performance. This study will propose that the issue of language should be highly considered when interpreting test results, especially when accounting for the historical, social and political roots of language development in African American communities. It is hypothesized that language differences found in this study will be statistically significant and will show a medium to large effect size in terms of performance difference between the sample group and the norm. As a result, language will in fact be a potential variable that should not be overlooked by test developers.

Literature Review

Measures of SES

SES is a multifaceted construct that often includes measures of educational level, income, and occupation (APA, 2021). In addition to economic and educational levels however, supplementary social factors such as power, hierarchical social status, psychosocial and neighborhood experiences, and access to resources are all contributors

to variations in SES (Duncan and Magnuson, 2012). Thus, measures of SES as it pertains to influences on child development, is complex and highly controversial. Although correlated, discrepancies among measures of SES do in fact exist and should therefore not be used interchangeably. As a way of example, childhood socioeconomic status is measured by caregivers' or parental SES, which alone is a strong predictor of child developmental outcomes (Duncan and Magnuson, 2012). Psychologists and sociologists have also suggested that the quality of family relationships is a strong predictor of child development, and has been found to be negatively impacted by economic disadvantage (Duncan and Magnuson, 2012). The literature is clear in demonstrating that high SES strongly correlates with factors such as, parental cognitive stimulation (quantity and quality of speech that children receive from caregivers), material wealth, occupation, access to education, high-quality neighborhoods, social networks, reduced stress, amongst many others (Demir & Küntay, 2014).

SES and Race

Research has established a significantly high correlation between socioeconomic status and race/ethnicity. Historically, indicators of SES are evidently patterned with race. To this end, the categorization of socioeconomic status and social groups into races/ ethnicities have historically mirrored oppression and social inequality (Pieterse, Chung, Khan, & Bissram, 2019). According to data from the 2020 US Census Bureau, among the major racial groups examined in the report, African Americans were found to have the highest poverty rate (19.5 percent), with 8.5 million individuals in poverty and not statistically different from the 2019 data (Shrider, Kollar, Chen, & Semega, 2021). Moreover, African American children are of the highest percentage of children living in

poverty (Shrider et al., 2021). Racial differences in SES have also been established to be strong predictors of racial disparities in educational outcomes among African Americans and Whites (Williams, Priest, & Anderson, 2019).

SES and Language Development

The implication of childhood SES to brain development lies in its evident relationship to cognitive ability as measured by cognitive assessments, and begins in early childhood (Hackman & Farah, 2009). As implied by results of numerous studies on childhood development and SES, neurocognitive systems are not homogeneously affected by SES. Particularly, language appears to be one of the main cognitive systems strongly impacted by socioeconomic status (Hackman & Farah, 2009). Influences of SES have been observed in vocabulary, phonological awareness, and syntax, at several stages of development, providing clear behavioral evidence for language system disparities between low and high SES children (Hackman and Farah, 2009). Hackman and Farah (2009) shared results of several studies that revealed disproportionate effects of poverty across different neurocognitive systems, including language. Specifically, one study demonstrated that language ability differs significantly as a function of SES. The study revealed that the average vocabulary of 3-year-old children from professional families were found to be two times as large as the vocabulary of 3-year-old children from families on welfare (Hackman & Farah, 2009). Another study found that with kindergarteners, middle-SES individuals performed better than their low-SES counterparts, particularly on tests of the left perisylvian/ language system, a system of the brain that encompasses semantic, syntactic, and phonological aspects of language. SES-related differences are evident in communication skills of children that emerge even

before speech production (Hackman & Farah, 2009; Demir and Kuntay, 2014; Rowe and Meadow, 2009). Skills such as gestures and eye gazing have been found to establish a strong base for children's later language development and growth (Demir and Kuntay, 2014). In a study by Rowe and Meadow (2009), the authors videotaped 50 children from families of varying SES, interacting with parents at the age of 14 months and later assessed their vocabulary skills at the age of 54 months. Results of the study revealed that children from high-SES had larger vocabulary at 54 months than those of low SES counterparts. According to the study, the large vocabulary was found to be explained by children's gesture use at 14 months. Children from high SES families were found to frequently use gesture to communicate at 14 months, an action explained by parent gesture use. In another study by Fernald, Marchman, and Weisleder (2013), SES-related differences were found on language processing skills and vocabulary of 18- and 24-month-olds. Language processing skills such as lexical processing efficiency was measured by how quickly the children turned toward the correct picture of a target object, once hearing the object's name (Fernald et al., 2013). The study found that children from high SES families looked at the target object sooner and more accurately than those from low SES families. In summary, children from high SES families have consistently demonstrated on average, larger vocabulary sizes. Previous research has related these differences to the quality and quantity of speech that children receive from their parents (Hackman & Farah, 2009; Duncan and Magnuson, 2012; Fernald et al., 2013). Rowe and Meadown (2009) suggested that evidently, high SES parents "talk more, use more diverse vocabulary, and use more complex syntax" with their children than parents of low SES. Further, the SES gap in vocabulary begins as early as pre-speech production ages, widens

through toddlerhood until about the age of four, and often remains constant throughout the school years (Rowe and Meadow, 2009). Additionally, Romeo et al. (2018) that it is quality, and not just the quantity that is important for early language development, which goes beyond the influence of SES. Similarly, Hart & Risley (2003)'s study showed that children's early language experiences were not as impactful in terms of numbers of words but instead, the quality of the words heard by the children. Hart & Risley (2003) also emphasizes the importance of quality of vocabulary in early years, which happens within the home and their families, prior to their experiences in social groups and with peers outside of the home. These findings have serious implications, as it has been established that vocabulary is a strong predictor of later linguistic skills, cognitive abilities, academic achievement (Romeo et al., 2003) and is a primary reason as to why many low SES children enter school at a higher risk for academic failure or misplacement in special education and remedial programs.

Culture and Language Development

Little research has been conducted that investigates the involvement of developmental psychology theories in language difference and cognitive test performance of African American children. Traditionally, the theoretical framework that has guided research in this area has been grounded in comparative studies of a difference versus deficit perspective (Harris & Shroeder, 2013) by looking at the "norm". A developmental theory that could aid in assisting with this gap in knowledge of language differences is the Vygotsky's sociocultural theory (Vygotsky, 1978; Harris & Shroeder, 2013). Crucial in this theory, is the recognition of the role that culture plays in cognition, education, and language development of a child. In Vygotsky's Zone of Proximal Development, or ZPD,

it is theorized that children develop cognitive skills and learn cultural and psychological tools while engaged in cooperative interactions with adults and their environments (Harris & Shroeder, 2013; Mutekwe, 2018). To this end, it is important to consider that children participate in language activities in their homes, churches, and neighborhoods, and those language activities are grounded in cultural rituals and traditions maintained by specific cultural socialization practices (Harris & Schoeder, 2013). Given the theory's strong emphasis on sociocultural roots of cognition, this framework is particularly crucial to consider when distinguishing between a deficit and a difference in cognitive test performance.

Bilingualism and Bidialectalism

Language is undoubtedly in many aspects considered a socially constructed behavior influenced by exposure, identity, and other environmental influences (Rickford et al., 2015). Sociolinguistics have spent decades studying and researching ways to best define language use in African American communities across the United States. What is most salient about the research on the topic is the vital implications of sociocultural and historical contexts that produce language variety (Lanehart, 2015).

Although there is no agreed-upon definition of "bilingualism" by researchers (Butler & Hakuta, 2004), bilingualism is often broadly defined as "the ability to utilize two languages to varying degrees across different contexts" (Lee-James & Washington, 2018). This definition should therefore be inclusive of use of different language varieties and bidialectalism. However, the research has shown that in educational and clinical practices, bidialectalism is often defined as low-status versions of mainstream, or the nonstandard language system of low SES communities (Lanehart, 2015; Lee-James &

Washington, 2018).

