Masters Theses, 2020-current   The Graduate School

5-12-2022

Writing While Black: African American Vernacular English (AAVE) and Perceived Writing Performance

Jaylin Nesbitt

Writing While Black: African American Vernacular English Use and Perceived Writing

Performance

Jaylin Nicole Nesbitt

A thesis submitted to the Graduate Faculty of

JAMES MADISON UNIVERSITY

In

Partial Fulfillment of the Requirements

for the degree of

Master of Arts

Department of Graduate Psychology

May 2022

FACULTY COMMITTEE:

Committee Chair: Deborah Bandalos, PhD

Committee Members/ Readers:

Jennifer Randall, PhD

Sara Finney, PhD

Dena Pastor, PhD

Table of Contents

List of Tables

List of Figures

Abstract

In the education system, there have historically been inequities that have severely disadvantaged Black students academically. One area in which these inequities surface is on writing assessments in the form of lower scores. I argue that because the U.S. education system is centered around Standard American English (SAE), it disadvantages those from different linguistic backgrounds, specifically Black students, as they are most likely to be speakers of African American Vernacular English (AAVE). Although there are theoretical justifications for this, past literature has not empirically tied inequities on writing assessments to Black students' use of AAVE. The current study used Natural Language Processing (NLP) to quantify students' use of grammatical features of AAVE (e.g., multiple negation, completive done) on the written component of the Indiana Statewide Testing for Educational Progress-Plus (N=21,420) to investigate how students' use of AAVE impacted writing scores. The relationships between race, gender, economic disadvantage, AAVE use, and writing scores were examined using a path model analysis in M*plus* Version 8.4. Based on the results of the path model, it was concluded that those who identified as males and those who were economically disadvantaged had higher frequencies of AAVE use and lower scores on writing in comparison to those who identified as females and those who were not economically disadvantaged. Contrary to the hypothesized negative effect, as AAVE increased writing score increased (b=.05, p <.001), but this parameter was small in magnitude and therefore ignorable. It was concluded that Black students received lower writing scores than white students, but based on the findings of the study, the score differences could not be attributed to Black students' higher use of AAVE. Explanations for these unanticipated findings are explored and include assessment

directions that may have limited the use of AAVE, concerns about the procedure used to measure AAVE, and assumptions made by the statistical model employed. Future research should continue to investigate the linguistic discrimination Black students experience as it relates to their use of AAVE and other factors mediating the relationship between race and writing, due to the presence of written components on consequential tests (SAT, GRE).

Chapter 1: Introduction

Language is used to communicate ideas, facts, or knowledge, and can be

considered a cultural marker of a specific ethnic and/or racial group (Kellogg, 2016). For

example, ethnic and/or racial groups share language grounded in similar cultural

experiences and use language to express experiences unique to them. Standard American

English (SAE) is the dominant language in the U.S., because it is the language used by

white Americans, the dominant group in U.S. society (Inoue, 2015; Thomas, 2018).

Schools, workplaces, and testing practices are centered around SAE (Baker-Bell, 2020b;

Ball, 1996; Inoue, 2015; Randall, 2021; Randall, et al., 2021; Wheeler, 2012), which

reinforces the message that this language is neutral, academic, professional, and superior

to other forms of English (Baker-Bell, 2020a; Mitri & Terry, 2014; Wheeler, 2019). In

U.S. society, power, prestige, and educational opportunity are associated with whiteness.

Because race and language are interconnected (Baker-Bell, 2020a), language has evolved

as one way those in power maintain their dominance. That is, power, prestige, and

educational opportunities are dominated by SAE speakers, and this has resulted in the

dependence of these privileges on SAE proficiency. Because SAE usage is more intuitive

for SAE speakers, that group maintains an advantage in pursuing educational and

employment opportunities that simultaneously disadvantages others.

Within the U.S education system, the centering of SAE, or the centering of

whiteness, disadvantages students of color who have linguistic backgrounds other than

SAE (Randall, 2021; Inoue, 2015). More specifically, most Black[1] students use African

American Vernacular English (AAVE) (Lippi-Green, 1998) -- a language with unique

grammatical, syntactical, pronunciation, and rhetorical patterns (Ivy & Masterson, 2011;

Johnson, 2013; Mitri et al., 2004; Thompson et al., 2004). Because this is a language

other than SAE, it is devalued in schools (Baker-Bell, 2020b; 2020a; Ball, 1996; Ball &

Lardner, 1998; Harris & Schroeder, 2013; Charity et al., 2004; Sealey-Ruiz, 2005;

Thomas, 2018; Wheeler, 2012;2019). It is devalued in schools by marking the use of

AAVE as wrong or discouraging students from using AAVE in academic spaces. This

disadvantages Black students who are speakers of AAVE and illustrates the manifestation

of racial hierarchies in society as linguistic hierarchies in the classroom.

        The treatment of language is a reflection of how specific groups of people are

disadvantaged in U.S. society due to their use of non-SAE languages such as AAVE

(Baker-Bell, 2020a; Mitri & Terry, 2014; Rashid, 2011; Thomas, 2018; Wheeler,2019).

Baker-Bell (2020a) stated that race and language are interconnected, because the

devaluation of AAVE is a direct reflection of the devaluation of Black lives in the United

States. Black peoples' lives are seen as inferior, unimportant, unintelligent, and unhuman.

Their use of AAVE is treated in the same way (Baker-Bell, 2020a; Inoue, 2015) and

penalization of AAVE (Ball, 1996; Johnson & VanBrackle, 2012; Richardson,1997;

---

[1] I capitalized Black throughout this thesis to encompass the myriad of cultures and ethnicities that make up the Black community (those of African descent, Caribbean descent, etc.). Black represents a shared identity and community, while white does not hold the same meaning, hence the reason it is lower cased. I avoid capitalizing white as a form of resistance against the culture of white supremacists (Coleman,2020; Laws,2020). This was a decision I made to prioritize Black people, who's person and knowledge have historically been devalued in this country. I encourage you to look at the following websites that validated this decision: https://www.cjr.org/analysis/capital-b-black-styleguide.php and https://www.nytimes.com/2020/07/05/insider/capitalized-black.html

Wheeler, 2012) can be seen as an attempt to erase Black experiences, culture, and students in the classroom. So, when the experiences of Black students are discussed in the classroom, their racial experience and linguistic experience cannot be separated.

Critical Race Theory (CRT) is the theoretical framework used in this study to investigate the manifestation of racial hierarchy in society (racism) through the preference given to SAE (i.e., whiteness) and the devaluation of AAVE (i.e., Blackness). Bell (1992) defined CRT as the acknowledgement of racism as an everyday aspect of life that can inform our understanding of how racism is manifested across numerous domains, including the education system. One example of such racism is the preferencing of SAE that arises in the form of *Anti-Black Linguistic Racism* in classroom practices used to teach language. *Anti-Black Linguistic Racism* is defined by Baker-Bell as "the linguistic violence, persecution, dehumanization, and marginalization that Black Language-speakers [AAVE] experience in schools and everyday life" (2020a, p. 11).

Previous work has highlighted that students of color (i.e., Black students) are the students who linguistically deviate from the language preferred and used in academia (Inoue, 2015). The racial hierarchy that places white people and SAE at the top of the hierarchy is directly reflected in the treatment of other languages (i.e., AAVE) in the classroom (Delgado & Stefancic, 2017). When students use languages other than SAE in the classroom, they are labeled as incorrect, insufficient, and unintelligent. This places AAVE speakers at an inherent disadvantage in the education system and puts them at odds with their academic environment. The stigma associated with non-SAE languages (i.e., AAVE), in the education system has drastic consequences for students of color, specifically for Black students (Baker-Bell, 2020b; Ball & Lardner, 1997; Mitri & Terry,

2014; Randall, et al., 2021; Sealey-Ruiz, 2005; Thomas, 2018). For example, teachers often do not acknowledge the linguistic background of Black students and instead treat their language expression as deficient. Additionally, if AAVE is acknowledged, students are encouraged to use AAVE as a bridge to learn and use SAE—a concept referred to as code-switching (Baker-Bell, 2020a).

CRT is also relevant to understanding the role that racism plays in the classroom and in the assessment of classroom performance (Ladson-Billings, 1998). Specifically, researchers have used CRT to examine how educational institutions inherently perpetuate the subordination of any group that does not identify as white (Rashid, 2011). For example, research has found that Black students are differentially penalized for their use of AAVE on assessments of literacy (Ball, 1996; Johnson & VanBrackle, 2012; Richardson,1997; Wheeler, 2012; 2019)

Although literacy is assessed through both reading and writing, the focus of this thesis is on writing assessment. Because writing is a necessary component for success on consequential admissions tests (SAT, GRE), classroom tests, and large-scale educational assessments, the educational success of students who are not proficient in SAE is adversely impacted. SAE can be seen as a form of gate-keeping that operates through tests and assessments to advantage SAE speakers and disadvantage those who deviate from SAE. Thus, instead of writing assessments serving as a steppingstone from one life stage (i.e., high school graduation) to another (i.e., college admission), they serve as a barrier for students of color. While writing assessments may propel one group (i.e., white students) into higher education, they may prevent other groups from getting as far.

An example of these adverse impacts for users of AAVE would be teachers' practice of marking AAVE use as wrong on writing assessments (Richardson, 1997; Wheeler, 2012; 2019) even though it is a valid way to display knowledge. This practice not only places Black students at an inherent disadvantage in the classroom overall, but also results in disadvantages in the form of lower scores on writing assessments in comparison to their white counterparts (Ball, 1996; Cronley et al., 2017; Johnson & VanBrackle, 2012; Richardson, 1997). Additionally, the centering and preferential treatment of SAE in the classroom paints the picture that AAVE is wrong, inferior, and an incorrect way to demonstrate proficiency in writing.

Why do Black students receive lower scores in comparison to white students on writing assessments? I argue that this is because the construct of writing as currently measured in educational assessments does not represent the experiences, cultural values, and linguistic backgrounds of people of color. This is because most K-12 writing assessments require the use of SAE (Inoue, 2015; Randall, 2021), a language that does not reflect the linguistic background of most Black students. The current rubrics used to score writing assessments (e.g., English Placement Exam (EPT); Inoue, 2015), were designed to measure the writing proficiency of SAE speakers and do not allow for the ways in which students from other language backgrounds demonstrate the construct of writing. The assumption of SAE fluency results in the differential penalization of culture-specific words, phrases, and grammatical structures.

The differential penalization mentioned previously has been investigated in past research studies. Richardson (1997) investigated the relationship between AAVE usage and writing scores and concluded that essays with no evidence of AAVE use received

higher scores than those with AAVE use. Mitri and Terry (2014) found that achievement measures that targeted reading and vocabulary skills resulted in significantly different scores between students with low and high levels of AAVE use. Students with high levels of AAVE usage scored lower on reading and vocabulary achievement measures in comparison to students with low levels of AAVE use. Additionally, Johnson and VanBrackle (2012) found that teachers had different perceptions of errors depending on whether errors were related to usage of SAE, English as a second language (ESL), or AAVE. Even when essays had the same number of errors, essays with errors related to AAVE were more likely to receive failing grades than those with errors related to ESL or SAE.

Because current methods used to score writing assessments preference SAE usage, they are unfair for students who are not SAE speakers, such as AAVE speakers and multilingual students. According to *The Standards for Educational and Psychological Testing*, a fair assessment represents the same construct across diverse groups of people, is free of bias, does not advantage one group over another, and yields valid interpretations for its intended uses for all groups (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 2014). All of these requirements are violated in writing assessments that are not committed to an anti-racist framework in which whiteness is decentered (Inoue, 2020; Randall, 2021; Randall et al., 2021). I provide examples of how current writing assessments violate these requirements in the paragraphs that follow.

Current assessments of writing only encompass the language (SAE) and experiences of white students (Inoue, 2015; Randall; 2021). In the language of validity theory, the

construct of writing is underrepresented for Black students because their writing is only

deemed correct when it aligns with SAE and is deemed incorrect when it has features of

AAVE. According to *The Standards*, construct underrepresentation is a specific threat to

measurement validity in which "a test fails to capture important aspects of the construct.

It implies a narrowed meaning of test scores because the test does not adequately sample

some types of content, engage some psychological processes, or elicit some ways of

responding that are encompassed by the intended construct" (AERA, APA, NCME, 2014,

p. 12). Because the construct of writing is underrepresented for Black students, it does

not yield a valid interpretation of writing proficiency for these students. Additionally

writing assessments are unfair because the preferencing of SAE use advantages students

with language backgrounds aligning with SAE while simultaneously disadvantaging

students who are AAVE speakers.

The lack of fairness and construct underrepresentation of writing assessments have

negative implications for Black students. For example, the EPT (English Placement

Exam) is an entrance exam with a writing component for students enrolling in Fresno

State University in California. Results on this exam can lead to students being classified

as in need of remediation of their English fluency and having to take additional classes

before enrolling. Students of color were disproportionally classified into remedial classes

in comparison to their white counterparts (Inoue, 2015). Additionally, low scores on the

written components of high stakes college admissions assessments such as the ACT,

SAT, and GRE, adversely impact entrance into undergraduate institutions and graduate

school programs for Black students. For example, in 2015, 61% of Black students' scores

did not meet the benchmarks on the ACT college readiness test, which includes writing

assessments (Bridges, n.d.). Although there is no research specifically demonstrating that these low writing scores are due to students' use of AAVE or lack of familiarity with SAE, it seems likely this may be the case given that academic assessments require SAE use. These low scores, in conjunction with being Black, make the college acceptance process more stressful and difficult than necessary.