For over four decades, disputes have existed among clinicians, scholars, and educators, regarding the legitimacy of dialects or language varieties of American English, such as African American Vernacular English (AAVE) as a language system (Harris & Schoeder, 2013). The first recognition of AAVE was noted through The Ann Harbor case in 1979, which addressed the influence of AAVE on the academic achievement of African American school children (Harris & Schoeder, 2013). The case specifically addressed the ongoing controversial theory of language deficit vs. language difference and ruled against AAVE being a language deficit. For the first time in history, the language difference theory was considered and asserted that language variation in AAVE speakers did not imply deficiency in language or cognitive skills. These differences were instead attributed to factors such as parenting, learning and language environments of low SES children (Harris & Schoeder, 2013). The debate was later revisited in 1996 through a resolution from the Oakland Unified School District in California, also known the *Ebonics Resolution* (Long, 2019). The resolution was the first to recognize AAVE as a legitimate "genetically based language system" (Long, 2019).

Little is left to dispute in regard to the existence of a linguistic system used by many African American individuals (Lanehart, 2015). Central to this idea, is the prevalence in use of African American Vernacular English (AAVE) and other language varieties of Standard American English (SAE) in low SES African American communities across the United States today (Lanehart, 2015). Although regional differences have been difficult to account for, greater importance has been given to grammatical features of AAVE, as well as features of other language varieties that are

unique to African American speakers (Charity, 2007). What we know, is that similar to bilingual speakers, children who speak AAVE or other language varieties of SAE have historically faced linguistic prejudice and have consistently underperformed on cognitive assessments compared to White, middle-class children (Harris & Schroeder, 2013).

SES, Language, and Norm Sample Stratification

Considering what we know about the implications of SES and language on child development, using norms yielded from current stratification procedures present as a highly inappropriate practice. Noteworthy, studies that have examined child outcome differences between African American and White children often attribute SES as the primary factor to these differences, without considering language differences. Contrastive analyses often used in assessment also compare culturally and linguistically diverse children with mainstream American English speakers (Cormier, McGrew, & Ysseldyke, 2014). This has significant implications on cognitive test performance of African American children, especially when accounting for the historical, social and political roots of culture and language in African American communities.

A central issue in the validity of assessments, is test fairness. Messick (2000) recognizes "test fairness" to be a social value not just a psychometric issue, that is free of prejudice and bias. Test fairness has been defined as having the ability to yield comparably valid scores from individual to individual and all subgroups including those from low SES backgrounds (Willingham, 1998; Messick, 2000; Ortiz, 2018). The Standards for Educational and Psychological Testing (2014) addresses the term "fair" to be characterized as "the absence of bias and as equitable treatment of all examinees in the testing process". Poe and Elliot (2019) cites different views on "test fairness", which

includes the elimination of bias, pursuit of validity, acknowledgement of social context, a legal responsibility and an ethical obligation.

Despite varying views on the concept of test fairness, most are attributed to the concept of construct validity (Messick, 2000). Test developers have longed to produce valid and fair assessments through attempting to control for factors such as SES and race. Despite indications from research, that race as single factor does not prove to impact test performance, race and SES are often found to be similarly represented in norm samples. A prominent issue with race and SES as a single representation, is that it likely echoes data from the U.S Census Bureau, which finds African Americans to have the highest poverty rate (Shrider, Kollar, Chen, & Semega, 2021).

As previously established, racial differences in SES are strong predictors of racial disparities in cognitive test performance (Harris & Schroeder, 2013) and educational outcomes among African Americans and Whites (Williams, Priest, & Anderson, 2019). Another factor that appears to violate the assumption of fairness in assessments, is the issue of language. Unlike many assessments, the Ortiz Picture Vocabulary Assessment Test (Ortiz PVAT; Ortiz, 2018) directly addressed the issue of language within the norm sample stratification. While the Ortiz PVAT was intentionally created to address language differences in English Language Leaners, it generally controlled for any effect that language differences may have on one's receptive vocabulary skills (Ortiz, 2018). Moreover, the Ortiz PVAT Technical Manual (Ortiz, 2018) reports that Ortiz controlled for SES in a manner that included sampling individuals of varying SES backgrounds, without specific norms for SES levels. Ortiz (2018) analyzed the impact of SES on performance above and beyond controlling for language and other stratification variables.

The results of his analysis indicated that although small, the difference in the performance between groups revealed that SES had a stronger effect on the performance of low SES individuals (Ortiz, 2018). These results support the literature that reveals that although low SES groups are often included in the norm sample, the norm sample is likely representative of middle-class SES (based on the average of all SES groups). This form of representation is likely to add to the disproportionality of low SES and African American students in special education and remedial programs. Thus, the literature continues to support the need for increased attention in the areas of development, administration, interpretation and use of standardized assessments with African American students.

Present Study

The present study will attempt to answer several questions regarding the effect of socioeconomic status (SES) and language differences on cognitive test performance of African American children when compared to normative performance on the WISC-V. It has been established that SES has a significant impact on child development and cognitive test performance. Therefore, test creators have included SES as a stratified variable that is highly controlled for in normative samples of cognitive assessments. Although, low SES groups are included in the norm sample, the norm sample is likely representative of middle-class SES, based on the average of all SES groups. Therefore, this form of representation is likely to yield inequitable and possible discriminatory results when assessing low SES students, especially when the variable of language is not being controlled for. Thus, the study will examine specific cognitive test performance, while separating the pathway of verbal versus nonverbal tasks, to demonstrate that the

effect of SES and language difference is not evenly distributed. If we continue to use normative samples that do not take into account the proper use of SES, as well as language differences, we will continue to produce unfair, possibly discriminatory results that will continue to lead to already existing issues, such as the overrepresentation of African American children in special education programs.

If the claims made by the author are presumed to be demonstrably factual, the study will demonstrate the lack of fairness of normative samples currently being utilized in standardization of cognitive assessments such as the WISC-V, specifically as it pertains to variables such as SES and language. Particularly, the study will attempt to answer the following questions: 1) When assessing African American children with the WISC-V, are normative samples utilized for comparison, leading to inequitable and potentially discriminatory results? 2) Are normative samples utilized in standardized tests properly accounting for differences in SES? and 3) Does language impact cognitive performance of African American children, despite being historically considered monolinguals?

The present study will attempt to demonstrate that we are not accounting for differences in SES properly and that normative samples are not representative of low SES individuals. The study will intentionally include a population sample of only low SES children, as they are the ones likely to be at risk for discriminatory outcomes. Utilizing this specific population will allow for measurement of potential differences in performance between low SES children and those represented in the normative sample. If differences in test performance are found to be statistically significant, these results

would indicate problematic issues in the stratification of SES, particularly as it relates to equity and fairness in cognitive assessments such as the WISC-V.

Research has demonstrated mixed views on language differences between African American and White children. Given the assumption that stratification of language is not necessary in development of normative samples, test performance between the sample group and the norm group should yield similar results. To further support the author's claims, the present study will attempt to examine the extent to which potential language differences in African American students of low SES backgrounds, play a role in their performance on cognitive tests such as the WISC-V. This study also proposed that the issue of language should be highly considered when interpreting test results, especially when accounting for the historical, social and political roots of language in African American communities. It is hypothesized that language differences found in this study will be statistically significant and show a medium to large effect size in terms of performance difference between the sample group and the norm. As a result, language will be a potential variable that should not be overlooked by test developers.