Purpose of Current Study

Given the literature up to this point, the final question I want to address is why this study is important and necessary. AAVE has been fighting for its rightful recognition in the education system since the 1979 Ann Arbor, Michigan Black English Court case. This court case concluded that because teachers did not acknowledge AAVE as the language of their students, the curriculum was not tailored in a way that facilitated Black students' learning and ability to thrive academically. Additionally, the negative attitudes of the teachers in the school were shown to have a negative impact on Black students and their ability to learn. Without the acknowledgement of the linguistic backgrounds of Black students, the pedagogical practices in the classroom were classified by the court as inadequate and ineffective (Ball & Lardner, 1998). The current study is necessary because the lack of acknowledgement of AAVE and the Black experience continues to run rampant in the U.S. education system in the current assessment of writing (Baker-Bell, 2020a;2020b; Inoue, 2015; Randall, 2021).

As mentioned previously, there are major differences between the AAVE and SAE as they are two completely different languages belonging to different linguistic groups. In writing specifically, these differences refer to style, grammar, spelling, and the inclusion of culturally specific words (Richardson, 1997). Students can be penalized for

these differences based on the traditional and inherently racist grading practices of the

education system. When assessments do not factor in syntactic, phonological,

and morphological differences inherent in students' use of AAVE, students can

potentially be mis-assessed. This refusal to acknowledge AAVE as a rule-governed

language has contributed to the development of negative attitudes in educators which

directly influence the educational well-being of Black students and play a huge role in

their futures (Ball & Lardner, 1998). More specifically, Black students tend to internalize

the negative attitudes towards AAVE and label it "bad, wrong, or incorrect" (Baker-Bell,

2020b; Sealey-Ruiz, 2005). Black students that experience this internalization show

outcomes such as silence or disengagement in the classroom (Thomas, 2018).

Additionally, not acknowledging, or penalizing AAVE serves as a form of cultural

rejection and reinforces the notion that AAVE does not have a place in academia,

regardless of its relevance to Black students.

  Additionally, I recognize that CRT has an activist dimension committed to not

only acknowledging inequalities in society but changing it by challenging societal power

structures and questioning their legitimacy for adequately representing **all** people

(Delgado & Stefancic, 2017). Not only is this study a valuable contribution to research on

Black students in the education system, but it is also my activism and commitment to

change the system to better represent Black students by providing a potential answer to

why these inequities exist.

  The goal of such research is to ensure that Black students have an equal chance of

being scored in an equitable manner even when they are not proficient in SAE or do not

feel comfortable using SAE. Given the evidence provided thus far, it seems likely that

students' use of AAVE is one of the reasons Black students tend to receive lower scores

on writing assessments. Because of the violence perpetuated by writing assessments

towards people of color, it is imperative that these inequities are investigated. I plan to

challenge these inequities in the writing scores of Black students; specifically, regarding

the lack of acknowledgement of the potential language use differences that may surface

during writing tasks but continue to go unacknowledged on the grading rubrics.

Positionality

I am also conducting this study because I am a Black woman who went through a

predominantly white education system, where my experiences felt as though they were

other. I felt like I would never belong. I did not read books that mirrored my experience,

and my friends and teachers did not talk how I talked with my family at home. I distanced

myself from my Blackness because I associated it with something negative, something I

was scared, and frankly embarrassed to represent. I wonder if you know what that does to

a child who does not like who they are because of what society perceives them to be. It

was not until I was 18 years old and was accepted into my Alma Mater, Spelman College,

that I felt proud to be who I am. I saw so many people who looked like me, talked like

me, and mirrored what I've been through. The structure of the K-12 education system

down to the faculty, curriculum, student body, and language, stunted my identity

development for 18 years. That being said, I am committed to being the voice I did not

have when I was in school, and I want to fight for the Black children that are continuing

to endure rejection in the education system for simply being who they are.

This thesis was therapeutic for me to write as it validated so much of my

experiences, past and present. It helped me hold onto my Blackness and who I am as I

returned to a predominantly white space after my time at Spelman. Spelman equipped me

with the confidence I needed to center my thesis around Blackness, in a space where

whiteness is centered. Claiming my Blackness out loud is something I owe to my

elementary, middle, and high school self.

Model Description and Justifications

 The specific relationships between the exogenous and endogenous variables are

discussed in the remainder of this section. Moving from the left to the right side of the

model (Figure 1), I discuss the hypothesized exogenous variable covariances, direct

effects on AAVE, direct effects on writing scores, and conclude with the hypothesized

indirect effects on writing scores.

Figure 1
Hypothesized Path Model
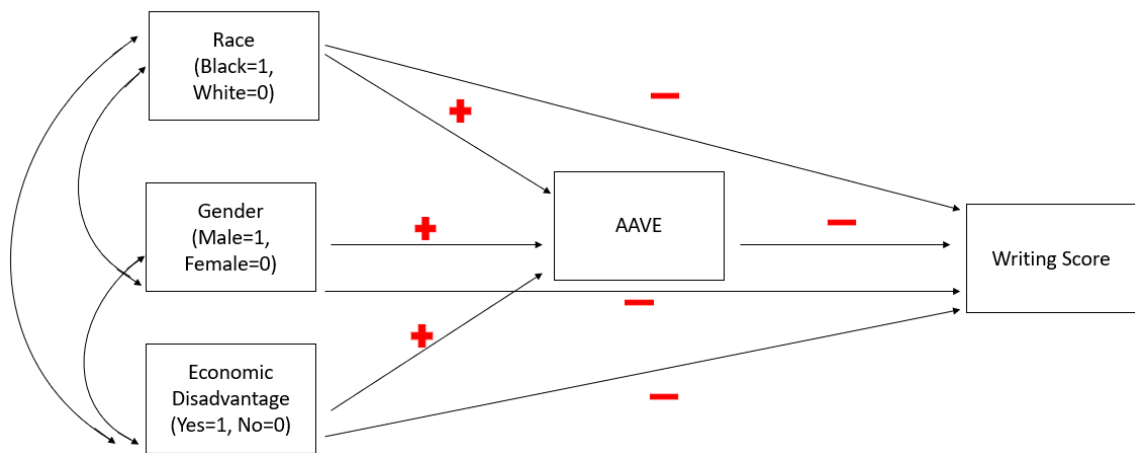
Covariances among Race, Gender, and Economic Disadvantage

 On the left-hand side of the model, I have included the following three exogenous

variables: race, gender, and economic disadvantage. I included these specific variables in

the model as research has linked race, gender, and economic disadvantage to AAVE use (Charity et al., 2004; Lippi-Green, 1997; Fitton et al., 2012; Wainwright, 2018) and writing performance (Cronley et al.; Kim et al., 2001; Mitri & Terry, 2014; Randall, 2021; Richardson, 1997).

Race and economic disadvantage covary as previous research found that a higher proportion of Black participants are of lower SES (Lippi-Green, 1997; Sautter et al., 2012). Although theoretically, race and gender are not correlated, Black boys and girls are disproportionately represented in the current data. Unequal proportions of girls and boys would result in a lack of independence and induce covariation among the two variables. Because of this, the covariance between race and gender is statistically justified. For the same reason, the covariance between gender and economic disadvantage is statistically justified because the proportions of girls and boys who are economically disadvantaged were not equal in this dataset.

Direct effects of Exogenous Variables on AAVE Use

About 80% of the Black community identify as AAVE speakers (Dorn, 2019; Lippi-Green, 1997). Based on this, I hypothesized a relationship between race and AAVE use in which Black students would be more likely to use AAVE than white students. Additionally, Wainwright (2018) studied the relationship between gender and the use of AAVE in narrative production and found that those who identified as men produced more features of AAVE in story retelling (M=3.13) in comparison to those who identified as women (M=2.15). In the current study, it is hypothesized that the relationship between gender and AAVE will be positive. Lippi-Green (1997) acknowledged that although AAVE is used by a variety of Black people regardless of

SES, when it comes to the grammatical features of the language, it is more commonly seen in poorer Black communities. This finding is supported by Fitton et al. (2012) who conducted a study that aimed to predict reading achievement of 207 children from the density of AAVE used in a written narrative sample. Density was defined as the total number of AAVE word usages divided by the total number of words in the essay. The researchers found that as self-reported household income increased, the use of AAVE in writing samples decreased. With these findings in mind, I hypothesized a positive relationship between economic disadvantage and AAVE. That is, AAVE use should higher for those who identify as economically disadvantaged compared to those who do not.

Direct Effects of Variables on Writing Score

Results of several studies suggest that higher levels of AAVE use may result in lower writing scores (Johnson & VanBrackle, 2012; Richardson, 1997), phonological performance (Mitri & Terry, 2014), reading achievement (Wheeler et al., 2012), and spelling performance (Treiman, 2004; Treiman & Bowman, 2015). Considering that this lack of acknowledgement of AAVE is mirrored in the grading rubrics used for the assessment of writing in this study, I would expect greater AAVE use to result in lower writing scores. Based on the previously mentioned literature, I hypothesized that the relationship between AAVE use and writing scores would be a negative one.

Turning to the link between economic disadvantage and writing scores, Charity et al. (2004) concluded that schools with a higher proportion of students enrolled in federal lunch programs (a proxy for economic disadvantage), obtained lower average scores on

phonological and grammatical tasks (Charity et al., 2004). This suggests a negative

relationship between economic disadvantage and performance on writing tasks.

Kim et al. (2001) studied 304 first-grade students (Mean age= 6.18) to investigate

the degree to which growth trajectories of writing performance differed depending on

students' socioeconomic status (SES), measured by whether children were enrolled in

free or reduced-price lunch programs. Students' writing was scored on the following four

criteria: quality of ideas, organization, number of words, and number of ideas.

Researchers found that free and reduced-price lunch status had a negative relationship to

writing scores. Overall, students enrolled in free and reduced-price-lunch status had

lower scores on quality of ideas (β= -.40, p < .001), organization (β= -.22, p < .05),

number of words (β= -4.82, p < .001) and number of ideas (β=-1.00, p < .01) compared to

students who were not enrolled in free and reduced-price-lunch status. Because of the

relationship researchers found between SES and writing outcomes, I hypothesized a

negative relationship between economic disadvantage and writing scores.

With regard to the link between gender and writing scores, Reynold et al. (2015)

investigated whether there were gender differences in reading, math, and writing

performance on achievement tests. For the current study, I focused on their results

surrounding gender and writing performance as measured by the Kaufman Test of

Educational Achievement, Second Edition- Comprehensive Form (KTEA-II). The

researchers utilized the norming sample of 2,027 participants between the ages of seven

and nineteen and analyzed the data using a multiple-indicator multiple-cause (MIMIC)

model to investigate gender differences on writing performance while controlling for

cognitive ability. The mean difference between males and females on written expression

had an effect size of .44 to .46, which can be classified as a medium effect according to Cohen's d (Cohen, 1988). Overall, females performed better on writing than males ($b=$ -6.87, 95% CI [-8.31, -5.43].

Reilly et al. (2019) conducted a study to determine if there were gender differences in the reading and writing achievement of fourth, eighth, and twelfth graders, using the National Assessment of Educational Progress (NAEP). They wanted to observe historical patterns of these gender differences, so they studied the performance on reading and writing tasks across the span of three decades (1988-2015, N = 3.035 million students) using a meta-analytic technique. Regardless of grade and time of assessment, girls performed significantly better than boys on writing achievement (d= -.54). This effect size can be classified as medium in magnitude.

Additionally, Cronley et al. (2017) conducted a study on writing scores of students enrolled in the Baccalaureate of Social Work (BSW) program at a U.S. school of social work to see if demographic variables (i.e., race, gender, age, GPA) significantly predicted writing scores. Six instructors of composition were recruited to pilot a new rubric to score the essays for response, organization, grammar, and APA style. Researchers found that female students scored higher on grammar in comparison to males, $t(85)=-2.2, p < .05$.

All studies mentioned above dummy-coded girls as 0 and boys as 1, which indicated that negative values of effect size, correlation, or parameter estimates were in favor of females. In the current study, I aligned my dummy coding with the previous studies (females=0, males=1) and hypothesized that the relationship between gender and

writing scores would be a negative one, indicating girls will obtain higher writing scores than boys, on average.

To justify the hypothesized relationship between race and writing, I bring your attention back to the results of Cronley et al. (2017). In addition to the influence of gender, those researchers also found that Black students (M=2.62), among other racial groups, received significantly lower scores on grammar than white students (M=2.95). Across all components of the rubric (response, organization, grammar, and APA style) except APA style, Black students scored lower than white students. Given this, I hypothesized a similar negative effect for Black students in this study.

Indirect Effects on Writing Score

I have previously established that a) Black students use AAVE more than white students (Lippi-Green, 1997) and b) greater use of AAVE is associated with lower writing scores (Inoue, 2015; Johnson & VanBrackle, 2012; Mitri & Terry, 2014; Randall, 2021; Richardson, 1997; Wheeler, 2012), so it is hypothesized that one reason for the lower writing scores of Black students is AAVE use, which suggests an indirect effect of race on writing scores through AAVE usage.

Current study

For the current study, I conducted a preliminary study using Natural language Processing (NLP) as a method to quantify features of AAVE in a large number of essays. Past studies have not utilized NLP methods to investigate patterns of AAVE use in the writing of Black students or how the frequency of AAVE use impacts writing scores on a standardized assessment. After I quantified the frequency of AAVE use with NLP, I used methods of Structural Equation Modeling (SEM) to test the model presented in Figure 1.

The paths hypothesized between my variables of interest served as my hypotheses for the expected relationships, based on previous literature. The next chapter provides extensive background literature to provide the reader with didactic information regarding the history, theoretical framework, and research used to investigate the relationship between AAVE, race, and perceived writing performance.