Method

Procedures

Data collected for this study was obtained directly from evaluations conducted independently by this researcher. Data was strictly collected for the purpose of completing this study. Recruitment of participants commenced after approval was sought and obtained from St. John's University's Institutional Review Board and from the Mount Vernon City School District. The author distributed information regarding the study and inform teachers about the study procedures, ensuring to minimize participants' time out from instructional time to complete the assessment. The author conducted the assessments with participants, using selected subtests from the Wechsler Intelligence Scales for Children, Fifth Edition (WISC-V). Prior to initiating any assessment, the author obtained parental consent from the parent or legal guardian of the participants, as all participants were minors and school aged children. Verbal student assent was also obtained, to assure that participants agreed to complete the assessment and were free to stop at any point if they did not wish to continue. As part of the recruitment efforts, all participants received an incentive of their choice from a prize bin provided by the researcher, for their participation in the study.

The study examined test performance of students across different areas of cognitive functioning. All assessment sessions took place between January 2022 and January 2023. All sessions took place during after-school hours and were conducted on an individual basis whereas the participants were assessed in a quiet, separate location which was free of visual and auditory distractions.

Participants

The participants of the study included 79 African American students, strictly

including public school students in grades Kindergarten to fifth grade, who attend the Mount Vernon City School District in Westchester County, NY. Participants were recruited and selected based on meeting the race criteria of Black or African American, as well as the criteria of being *flagged* as "Poverty-from low- income family". This information was obtained from the student portal on *Infinite Campus*, an online database of Student Information System, used by the Mount Vernon City School District. Participants in the study were non-disabled students who did not have an Individual Educational Program (IEP) or a Section 504 Accommodation Plan. All participants were general education students who spoke no other language at home, aside from "English", according to Infinite Campus.

Based on the most recent enrollment data reported by the New York State Education Department (NYSED), roughly 7,326 students are enrolled in grades Kindergarten through 12th in the Mount Vernon City School District. Approximately 4,951 of these students, or 68% are African American students, and approximately 5,343 or 73% of the students are considered to be of low socioeconomic status (also characterized as "economically disadvantaged"; New York State Education Department [NYSED], 2020). As defined by New York State Education Department (NYSED), "economically disadvantaged" students are those who "participate in, or whose family participates in, economic assistance programs, such as the free or reduced-price lunch programs, Social Security Insurance (SSI), Food Stamps, Foster Care, Refugee Assistance (cash or medical assistance), Earned Income Tax Credit (EITC), Home Energy Assistance Program (HEAP), Safety Net Assistance (SNA), Bureau of Indian Affairs (BIA), or Family Assistance: Temporary Assistance for Needy Families

(TANF)."

Instruments

WISC-V. Participants were administered six subtests from the Wechsler Intelligence Scales for Children, Fifth Edition (WISC-V). The WISC-V is an individually administered, comprehensive clinical instrument for assessing cognitive functioning of children ages 6:0 through 16:11 (Wechsler, 2014). Normative information from the WISC-V was based on data collection which took place from April 2013 through March 2014 (Wechsler, 2014). The normative sample included 2,200 children and was stratified to represent consensus data based on demographic information such as age, sex, race/ethnicity, parent education level, and geographic region (Wechsler, 2014). Representativeness of the sample revealed that only 13% of these children were considered to be African American.

The WISC-V assesses intellectual functioning in five cognitive areas: Verbal Comprehension Index (VCI), Visual Spatial Index (VSI), Fluid Reasoning Index (FRI), Working Memory Index (WMI), and the Processing Speed Index (PSI). The primary subtest scores produce a Full-Scale IQ (FSIQ) composite score, which represents general intellectual ability (Wechsler, 2014). The primary index scores and the FSIQ are on a standard score metric with a mean of 100 and a standard deviation of 15. The subtests pertaining to each index is on a scaled score metric with a mean of 10 and a standard deviation of 3 (Wechsler, 2014).

For the purpose of this study, students were administered three indices of the WISC-V, which were comprised of a total of six administered subtests. Data was obtained from the Verbal Comprehension Index (VCI), using the Similarities (SI) and

Vocabulary (VC) subtests; the Visual Spatial Index (VSI), using Block Design (BD) and Visual Puzzles (VP); the Fluid Reasoning Index (FRI), using Matrix Reasoning (MR) and Figure Weights (FW); and the General Ability Index (GAI). The General Ability Index (GAI; Wechsler, 2014), is a measure of global intellectual functioning for use with the WISC-V. The GAI consists of a total of five subtests from the verbal comprehension, visual spatial, and fluid reasoning indices (Wechsler, 2014). These include the WISC-V BD, WISC-V SI, WISC-V MR, WISC-V VC, and WISC-V FW. The Standard Error of Measurement (SEM) on the WISC-V differs by the individual measure (e.g., VCI - 4.2, VC - 5.4, SI - 5.5). For the purpose of this study, the GAI score was utilized in place of a FSIQ score, to represent general cognitive performance. Previous research has strongly supported the use of the GAI as a good predictor for the FSIQ of the WISC-V (Weiss, Saklofske, Holdnack, & Prifitera, 2016). Moreover, what the literature suggests is that the GAI has a higher loading of "g", or general intellectual functioning, as it is composed of subtests which are more interrelated; and has a higher correlation with the FSIQ, when compared to the working memory (WMI) and processing speed (PSI) indices (Scott, 2006). A study by Scott (2006) found that the GAI and FSIQ scores correlate positively ($r = .96$, $p < .01$), and found no significant difference ($t(30) = 1.0$ $p < .01$) between the scores of the GAI and the FSIQ. Conceptually, the GAI is an ancillary index that provides an estimate of general intelligence that is less impacted by working memory and processing speed, relative to the FSIQ (Wechsler, 2014). Generally, the WISC-V is designed to be administered individually utilizing a paper and pencil format or using an iPad. The only format used during this study was paper and pencil.

Results

Descriptive statistics (means, standard deviations, maximum, and minimum) were calculated for standard scores and scaled scores obtained by participants on the WISC-V VCI, WISC-V VC/SI subtests, WISC-V FRI, WISC-V MR/FW, WISC-V VSI, WISC-V BD/VP, and WISC-V GAI (see Table 1). The average age of participants in the sample group was 8.34 years (see Table 1).

A one-sample t-test was used to analyze whether there was a statistically significant difference between the mean performance of the sample group when compared to the normative mean performance on the Gc, Gf, Gv clinical composites of the WISC-V (VCI, FRI, VSI) and mean performance on the GAI. Furthermore, a one-sample t-test was also used to analyze whether there was a statistically significant difference between the mean performance of the sample group when compared to the normative mean performance on individual subtests. A coefficient alpha level of $p < .05$ was used to control for potential Type I errors. As indicated by previous research, socioeconomic status and race are stratified variables, tightly controlled for in the norm sample of the WISC-V. Additionally, as proposed by previous studies, language does not need to be controlled for because there is no proposed language difference in the African American population when compared to the norm sample. Assuming the sample group are monolingual and non-disabled students, both groups should therefore, perform evenly and have a mean of 100. To calculate the difference or the effect size between the composites and the normative mean, Cohen's *d* was used to determine whether the effect was large enough to be considered meaningful or significant. A pairwise analysis was used to analyze whether there was a statistically significant difference between the mean

performance of the sample group across the three measured indices (WISC-V VCI and WISC-V FRI, WISC- V VCI and WISC-V VSI, and lastly, WISC-V- FRI and WISC-V VSI). Cohen's d was also used to determine whether the effect size was large enough to considered significant between the performance across the composites measured (WISC-V VCI and WISC-V FRI, WISC-V VCI and WISC-V VSI, and lastly, WISC-V- FRI and WISC-V VSI) for the sample group. According to Cohen (1988), d=0.2 is considered a small effect size, while d=0.5 is a medium effect size, and d=0.8, a large effect size.