Chapter 2: Literature Review

This chapter serves as a didactic presentation of the background literature encompassing the history, theoretical framework, and research utilized to investigate the relationship between race, AAVE, and perceived writing performance. The following paragraphs provide a brief overview of what will be discussed in this chapter. I begin with an explanation of what AAVE is, its historical importance, the mistreatment of the language in the education system, and the implications for this mistreatment. This will provide the reader with a better understanding of AAVE, the focus of this study, and the dangers that arise when this language is misunderstood.

Because AAVE is a language mostly used by individuals in the Black community, it is imperative to explain how race and language are interconnected for Black students. As explained in the introduction, I used CRT as a theoretical framework to explain how the treatment of AAVE is representative of the manifestation of the racism Black students experience in society. Anti-Black Linguistic Racism (Baker-Bell, 2020) is a manifestation of systematic racism that surfaces in the classroom through the ways in which students' use of AAVE is penalized. I use the tenets of CRT to discuss the historical and current experiences of Black students, and I use Anti-Black Linguistic Racism to provide a name for how racism functions through language discrimination in the education system.

After the establishment of this relationship, I argue that one reason for the educational inequities in writing performance seen in Black students is the penalization of their use of AAVE. To justify this, I provide an overview of literature that investigated the negative impact of AAVE use on writing scores. I argue that this negative impact

could be a result of the grading practices used in current writing assessments. In psychometric terms, I argue that the construct of writing is underrepresented for Black students because it ignores their linguistic background. I argue that the reliance on SAE seems implicit on writing rubrics and is central to my argument that current scoring practices in writing penalize the use of AAVE. I highlight the importance of investigating the preferencing of SAE in writing as part of the purpose of my study. I introduce my current study as a preliminary analysis that utilized NLP to quantify AAVE use in the writing of Black students. I conclude this chapter with a discussion of my use of methods of Structural Equation Modeling (SEM) to investigate the nature of the hypothesized relationship between race, AAVE, and scores on a standardized writing assessment.

What is AAVE?

In this section, I provide background information on AAVE to orient the reader to key features and markers of AAVE that differentiate it from other languages and may contribute to its penalization in writing performance. Features of AAVE often arise in the writing of Black students and were used as a guide to identify markers of AAVE in the current study. I will first discuss AAVE as a cultural marker of the Black community, why and when it was first developed, and end with the prominent features of AAVE discussed throughout this study.

As mentioned in the introduction of this paper, language is not only used to communicate, but is also often seen as a cultural marker for specific groups (Kellogg, 2016). More specifically, AAVE is the cultural marker of the Black community with about 80% of the Black community identifying as speakers of AAVE (Lippi-Green, 1997). AAVE emerged from U.S. slave descendants, as Black people were violently

forced into the United States from different countries and linguistic backgrounds. There

was not a standardized way for enslaved people to communicate, so AAVE emerged out

of necessity to establish community centered around shared experience and language

practices (Baker-Bell, 2020a; Baldwin, 1979; Smitherman, 2006). U.S. laws banned

enslaved people from reading, so Black people developed another way to communicate

that was not represented in the books they would be murdered for opening. White people

kept their language from Black people and now punish them for using their own language

that was developed out of pure necessity.

As AAVE has evolved, you may have heard it referred to as Ebonics, Black

English, Black Language, African American English (Baker-Bell, 2020a), or incorrectly

referred to as slang and a "spin-off" of Standard English. Although AAVE has

similarities to Standard English, it has unique "semantic, grammatical, pronunciation, and

rhetorical patterns" (Smitherman, 2006, p.3; see also Mitri & Terry, 2014) that dictate its

use and distinguish it from other variations of English. AAVE is a rule-governed

language with grammatical features and rules that differentiate it from SAE. To illustrate

this, I use this section to provide an overview of common rules, features, and

characteristics of AAVE. This list is limited to features most commonly seen in the

writing of Black students and includes the following: (a) copula absence of is and are

(Ivy & Masterson, 2011) (e.g., We talking to the professor after class), (b) completive

done (Thompson, Craig, & Washington, 2004) (e.g., I done told her to put a jacket on),

(c) multiple negation (Thompson et al., 2004) (e.g., she don't walk to school no more),

(d) indefinite article-a (e.g., I forgot a umbrella), (e) Habitual/Invariant "be" (e.g., They

be going home on Thanksgiving) (Thompson et al., 2004), (f) absence of regular tense -

ed (e.g., They play yesterday), (g) the use of ain't, (Ivy & Masterson, 2011), (h) absence of "be" auxiliary (e.g., She going to office hours), (i) "g" dropping (e.g.,, They playin at the park), (j) consonant cluster reduction (e.g., The pirate has a treasure [ches] (chest)), (k) devoicing final consonants (e.g., devoicing final consonant [d];"She sleeping on the /bet/ (bed) (Mitri & Terry, 2014, p. 557)), and (l) zero possessive (e.g., "The bird beak is yellow" (Mitri & Terry, 2014, p.557)). These features are provided along with examples in the methods section in Table 3.

As mentioned previously, the prevalence of the use of AAVE in the Black community is about 80%, so most Black students are entering school with this linguistic background. In the section that follows, I transition to a discussion of the stigma towards the use of this language in the classroom.

Mistreatment of AAVE in the Education System

A commonly cited court case associated with AAVE use in the education system is the 1979 Ann Arbor, Michigan Black English Court case (Ball & Lardner, 1998). The plaintiffs, on behalf of fifteen elementary school aged children, challenged the school in court to take responsibility for failing to teach the children how to read, among other academic inequities. The court concluded that the home language (AAVE) of the students was not acknowledged in the classroom and was not considered in assessing their ability to learn, which prevented teachers from tailoring the course material to encompass non-SAE speakers. The outcome of this case was the development of programs and curriculum that were more inclusive of AAVE speakers. This was one of the first cases that highlighted the influence of teachers' negative attitudes toward AAVE on the creation of language barriers for Black students and the negative effects these attitudes

had on students' learning. This case served as a "legal intervention" to hold educators

accountable for the underachievement of their Black students due to "inadequate,

ineffective curricular and pedagogical routines" (Ball & Lardner, 1998, p.470).

The decision of the Ann Arbor Black English case placed the responsibility for

the underachievement of Black students in the hands of the Ann Arbor School District

Board and acknowledged that the negative attitudes teachers held toward AAVE was an

impediment to student academic progression. This case resulted in a historic shift by

"drawing national and international attention to the role of language variation in the

education of Black children" (Ball & Lardner, 1998, p.471). The teachers' refusal to

acknowledge AAVE as a rule-governed language contributed to the development of

negative attitudes toward that language in the education system, which directly

influenced the educational well-being of Black students and played a huge role in their

futures (Ball & Lardner, 1998, p. 472).