Table 1

Descriptive Statistics of the Sample Group

n = 79	Minimum	Maximum	Mean	SD
Grade	0	5	3.23	1.44
Age	6	11	8.34	1.46
WISC-V VCI (SS)	76	92	83.96	3.48
WISC-V Vocabulary (Ss)	4	8	6.46	0.98
WISC-V Similarities (Ss)	6	9	7.63	0.74
WISC-V FRI (SS)	85	100	92.78	2.78
WISC-V Matrix Reasoning (Ss)	7	10	8.89	0.72
WISC-V Figure Weights (Ss)	7	10	8.71	0.79
WISC-V VSI (SS)	92	105	98.35	2.67
WISC-V Block Design (Ss)	8	10	9.13	0.56
WISC-V Visual Puzzles (Ss)	8	13	10.35	0.88
WISC-V GAI (SS)	81	92	87.23	2.66

Note. SD = Standard Deviation; SS = Standard Score; Ss = Scaled Score

Hypothesis 1

Socioeconomic status is one of the most extensively researched variables in the study of cognitive performance. Thus, test developers include SES as a stratified variable that is highly controlled for in normative samples of cognitive assessments. While low SES groups are reportedly included in the norm sample, the norm sample is likely representative of middle-class SES, as it is based on the average of all SES groups. Thus, if SES is truly properly controlled for, then general performance (WISC-V GAI) of the sample group (low SES subjects) should be comparable to the overall normative mean of 100 and should therefore, also be comparable across composites and subtests in general. However, the author proposed a different hypothesis, which is that performance across composites and the subtests would be generally lower when compared to the normative means. To examine if SES would affect cognitive performance differential, the study measured if there was any significant difference from the mean for the GAI, the composites and for the subtests. A one sample *t*-test with an alpha level of .05 was conducted using the normative mean standard score of 100 for all cases against the individual performance of the sample group across general performance and composites (WISC-V VCI, WISC-V VSI, and WISC-V FRI). In accord with the author's hypothesis, the *t*-test yielded statistically significant differences between the obtained mean scores between the WISC-V VCI and the normative mean (t = -40.99, df = 78, p < .001*, d = -1.09), the WISC-V FRI and the normative mean (t = -23.06, df = 78, p < .001*, d = -0.49), and the WISC-V VSI and the normative mean (t = -5.48, df = 78, p < .001*, d = -0.11). This analysis also indicated a statistically significant difference between the WISC-V GAI and the normative mean (t = -42.75, df = 78, p < .001*, d = -0.87).

A summary of this analysis can be found in Table 2.

Conversely, noted differences were observed in effect sizes expressed by Cohen's d during this analysis. Supporting the author's hypothesis, a large difference (-0.87) was noted between performance on the WISC-V GAI when compared to the norm. An exceptionally large difference was found between the WISC-V VCI and the normative mean, demonstrated by the obtained effect size of -1.09, and a moderate effect size (-0.49) was found between the WISC-V FRI and the normative mean. In contrast, the effect size found from the comparison of the WISC-V VSI and the normative mean suggests merely a small difference (-0.11). The observed differences and effect sizes were further evaluated and supported by the noted differences between the indices and normative means. The mean performance of the sample group on the WISC-V GAI and WISC-V VCI were 12.77 and 16.04 points lower than the normative mean, respectively. These differences support the large to exceptionally large effect sizes which were found. Similarly, a moderate effect size of -0.49 between the WISC-V FRI and the normative mean is supported by the mean difference of 7.22 points. However, the mean performance of the sample group on the WISC-V VSI was only 1.65 points lower than the normative mean, supporting the small effect size of -0.11. A visual summary of these results are presented in *Figure 1* below.

To further assess performance differences between the sample group and the normative mean group, a one-sample *t*-test with an alpha level of .05 was conducted using the normative mean scaled score of 10 for all cases against the individual performance of the sample group across subtests pertaining to the indices and overall, GAI (WISC-V VC, SI, MR, FW, BD, VP). In accord with the author's hypothesis, the

analysis also suggested statistically significant differences between the obtained mean scores between the WISC-V VC and the normative mean (t = -31.99, df = 78, p < .001*), the WISC-V SI and the normative mean (t = -28.55, df = 78, p < .001*), the WISC-V MR and the normative mean (t = -13.83, df = 78, p < .001*), the WISC-V FW and the normative mean (t = -14.58, df = 78, p < .001*), and the WISC-V BD and the normative mean (t = -13.79, df = 78, p < .001*). Of note, however, is that the analysis did not indicate a significant difference between the sample group's performance on the WISC-V VP when compared to the normative mean (t=3.59, df=78, p =0.999). A summary of this analysis can be found in Table 3. Overall, this analysis supports that subtests scores pertaining to the WISC-V VCI, which are VC and SI were far below the normative mean performance. Additionally, performance on the WISC-V VP and the normative mean sample, did not indicate a significant difference and supports the effect size found between the GAI and the normative mean, as the WISC-V VP does not pertain to the overall GAI. Please refer to *Figure 2* for a visual presentation of these results.

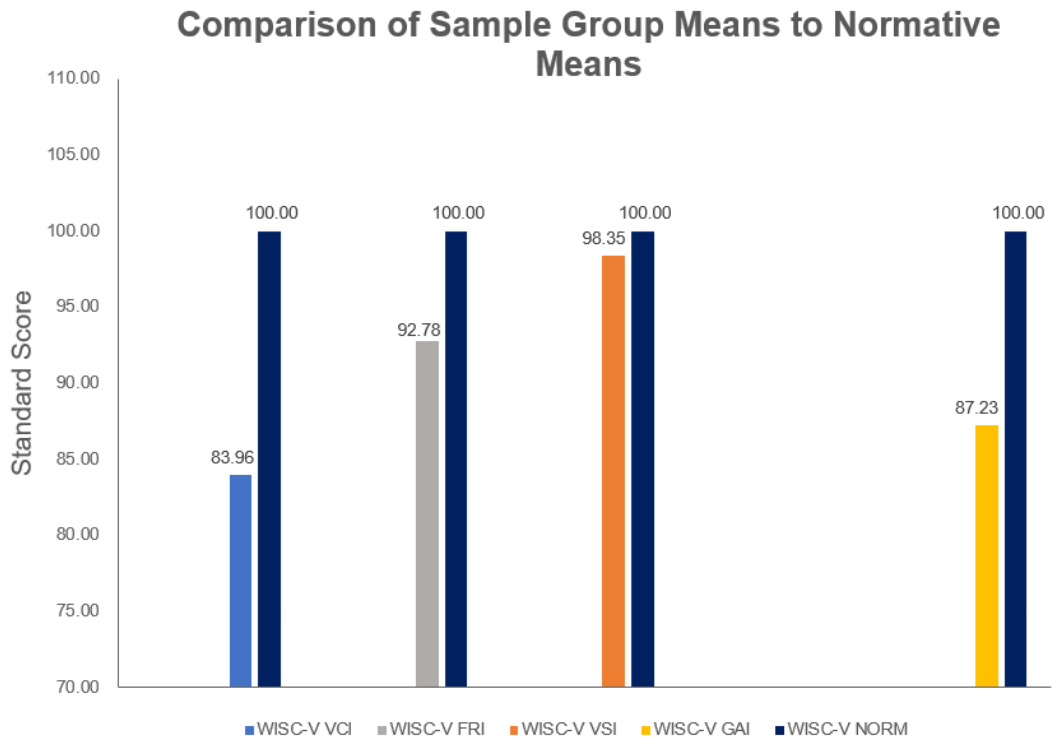

Figure 1. Mean differences in Standard Scores between the sample group index performance and the WISC-V normative mean of 100.