About 20 years later in Oakland, California, the 1996 Oakland Ebonics

Controversy surfaced when the Oakland Unified School District challenged the Oakland

School board to recognize Ebonics [AAVE] as the primary language of African

American students (Rickford, 1999). Oakland Unified School District argued that Black

students were not adequately being educated due to AAVE not being acknowledged in

schools. With the support of numerous linguists, the resolution of the court case was to

recognize AAVE as a systematic, rule-governed language and to take Black students'

linguistic backgrounds into account in their instruction. The decision to recognize AAVE

as a rule-based language in California was faced with a lot of controversy on the part of

the school district, the teachers, and the general public. This decision was accompanied

with jokes about the legitimacy of AAVE and the linguists who aided in the resolution

received a plethora of hate mail. In "The Ebonics controversy in my backyard: A

sociolinguist's experiences and reflections", John R. Rickford (1999) stated the

following:

> In cases like these, language was no longer at issue; 'Ebonics' had become a
>
> proxy for African Americans, and the most racist stereotypes were being
>
> promulgated. This cruel humor might remind us, however, that behind peoples'
>
> expressed attitudes to vernacular varieties, there are often deep-seated social and
>
> political fears and prejudices about their speakers. If we don't take the 'socio' part
>
> of sociolinguistics seriously, we won't be prepared to understand or respond to
>
> such attitudes effectively. (p. 272)

Harris and Schroeder (2013) provided an overview of research on the literacy

performance of AAVE speakers with a focus on the tendency for AAVE to be viewed

from a deficit approach versus a difference perspective of language. The overview

concluded with a discussion of the 1974 Ann Arbor AAVE court case and the Oakland

Ebonics controversy and how these two cases illustrated how harmful the deficit

approach was when assessing Blacks students' literacy performance. An example of the

deficit approach would be how the education system has historically viewed

the underperformance of Black children on standardized tasks in comparison to white

children as a representation of deficient language skills rather than as representative of

their use of AAVE (Harris & Schroeder, 2013; Wheeler, 2019). By recognizing AAVE as

the language of Black students, the Ann Arbor and Oakland court cases viewed the

performance of Black students from a difference perspective rather than a deficit one. In

contrast, by ignoring the cultural and linguistic background of their students, teachers in

the two districts viewed the students' performance from a deficit approach, thus

contributing to their perception of underachievement of Black students. The schools and

the school boards were urged to recognize that the variation in language and academic

performance of Black students was not indicative of intellectual weaknesses, but of

cultural and linguistic factors unique to the Black community. Considering differences in

language, SES, culture, and other environmental factors would provide a better and more

accurate understanding of literacy performance. Given this deficit view, students who use

AAVE in the classroom can be seen as entering school at a disadvantage with regard to

language

Unfortunately, the perception of AAVE in the classroom has remained the same

even after AAVE was classified as a legitimate language and the language of the Black

community in multiple court cases. Its legitimacy is still challenged in the classroom, as it

is not perceived as academic because it is not Standard American English (SAE), the

language the education system is centered around. In 2019, Wheeler conducted a study

that analyzed the data of in-service teachers and pre-service teacher education students as

they navigated a dialect awareness curriculum. In-service teachers were middle school

and high school teachers located on Virginia's Eastern Shore and pre-service teacher

education students were enrolled in the Master of Arts Teaching program at Christopher

Newport University. As part of the curriculum, participants were equipped with

alternatives to marking differences in dialect as incorrect. One alternative was a T-chart

graphic organizer. The left side of the chart contained an AAVE phrase, and the right side

was the SAE alternative. At the bottom of each category the pattern of the grammatical

rule was indicated. Rather than marking the AAVE version as incorrect, the graphic

organizer allowed the teacher to recognize the rule of AAVE associated with the response

and the resources to provide the SAE equivalent.

As a result of the workshop teachers claimed they developed new attitudes

towards language use in the classroom and were appreciative of the workshop

highlighting that the language patterns seen in AAVE were not indicative of a lack of

language ability. Even with said attitude shift, when teachers returned to their respective

classrooms, they continued to mark errors related to dialect as incorrect forms of SAE

even after the workshop had provided them with methods to be more inclusive in their

grading practices. The education students acknowledged that considering dialect patterns

was "too hard" and said they would just "count 'em wrong" (Wheeler, 2019, p.9).

Wheeler (2019) addressed the fact that deep-rooted dialect prejudice was

prevalent in the education system. Her findings provided evidence that current educators

and students preparing to enter the education system as teachers did not consider the

association of race and language in the assessment of writing. The participants held the

privilege to decide not to understand students' linguistic backgrounds despite the

negative consequences that decision had on Black students. Even though they found it

difficult to learn the rules of AAVE, they could not empathize with a Black student

struggling to articulate themselves in SAE due to their familiarity and comfort with

AAVE. Why must one group navigate discomfort while another group can choose not to?

Effects of Devaluing AAVE on Student Attitudes

Baker-Bell (2020b) further investigated the historical background of the

persecution and marginalization of AAVE in the classroom and introduced the term *Anti*

Black Linguistic Racism. This is defined as the "linguistic violence, persecution,

dehumanization, and marginalization that Black language (BL) [AAVE] speakers endure

when using their language in schools and in everyday life" (Baker-Bell, 2020b, p. 9). For

this study, Baker-Bell (2020b) observed a ninth grade English Language Arts (ELA)

class in Detroit, Michigan. Students in the class identified as AAVE speakers. Baker-Bell

used an activity that consisted of the following instructions: "1) Read two language

samples, 2) Draw an image, cartoon, or character that reflects each language sample, and

3) Write a paragraph that expresses your thoughts about both languages and the speakers

of those languages" (Baker-Bell, 2020b, p. 11). One language sample represented AAVE

and the other reflected SAE, but the students were not made aware of this distinction.

Upon completion of the activity, a group dialogue was facilitated to discuss

student perspectives of each language example. The comments students made

surrounding Black English reflected an internalization of the negative misconceptions

and perceptions society holds toward the language, and toward Black people in general.

The students described the AAVE language sample as representing a student who was

"bad, gets bad grades, disrespectful, etc." and described the White Mainstream English

[SAE] sample with the words "proper, school language, teacher, and knowledge" (p.12).

These negative misconceptions surrounding AAVE in the classroom perpetuate the idea

that White Mainstream English [SAE] is the academic standard and that anyone not

abiding by the linguistic rules of the dominant culture are academically incorrect. Backer-

Bell urged educators to shift towards an anti-racist Black language pedagogy and develop

more inclusive teaching practices for students of color. As noted by Thomas (2018), the

lack of acknowledgement of AAVE as a language and the negative attitudes towards it

have negative effects on students in the classroom.

Students' internalization of negative attitudes toward their use of AAVE can also

affect their attitudes toward learning. For example, Sealey-Ruiz (2005) recounted seven

years of teaching a composition class to Black adult learners. Students in the class were

encouraged to participate in writing exercises as well as class discussions. In her paper,

the author provided a few excerpts from a previous student who depicted an

internalization of the negative attitudes and stereotypes toward their home language and

believed their language was "wrong" or "incorrect." The lack of acknowledgement of

AAVE's legitimacy and the marking of AAVE as "incorrect" served as a form of cultural

rejection and reinforced the notion that AAVE did not have a place in academia,

regardless of its cultural relevance to Black students. One outcome of such rejection was

that Black students resorted to silence and disengagement in the classroom (Sealey-Ruiz,

2005).