Table 2

Comparison of WISC-V Index Means to Normative Means

	Mean	Diff.	SD	*t*	*p*	*d*
WISC-V GAI	87.23	12.77	2.65	-42.75	< 0.00	-0.87
WISC-V VCI	83.96	16.04	3.48	-40.99	< 0.00	-1.09
WISC-V FRI	92.78	7.22	2.78	-23.06	< 0.00	-0.49
WISC-V VSI	98.35	1.65	2.67	-5.48	< 0.00	-0.11

Note. Hypothesized mean of 100 set for all indices. GAI = General Ability Index; VCI = Verbal Comprehension Index; FRI = Fluid Reasoning Index; VSI = Visual Spatial Index; Mean = Mean Performance; Diff = Mean Difference; SD = Standard Deviation; *t* = *t*-value; *p* = *p*-value; *d* = Cohen's *d*. **p* significant at the .05 level

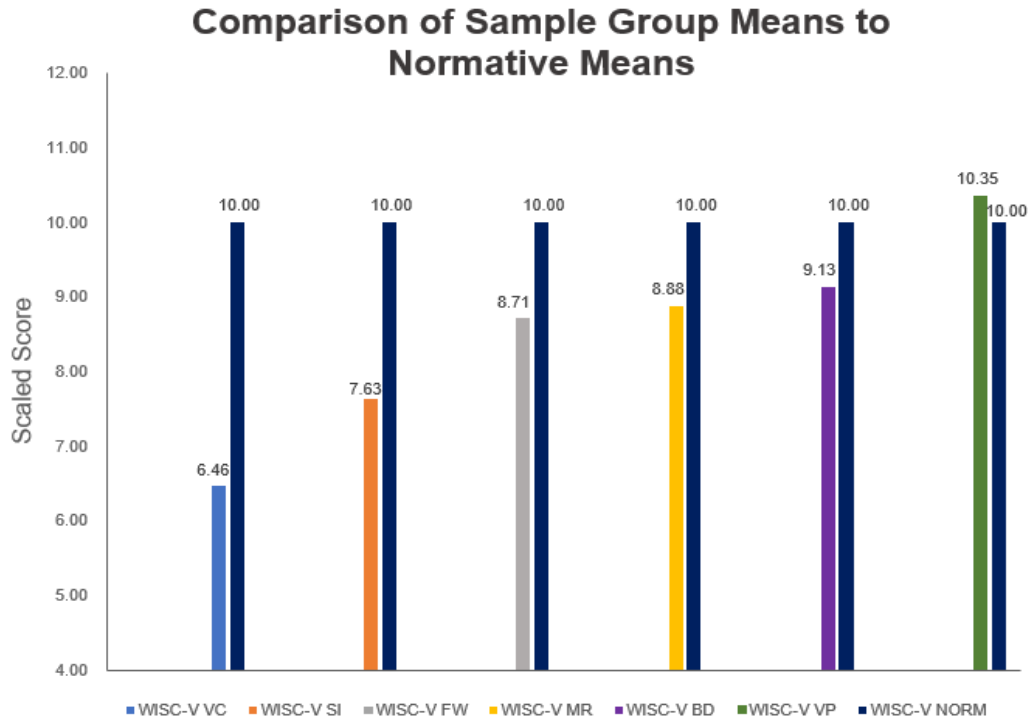

Figure 2. Mean differences in Scaled Scores between the sample group individual subtest performance and the WISC-V normative mean of 10.

Table 3

Comparison of WISC-V Subtest Means to Normative Means

	Mean	Diff.	SD	*t*	*p*
WISC-V VC	6.46	3.54	0.98	-31.99	< 0.00
WISC-V SI	7.63	2.37	0.74	-28.55	< 0.00
WISC-V MR	8.88	1.12	0.72	-13.83	< 0.00
WISC-V FW	8.71	1.29	0.79	-14.58	< 0.00
WISC-V BD	9.13	0.87	0.56	-13.79	< 0.00
WISC-V VP	10.35	0.35	0.88	3.59	0.99

Note. Hypothesized mean of 10 set for all subtests. VC = Vocabulary subtest; SI = Similarities subtest; MR = Matrix Reasoning; FW = Figure Weights; BD = Block Design;

VP = Visual Puzzles; Mean = Mean Performance; Diff = Mean Difference; SD = Standard Deviation; *t* = *t*-value; *p* = *p*-value; **p* significant at the .05 level

Hypothesis 2

Given the assumption that stratification of language, above and beyond SES, has been deemed unnecessary in development of normative samples, test performance between the two groups should yield similar results. However, the author proposes that when accounting for the historical, social, and political roots of language development in African American communities, language differences found in this study will be statistically significant and will show a medium to large effect size in terms of performance difference between the two groups. Thus, language differences will affect cognitive performance differential. Specifically, the performance of the sample group on the WISC-V VCI was predicted by the author to be lower when compared to the sample group performance on both the WISC-V VSI and WISC-V FRI, and that these differences would be significant. The analysis between the VCI and VSI corresponds to a comparison between mental processing of visual versus verbal information (Wechsler, 2014). Additionally, distinct from the FRI, the VSI does not explicitly evaluate aspects of conceptual reasoning, which consists of a semantic/linguistic factor (Wechsler, 2014). Thus, the author further hypothesized that due to the higher semantic and linguistic demands in the FRI, compared to the VSI, FRI performance of the sample group would be lower than the performance on the VSI. To test these hypotheses, a pairwise analysis was conducted to assess differences in performance between the index scores (WISC-V VCI, WISC-V VSI, and WISC-V FRI). With these results, the mean index scores were arranged, in order to assess performance rank. If performance was noted to be lower on composites of increasingly higher levels of linguistic demands, it would reinforce the

proposed language effect. See Table 4 for a summary of the analysis.

The results indicated a mean difference of -8.82 between the WISC-V VCI and WISC-V FRI obtained scores. This suggests that this difference is statistically significant (t = -17.64, df = 78, p < .001*, d = -2.80). A mean difference of -14.39 was found between the WISC-V VCI and WISC-V VSI suggesting that this difference was also noted to be statistically significant (t = -31.79, df = 78, p < .001*, d = -4.64). The analysis for both comparisons demonstrated exceptionally large effect sizes as estimated by Cohen's d, of –2.80 and -4.64 between the WISC-V VCI and FRI and WISC-V VCI and VSI, respectively. To further support the author's hypothesis that language differences would affect cognitive performance differential, a statistically significant difference was found between performance on the WISC-V FRI and VSI (t = -14.25, df = 78, p < .001*, d = -2.04). The mean difference between the two indices was -5.57. The measured effect size of -2.04 as expressed by Cohen's d suggests that the difference in performance was also found to be exceptionally large. Overall, the analysis supports the author's prediction that while all index scores were found to be below the normative mean, those which consisted of higher semantic and linguistic demands were far lower than those with less verbal information. Specifically, scores on the WISC-V VCI when compared to other indices, demonstrated differences that were significant enough to reinforce the proposed language effect consistent with the author's hypothesis. *Figure 3* presents a visual representation of the aforementioned results.