Thomas (2018) constructed an autoethnography of her linguistic experience as an

AAVE and SAE speaker from a student and teacher perspective. She provided separate

narratives describing five crucial life stages (i.e., ten years old, adolescence/teenage

years, undergraduate college student, teacher, and doctoral student) and how each served

as moments in her linguistic and cultural development. Her self-reflection addressed the

linguistic harm African American students experience in the education system and how

the harm she experienced informed her treatment of AAVE as a teacher (i.e., encouraging

the use of SAE and the penalization of AAVE). One example of this type of linguistic

harm was the internalization of the stereotypes associated with AAVE which led Thomas

to fall silent and lack confidence in classroom settings among predominantly white peers.

Another example was the feeling of confusion and uncertainty when switching between

SAE and AAVE (code-switching). Thomas' fear of not sounding smart due to her AAVE

use led her to question whether she belonged at her university.

Theoretical Framework: Critical Race Theory

Up until this point, I have described what AAVE is, the historical mistreatment of

the language, and the impact this treatment has on the attitudes of Black students. The

next question I want to address is the following: Why does the devaluation and

mistreatment of AAVE occur? There are several answers to this question, but one has to

do with the relationship between race and language for Black students. I use Critical Race

Theory (CRT) to investigate the manifestation of racial hierarchy in society (racism)

through the preference given to SAE (i.e., whiteness) and the devaluation of AAVE (i.e.,

Blackness).

Before I elaborate on the manifestation of racial hierarchies in society as the

mistreatment of AAVE in the classroom, I will first define what CRT is and how it

informs this relationship. Derrick Bell, Kimberlé Crenshaw, Alan Freeman, and Patricia

Williams were a few of the many scholars involved in developing CRT. CRT was first

introduced as a framework influenced by research in both radical feminism and critical

legal studies through its historical relevance to court cases centered around race (Bell,

1995; Delgado & Stefancic, 2017). CRT was developed as scholars found that current

framings for inequities in the justice system were inadequate. In response, they proposed

several tenets to provide a foundation for CRT. Although there are several tenets, the

following are most applicable to education: (1) race is socially constructed with no

biological basis; (2) racism is not an aberration, but rather typical and the norm; (3) the

construct of race has been differentially applied historically; (4) interest convergence; (5)

narrative and counter narratives (Delgado & Stefancic, 2017). I will use the following

paragraphs to describe each of these tenets in detail in the order in which they were

mentioned.

The first tenet I mentioned was that race is socially constructed with no biological

basis. Race cannot be tied to any biological or genetic component but was constructed as

a way for society to categorize groups and consequently resulted in some racial groups

being labeled inferior and superior. Attaching biological basis to race contributed to the

argument that intelligence was hereditary. Because race is socially constructed, functions

such as personality, or intelligence cannot be attributed to someone's racial identity using

any form of scientific evidence.

Second, racism is not an aberration, but rather is typical and the norm. As

Delgado and Stefancic said, it is the "way society does business" (2017, p.8). Racism is

so heavily ingrained in U.S. society that it is perceived as a normal experience of people

of color (Bell, 1992). Additionally, Bell (1992) defined CRT as the acknowledgement of

racism as an everyday aspect of life and an active construct in the daily lives of people of

color. It is important to note that racism is not something of the past, nor is its presence in

society debatable, as it is a "permanent fixture in American society" (Ladson-Billings,

1998). It is not something one can ignore or address passively, but it is something that

must be actively dismantled through explicitly naming racism in all spaces.

Third, the construct of race has been differentially applied historically. Depending on the objectives of the dominant group, society "racializes different minority groups at different times" (Delgado & Stefancic, 2017). An example of this could be the shift in treatment towards Middle Eastern people before and after September 11th, 2001 and mislabeling an entire group of people as terrorists. Following the murder of George Floyd, Black people received another narrative and were depicted as violent protestors, inconsolable, and out of control. This tenet also supports that race is socially constructed and has historically been shifted to fit the narrative of the dominant group.

Fourth, the concept of interest convergence was theorized by Derrick Bell (1990) and is exemplified by the court case *Brown v. The Board of Education*, a historic case of the civil rights movement that contributed heavily to the desegregation of schools. This case is often highlighted when discussing the development of CRT from critical legal studies (Delgado & Stefancic, 2017; Ladson-Billings, 1998). Derrick Bell, a professor of law at the time, studied this case and concluded that the acknowledgement of racism in this court case was more for the advancement of elitist whites than for the benefit of Black students. He theorized this as interest convergence in which the motivation to combat racism was for the benefit of both elite whites and working-class whites, concluding that white people will only be willing to eradicate aspects of racism when it will benefit them (i.e., desegregation of schools). On the surface it may not be obvious that school desegregation advanced white interests. However, interest convergence, as illustrated in the Brown vs. Board of Education case discussed above, questions the timing of this "civil rights victory". Years before this specific case was brought to the Supreme Court, there had been many other desegregation cases that did not receive the

same urgency as the Brown V. Board of Education case. It was not until the Black

community's interests of desegregation aligned with the need for an image change for the

American government that the Supreme Court found it necessary to appear concerned

about the well-being of its minority populations. The United States was in the midst of

the Cold War during the decision to desegregate schools. The United States and the

Soviet Union were in a competition to develop partnerships with uncommitted nations.

Because of this, it was beneficial for the United States to combat the violent racism in the

country by "softening its stance towards domestic minorities" (Delgado & Stefancic,

2017, p.23) through the desegregation of schools, to better appeal to surrounding nations.

And last, but certainly not least are narratives and counternarratives. Delgado and

Stefancic (2017) emphasized the need for people of color to tell their stories, as the world

is experienced very differently depending on which racial group someone identifies with.

CRT encourages storytelling as a method to process and heal from racial trauma (Ladson-

Billings, 1998) and the first step to truly understand how racism operates in society.

Narratives and counternarratives are crucial in understanding the experiences of Black

students in the education system, as their experiences do not align with those of the norm

group. Methods of storytelling highlight deep rooted areas of racism that need to be

combatted, specifically stories that relate to Black students and their education.

Critical Race Theory and Education

As can be gathered from the section above, racism is heavily embedded in our

society and has manifested in many ways throughout history. Specifically, racism is

embedded in the American education system in the way in which students are taught and

assessed. Racism manifests in education in the following ways: (a) the treatment of the

white students' experiences, language, and values as neutral, superior, and correct; (b)

educational inequities specific to Black students; (c) the existence of a linguistic

hierarchy.

The first manifestation of racism is the treatment of white students' classroom

performance, experience, and language as neutral, superior, and correct. This can be

explained by the color-blind approach present in the United States education system.