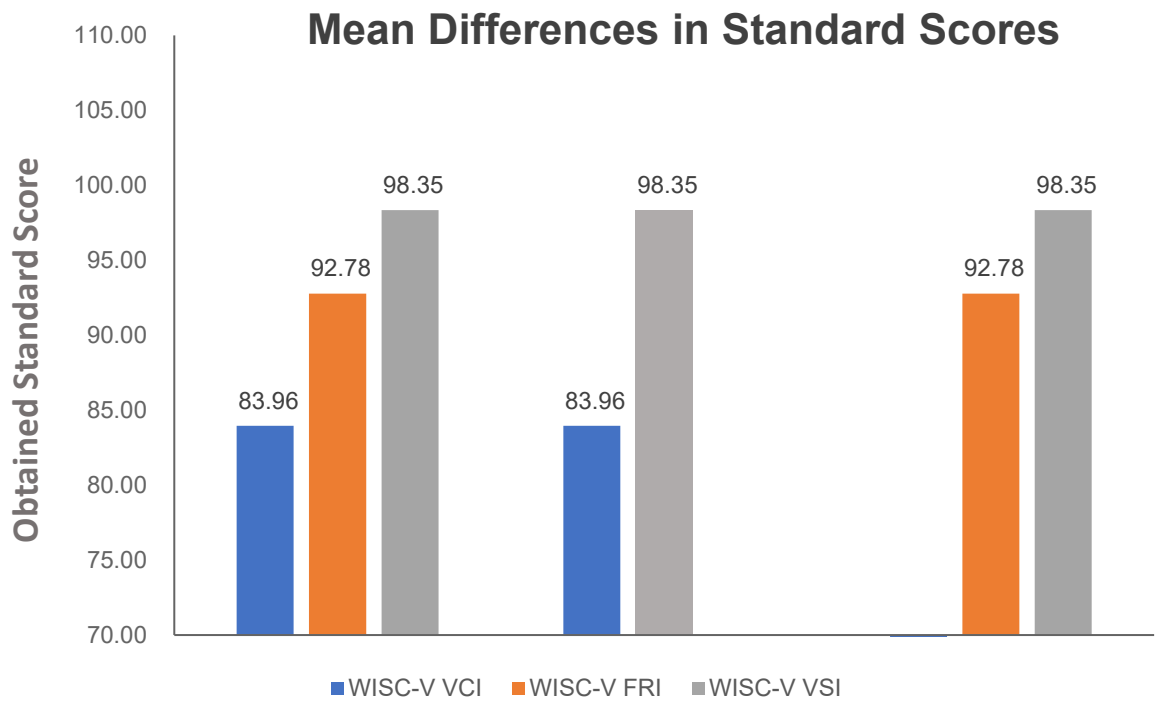

Figure 3. Mean differences in Standard Scores between the WISC-V VCI and FRI, WISC-V VCI and VSI, and between the WISC-V FRI and VSI.

Table 4

Pairwise Analysis Comparison of Standard Scores

	Mean	Diff.	*pooled* SD	*t*	*p*	*d*
WISC-V VCI	83.96					
vs. WISC-V FRI	92.78	-8.82	3.48	-17.64	< 0.00	-2.80
vs. WISC-V VSI	98.35	-14.39	3.09	-31.79	< 0.00	-4.64
WISC-V FRI	92.78					
vs. WISC-V VSI	98.35	-5.57	2.73	-14.25	< 0.00	-2.04

Note. VCI = Verbal Comprehension Index; FRI = Fluid Reasoning Index; VSI = Visual Spatial Index; Mean = Mean Performance; Diff = Mean Difference; pooled SD = Pooled Standard Deviation; *t* = *t*-value; *p* = *p*-value; *d* = Cohen's *d*. **p* significant at the .05 level

Discussion

The current study examined the extent to which potential language differences in African American students of low SES backgrounds, play a role in cognitive test performance. Specifically, the study investigated the influence of language on performance of African American children when compared to normative performance on the WISC-V, a common and widely used cognitive assessment tool. Test creators claim to have included SES as a stratified variable that is highly controlled for in normative samples of cognitive assessments. However, although low SES groups may be included in the norm sample, the norm sample is likely representative of middle-class SES, based on the average of all SES groups. Therefore, the author of this study claimed that this form of representation is likely to yield inequitable and possible discriminatory results when assessing low SES students, especially when the variable of language is not being controlled for. The present study examined specific cognitive test performance of a sample group, while separating the pathway of verbal versus nonverbal tasks, to demonstrate that the effect of SES and language difference is not evenly distributed.

Prior research has established significantly high correlations between socioeconomic status and race/ethnicity. Historically, indicators of SES are evidently patterned with race, with African American children generally being the highest percentage of children living in poverty. The implication of childhood SES to brain development lies in its evident relationship to cognitive ability as measured by cognitive assessments, particularly in the area of language. The participants of the study included African American students of low socioeconomic status (also characterized as "economically disadvantaged".) Participants were general education students whose

home language was determined to be *English*. No other language was reported to be spoken at home. Although there is no agreed-upon definition of *bilingualism* by researchers, the author emphasized the idea that language varieties of English should be highly considered as the ability to utilize language may vary across different contexts. Thus, the sample used in this study was unique in that it took into consideration the prevalence and vital implications of sociocultural and historical contexts that produce language variety among African American individuals. The study proposed that the issue of language should be highly considered when assessing cognitive test performance, especially when accounting for the historical, social and political roots of language development in African American communities.

Effects of SES and Language Differences

Given the assumption that stratification of language, above and beyond SES, has been deemed unnecessary in development of normative samples, test performance between the two groups should have yield similar results. To examine if SES and language would affect cognitive performance differential, the study measured if there was any significant difference from the mean for the GAI, the composites and for the subtests when compared to the normative sample of the WISC-V. Interestingly, the results did support the author's hypothesis that it would affect cognitive performance differential. Findings revealed statistically significant differences in performance between the sample group and the normative sample, which will be discussed below.

The study findings revealed, as predicted, a large difference between the general performance (WISC-V GAI) of the sample group when compared to the norm, which was reflected in the effect size. Additionally, an exceptionally large difference was found

between the mean performance of the sample group on verbal tasks (WISC-V VCI) and the normative mean. A medium difference was also found between the WISC-V FRI and the normative mean. In contrast, only a small difference was found between the sample group and the norm performance on the WISC-V VSI. In general, the results of the study supports the author's prediction that while all index scores were found to be below the normative mean, those which consisted of higher semantic and linguistic demands were far lower than those with less verbal information. Specifically, scores on the WISC-V VCI demonstrated differences that were significant enough to reinforce the proposed language effect consistent with the author's hypotheses.

To better understand the language effect, the author analyzed whether there was a statistically significant difference between the mean performance of the sample group when compared to the normative mean performance on individual subtests. Considering that the WISC-V does not account for low SES and language variations, scores on individual subtests were expected to be significantly lower than the norm, specifically on those with higher verbal demands. Consistent with previous findings, the author predicted that performance on verbal subtests such as WISC-V VC and SI would be the lowest, as vocabulary has been shown to be the trait of language most susceptible to the effects of SES (Hoff, 2013). Supporting the author's prediction, individual subtest analysis revealed statistically significant differences between the obtained mean scores between the WISC-V VC and the normative mean and the WISC-V SI and the normative mean. These differences were found to be the most significant. However, statistically significant differences were also found between the WISC-V MR and the normative mean, the WISC-V FW and the normative mean, and the WISC-V BD and the normative mean,

respectively. Of note however, the analysis did not indicate a significant difference between the sample group's performance on the WISC-V VP when compared to the normative mean. Interestingly, the WISC-V VP is the only subtest that did not contribute to the WISC-V GAI overall score and was also found to be generally the highest score of the sample group.

The author was also interested in further evaluating the proposed language differential by analyzing the performance of the sample group across different composites. The mean index scores were arranged, in order to assess performance rank. Performance of the sample group on the WISC-V VCI was predicted by the author to be lower when compared to the performance on both the WISC-V VSI and WISC-V FRI, and that these differences would be significant. This was predicted as the analysis between the VCI and VSI corresponds to a comparison between mental processing of visual versus verbal information (Wechsler, 2014), which supports the proposed language differential. Additionally, distinct from the FRI, the VSI does not explicitly evaluate aspects of conceptual reasoning, which consists of a semantic/linguistic factor (Wechsler, 2014). The results suggested statistically significant differences between the WISC-V VCI and WISC-V FRI, and between the WISC-V VCI and WISC-V VSI. Furthermore, a statistically significant difference was also found between performance on the WISC-V FRI and VSI. Overall, the analysis supports the author's prediction that while all index scores were found to be below the normative mean, those which consisted of higher semantic and linguistic demands were far lower than those with less verbal information. Scores on the WISC-V VCI when compared to other indices, demonstrated differences that were significant enough to reinforce the proposed language effect consistent with the

author's hypothesis.

Effects of Norming Implications

These findings have serious implications as it demonstrates the lack of fairness of normative samples currently being utilized in standardization of cognitive assessments such as the WISC-V, specifically as it pertains to variables such as SES and language. This suggests that if we continue to use normative samples that do not take into account the proper use of SES, as well as language differences, we will continue to produce unfair, possibly discriminatory results that will continue to lead to already existing issues, such as the overrepresentation of African American children in special education programs. Undoubtedly, based on the results of this study, language differences and variations in English speakers is in fact a potential variable that should not be overlooked by test developers. Within a broader aspect, the study addresses study addresses the ongoing controversial theory of language deficit vs. language difference. Consistent with earlier studies, differences in performance, specifically as it pertains to higher verbal demands, are likely attributed to factors such as limitations on the development of low SES children including language skills.

Overall, these results support that the GAI, similar to the FSIQ, is a poor measure of cognitive ability with African American children due to the heavy reliance on language related subtests. However, it is critical that while evaluating culturally and linguistically diverse children, that we do not eliminate assessment of "language" as it is the most important (highest *g*-weighted ability) portion of the evaluation. Language is a moderating factor above and beyond SES, and such skills are strong predictors of cognitive abilities and academic achievement as suggested by Romeo et al. (2003).

Eliminating language altogether, is therefore not a good solution. Clinicians can account for the differential of language on test performance by using effective tools such as the Culture-Language Interpretive Matrix (CLIM; Flanagan, Ortiz, & Alfonso, 2013). The C-LIM was developed to assist clinicians in determining validity of cognitive tests administered to culturally and linguistically diverse, which should be interpreted as reflecting such language differences and/or cultural influences (Flanagan et al., 2013).

Limitations of the Study

As with any study, results from the present study should be interpreted in light of and in the context of its limitations. While the preliminary supporting findings of this study could be utilized to make significant contributions to current assessment methods of African American children of low socioeconomic status, a few limitations of the study should be noted. First, the results reported in the study included merely a small sample group of 79 students versus the WISC-V normative sample of 2,200 children utilized for comparison. Additionally, the study did not control for the effects of age. The WISC-V normative sample included assessing the cognitive functioning of children ages 6:0 through 16:11. The sample group utilized in this study included students ages 6:0 through 11:0, precluding the ability to determine if age is a factor in cognitive performance due to possible extended exposure to mainstream English. Furthermore, although the present study was inclusive of both female and male students, it did not control for gender differences in the results, which further limits its generalizability. Future research should explore and consider the differences in effects when controlling for age and gender, while also using a larger sample.

Moreover, identifying the unique relationship between cognitive performance and SES and race can be challenging due to confounding variables in societies, such as which the present study was conducted. As with most previous studies, African American children are often directly correlated with low socioeconomic status when compared to other SES groups. Future research should include group control for SES that contains African American students from high SES backgrounds and compare their performance to those of low SES backgrounds. Additionally, the present study was limited to research

in one single area/ school district wherein students are predominantly African American children from low SES. Future studies can and should replicate analyses of the present study with similar groups but across different geographic regions to evaluate whether the results remain the same.

While the study provided encouraging findings and an extension of knowledge on lack of assessment fairness, another caveat of the study should be noted. Evidence from previous study is robust as it pertains to positive effects of high-quality early childhood programs in low SES children's cognitive development. Future research should focus on examining the extent, if any, to which participation in early childhood programs impact cognitive performance of low SES African American children on the WISC-V during school age, specifically in the area of language. In addition, the study also did not control for parent educational attainment. This is an important factor when considering parental education, and socioeconomic status as predictors of language development, school readiness, and educational outcomes (Mistry et al., 2010). Replicating analysis from the current study while separating parental education attainment and socioeconomic status may yield meaningful results in future studies.

Implications for School Psychologists

Based on the current study findings, implications will be discussed in terms for which school psychologists may increase best practices when selecting assessment tools that may be more appropriate with African American and other culturally and linguistically diverse children. Considering what we know about the implications of SES and language on child development, using norms yielded from current stratification procedures present as a highly inappropriate practice. Noteworthy, studies that have examined child outcome differences between African American and White children often attribute SES as the primary factor to these differences, without considering language differences.

There are serious implications that come with comparing culturally and linguistically diverse children with mainstream American English speakers on cognitive tests. Particularly, cognitive assessments are often utilized to determine special education eligibility and programming outcomes (Scheiber, 2016). Widely used are Wechsler tests such as the WISC-V. A central issue in the validity of such assessments is test fairness. Messick (2000) recognizes "test fairness" to be a social value not just a psychometric issue, that is free of prejudice and bias. Test fairness has been defined as having the ability to yield comparably valid scores from individual to individual and all subgroups including those from low SES backgrounds (Willingham, 1998; Messick, 2000; Ortiz, 2018). Test developers have longed to produce fair assessments through attempting to control for factors such as SES and race. Despite indications from research, that race as single factor does not prove to impact test performance, race and SES are often found to be similarly represented in norm samples. Another factor that appears to violate the

assumption of fairness in assessments, is the issue of language. Results of the current study support previous arguments that suggest that low SES, culturally diverse children may have unique verbal strengths that may not be reflected or captured through the use of cognitive assessments such as the WISC-V.

The present study supports evidence for use of other assessment tools with culturally diverse/ ethnic minority children. The literature continues to support the need for increased attention in the areas of development, administration, interpretation and use of standardized assessments with such students. Particularly, examination of test bias continues to be at the forefront of debates on the use of cognitive assessments. The results of the current study are extremely valuable given the frequent use of assessments such as the WISC-V as the sole method of making life-changing decisions for the educational future of minority children (Scheiber, 2015). The use of other assessments that take into consideration cultural and linguistic differences of ethnic minority children is highly recommended.

Unlike many assessments, the Ortiz Picture Vocabulary Assessment Test (Ortiz PVAT; Ortiz, 2018) directly addressed the issue of language within the norm sample stratification. While the Ortiz PVAT was intentionally created to address language differences in English Language Leaners, it generally controlled for any effect that language differences may have on one's receptive vocabulary skills (Ortiz, 2018). Moreover, SES stratification in norm samples is likely not useful as it does not control for differences that may appear to be more language related. The Ortiz PVAT Technical Manual (Ortiz, 2018) reports that Ortiz controlled for SES in a manner that included sampling individuals of varying SES backgrounds, without specific norms for SES levels.

It analyzed the impact of SES on performance above and beyond controlling for language and other stratification variables. The results of the analysis indicated that the difference in the performance between groups revealed that SES had a stronger effect on the performance of low SES individuals (Ortiz, 2018). However, it also found that SES remained important only when examining extremes, such as differences between low and high SES. Nonetheless, the difference was still relatively small ($d = 0.10 – 0.18$). Yet, when language was removed completely from the stratification by ensuring that all subjects are monolingual English speakers, differences in race and ethnicity disappear completely. This suggests that use of tests with African American students is valid only when they are monolingual English speakers to the third generation. In other cases, test score validity must be established by use of analysis such as the C-LIM before test results can be interpreted and reported. When such analysis does not take place, interpretation of test results is likely to add to the disproportionality of low SES and African American children in special education and remedial programs. Additionally, another measure of assessment of ethnic minority groups that aimed to create test fairness, is the Kaufman Assessment Battery for Children–Second Edition (Kaufman et al., 2005). The test authors of the KABC-II intentionally created different global indexes that excluded the heavy reliance on subtests that assess crystallized knowledge and that requires verbal expression skills (Scheiber, 2015). Similar to the results of the present study, the authors of the KABC-II recognize the implication that verbal subtests are significantly more biased than its nonverbal components when assessing children from ethnic minorities (Scheiber, 2015); and endorse the use of less biased and more neutral tools in the assessment of minority children. The use of less biased assessments is widely considered best practices,

and are based on the notion that racial differences in SES are strong predictors of racial disparities in cognitive test performance among African Americans and Whites (Williams, Priest, & Anderson, 2019), especially those that emphasize verbal knowledge without taking into consideration the historical, social and political roots of their ethnic or linguistic backgrounds.

References

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). AERA, APA, & NCME. *Standards for educational and psychological testing. Washington, DC: American Educational Research Association*.

American Psychological Association (2021). Socioeconomic Status. Retrieved October 15, 2021, from https://www.apa.org/topics/socioeconomic-status

Butler, Y. G., & Hakuta, K. (2004). Bilingualism and second language acquisition. *The handbook of bilingualism*, 114-144.

Charity, A. H. (2007). Regional differences in low SES African American children's speech in the school setting. Language Variation and Change, 19(3), 281-293.

Cohen, J. (1988). Statistical Power Analysis for the Behavioral Sciences, 2nd ed. Hillsdale, New Jersey: Erlbaum.

Cormier, D. C., Bulut, O., McGrew, K. S., & Kennedy, K. (2022). Linguistic influences on cognitive test performance: Examinee characteristics are more important than test characteristics. *Journal of Intelligence*, *10*(1), 8.

Cormier, D. C., McGrew, K.S., & Ysseldyke, J. E. (2014). The influence of linguistic demand and cultural loading on cognitive test scores. *Journal of Psychoeducational Assessment, 32*(7), 610-623.

Demir, Ö. E., & Küntay, A. C. (2014). Cognitive and neural mechanisms underlying socioeconomic gradients in language development: New answers to old questions. *Child Development Perspectives*, *8*(2), 113-118.

Duncan, G. J., & Magnuson, K. (2012). Socioeconomic status and cognitive functioning:

moving from correlation to causation. *Wiley Interdisciplinary Reviews: Cognitive Science*, *3*(3), 377-386.

Fernald, A., Marchman, V. A., & Weisleder, A. (2013). SES differences in language processing skill and vocabulary are evident at 18 months. *Developmental science*, *16*(2), 234–248. https://doi.org/10.1111/desc.12019

Flanagan, D. P., Ortiz, S. O., & Alfonso, V. C. (2013). Cross-battery assessment of individuals from culturally and linguistically diverse backgrounds. In A. S. Kaufman & N. L. Kaufman (Series Ed.), Essentials of cross-battery assessment (3rd ed., pp. 287–350). Hoboken, NJ: Wiley.

Hackman, D. A., & Farah, M. J. (2009). Socioeconomic status and the developing brain. Trends in Cognitive Science, 13(2), 65–73. doi: 10.1016/j.tics.2008.11.003.

Harris, Y. R., & Schroeder, V. M. (2013). Language Deficits or Differences: What We Know about African American Vernacular English in the 21st Century. *International Education Studies*, *6*(4), 194-204.

Hart, B., & Risley, T. R. (2003). The early catastrophe: The 30 million word gap by age 3. *American educator*, *27*(1), 4-9.

Kaufman, A. S., Lichtenberger, E. O., Fletcher-Janzen, E., & Kaufman, N. L. (2005). *Essentials of KABC-II assessment* (Vol. 95). John Wiley & Sons.

Lanehart, S. (Ed.). (2015). The Oxford Handbook of African American Language. Oxford University Press.

Lee-James, R., & Washington, J. A. (2018). Language Skills of Bidialectal and Bilingual Children. *Topics in Language Disorders*, *38*(1), 5-26.

Long, M. (2019). Ebonics, language, and power. In *Race and Ethnic Conflict* (pp. 331-

345). Routledge.

Messick, S. (2000). Consequences of test interpretation and use: The fusion of validity and values in psychological assessment. In *Problems and solutions in human assessment* (pp. 3-20). Springer, Boston, MA.

Mistry, R. S., Benner, A. D., Biesanz, J. C., Clark, S. L., & Howes, C. (2010). Family and social risk, and parental investments during the early childhood years as predictors of low-income children's school readiness outcomes. *Early childhood research quarterly*, *25*(4), 432-449.

Mutekwe, E. (2018). Using a Vygotskian sociocultural approach to pedagogy: Insights from some teachers in South Africa. Journal of Education (University of KwaZulu-Natal), (71), 58-72.

NYSED. "District and Student Data." New York State Education Department, 30 Jun. 2020.

Ortiz, S.O. (2018). *Ortiz PVAT Technical Manual*. Toronto, Canada: Multi-Health Systems.

Poe, M., & Elliot, N. (2019). Evidence of fairness: Twenty-five years of research in assessing writing. *Assessing Writing*, *42*, 100418.

Pieterse, A. L., Chung, S., Khan, T., & Bissram, J. (2013). Social class and racial and ethnic identity: Perspectives on the psychological impact of social stratification and inequality. *The Oxford handbook of social class in counseling*, *379*.

Rickford, John R et al. “Neighborhood effects on use of African-American Vernacular English.” Proceedings of the National Academy of Sciences of the United States of America vol. 112,38 (2015): 11817-22. doi:10.1073/pnas.1500176112

Romeo, R. R., Leonard, J. A., Robinson, S. T., West, M. R., Mackey, A. P., Rowe, M. L., & Gabrieli, J. D. (2018). Beyond the 30-million-word gap: Children's conversational exposure is associated with language-related brain function. *Psychological science*, *29*(5), 700-710.

Rowe, M. L., & Goldin-Meadow, S. (2009). Differences in early gesture explain SES disparities in child vocabulary size at school entry. *Science*, *323*(5916), 951-953.

Scheiber, C. (2016). Is the Cattell–Horn–Carroll-based factor structure of the Wechsler Intelligence Scale for Children—Fifth Edition (WISC-V) construct invariant for a representative sample of African–American, Hispanic, and Caucasian male and female students ages 6 to 16 years?. *Journal of Pediatric Neuropsychology*, *2*, 79-88.

Scheiber, C., & Kaufman, A. S. (2015). Which of the three KABC-II global scores is the least biased?. *Journal of Pediatric Neuropsychology*, *1*, 21-35.

Scott, K. A. (2006). *Is the GAI a good short form of the WISC-IV?*(Order No. 1434505). Available from ProQuest Dissertations & Theses Global. (304935629). https://jerome.stjohns.edu/login?url=https://www.proquest.com/dissertations-theses/is-gai-good-short-form-wisc-iv/docview/304935629/se-2?accountid=14068

Shrider, E. A., Kollar, M., Chen, F., & Semega, J. (2021). Income and Poverty in the United States: 2020. *US Census Bureau, Current Population Reports P60-273. US Government Publishing Office, Washington, DC, September. https://www. census.gov/content/dam/Census/library/publications/2021/demo/p60-273. pdf*.

Wechsler, D. (2014). *WISC-V: Technical and interpretive manual*. NCS Pearson, Incorporated.

Weiss, L. G., Saklofske, D. H., Holdnack, J. A., & Prifitera, A. (2016). WISC-V: Advances in the assessment of intelligence. *San Diego, CA*.

Williams, D. R., Priest, N., & Anderson, N. (2019). Understanding associations between race, socioeconomic status, and health: patterns and prospects. In *The Social Medicine Reader, Volume II, Third Edition* (pp. 258-267). Duke University Press.

Willingham, W. W. (1998). A systemic view of test validity. In S. Messick (Ed.), Assessment in higher education (213-242). Mahwah, NJ: Erlbaum

VITA

Name:	*Fernanda Carvalho*
Baccalaureate Degree:	*Bachelor of Science/Arts, Hunter College, City University of New York, New York, Major: Psychology*
Date Graduated:	*May, 2012*
Other Degrees and Certificates*	*Master of Arts, IONA University, New York, Major: School Psychology*
Date Graduated:	*May, 2015*