Color-blind is defined as the belief that one does not see color (i.e., race) and the belief in

"…equality, especially equal treatment of all persons regardless of their different

histories or current situations" (Degaldo & Stefancic, 2017, p. 26). APA's Inclusive

Language Guidelines state that

> Color-blind beliefs in a racial sense purportedly are based on the assumption that
>
> acknowledgement of race reifies racial divisions in society. People who endorse
>
> color-blind beliefs believe that individual effort is sufficient for achievement in a
>
> meritocracy predicated on the assumption that everyone has equal opportunity for
>
> life success. People who hold color-blind attitudes believe that doing so reduces
>
> racial and ethnic prejudice and that acknowledging racial and ethnic differences
>
> promotes racial division. (p.11)

In other words, it is a belief that racism can be reduced to the acknowledgement

of differences in skin pigmentation and neglects the systematic structure of racism. But

racism cannot be dismantled using such a passive approach because it does not absolve

the systematic disadvantage that will still be endured by people of color. Color-blind

beliefs do, however, "enable white people to continue theoretically, and morally, and/or

otherwise objecting to racism and racial inequality, while rejecting any real actions,

policies, behaviors, and understandings that could work toward dismantling systematic

racial inequality" (Jayakumar & Adamian, 2017, p. 915).

A color-blind approach to the education system does not allow the

acknowledgement of the educational inequities of Black students and removes

accountability from the education system. Rashid (2011) discussed the dangers of a

color-blind approach to education. "As a consequence of this approach, the massive

underachievement of African American students is not viewed as a systemic problem

stemming from poor teaching, insufficient academic support, lack of adequate

educational resources, or other factors that would warrant reform of the educational

establishment" (p. 589). Instead, poor performance of Black students tends to be ignored,

dismissed, or blamed on the students. Because the education system is centered around

whiteness, and education is approached with a color-blind lens, the performance of white

students is viewed as neutral and the norm, and therefore the academic standard (Randall,

2021). So rather than acknowledging that the performance of Black students is informed

by their cultural background, linguistic background, and experiences and inherent

differences, their performance is considered deficient because it does not align with that

of white students. This can be seen in the tendency for white students to receive higher

scores on consequential tests in comparison to Black students. It is important to shift

from a color-blind perspective to acknowledge the inequities of Black students as a

systemic issue.

In support of the second manifestation of racism in education, Inoue (2015)

expounds on the inequities Black students experience in the education system through the

investigation of performance on the written component of the English Placement Test

(EPT). This test was administered in California to incoming students at Fresno State University to determine if students needed to take a remedial class prior to beginning the semester. Additional classes (i.e., remedial classes) were required if students scored at the remedial level, which, based on their scores, implied that they needed more help with their English to succeed in their first year, and college in general. In 2013, it was found that students of color, specifically Black students, were classified as remedial at higher rates (43.7%) than their white classmates (23.6%) on the EPT. In addition to this finding, it was also seen that Black, Latino/a, and Hmong students received lower scores on their final portfolio in a first-year writing program in comparison to white students.

I argue that the vastly different performance of students of color is not a coincidence nor is it surprising, but rather is a systemic issue. One reason that could be inferred for these findings is the expectation that students are proficient in SAE when taking the EPT and must be proficient in SAE to succeed in their college career (i.e., requiring remedial classes to improve their "English"). This implies that higher education and the EPT require SAE, and this requirement differentially impacted Black students more than it did white students. Because Black students are likely to come from different linguistic backgrounds than SAE, I argue that one reason for these lower scores could be a result of their linguistic background (i.e., AAVE).

Finally, the implicit expectation that all students should use SAE in academic environments engenders a linguistic hierarchy in which students are expected to use SAE in school even though this may not be the language they use at home or in their social environment. The existence of this linguistic hierarchy constitutes the third manifestation of racism in the U.S. education system.

The comparison to the dominant language (i.e., SAE) is what characterizes the treatment of Black students' language as a deficit one. Deficit theory is typically applied to minority students in the education system (Delgado & Stefancic, 2017). An example of a deficit approach to language skills would be viewing the features of AAVE as a "deficient language system" and an incorrect way to demonstrate language skills. However, as noted previously AAVE is a rule-governed language that is a legitimate way in which someone can choose to express themselves (Harris & Schroeder, 2013). A differences approach to language skills would view AAVE as a unique language system and an acceptable demonstration of language proficiency. Under a differences approach, the fact that AAVE is different from SAE does not mean the latter language is right while the former is wrong. Yet, this is how racism manifests in the education system, and is one way in which white supremacy is upheld in the education system. Because SAE is associated with whiteness, and white supremacy deems whiteness superior, the education system inherently preferences whiteness (i.e., SAE) over Blackness (i.e., AAVE).

The history of white supremacy contributes to the manifestation of this linguistic hierarchy. Baker-Bell (2020a) made the argument that "Black children continue to be 'unconsciously trained to correlate blackness with wrongness and whiteness with rightness" (p. 24), which further supported the idea that AAVE is associated with wrongness in the classroom and SAE is deemed the "gold" academic standard. The enforcement of SAE and the devaluation of any language deviating from SAE in the education system is a way to maintain white supremacy.

The Connection between Race and Language

I open this section with the following quote from April Baker-Bell (2020a):

By linking the racial classifications Black and white to language, I am

challenging you, the reader, to see how linguistic hierarchies and racial

hierarchies are interconnected. That is, people's language experiences are

not separate from their racial experiences. Indeed, the way a Black child's

language is devalued in school reflects how Black lives are devalued in the

world. Similarly, the way a white child's language is privileged and

deemed the norm in schools is directly connected to the invisible ways that

white culture is deemed normal, neutral, and superior in the world. (p. 2)

Baker-Bell (2020a) named the linguistic experience of Black students in the

classroom *Anti-Black Linguistic Racism*. *Anti-Black Linguistic Racism* coupled with CRT

allowed the exploration of racial hierarchies as well as linguistic hierarchies. CRT

provided the tenets to provide the historical background of the racial hierarchies Black

students experienced in the education system (i.e., educational inequities). It provided the

answer to the question of why AAVE is mistreated. It is mistreated because it is attached

to Black people, so it was important to discuss the mistreatment of Black people within a

CRT framework. Anti-Black Linguistic Racism is the manifestation of these racial

hierarchies as linguistic hierarchies in the way literacy is taught and demonstrated in the

classroom. For example, it would be inadequate to label Black students' experiences in

the education system as just racism, because not only are they discriminated against by

the color of their skin but also by the way in which they speak. Because of this, Baker-

Bell (2020a) found it necessary to develop a term to describe the unique linguistic

experience of Black students. She defines *Anti-Black Linguistic Racism* as "…the

linguistic violence, persecution, dehumanization, and marginalization that Black

Language-speakers [AAVE] experience in schools and everyday life" (p. 11).

I agree with Baker-Bell's conclusion that race and language are interconnected

through the manifestation of racial hierarchies as linguistic hierarchies (2020a). These

hierarchies are inherent in the preferencing of SAE in the education system and the

devaluation of AAVE. SAE has a dominant presence in society which has contributed to

the stigmatization of "nonstandard American English" (Mitri & Terry, 2014). One of the

best-researched "nonstandard forms of Standard English" is AAVE. Most peoples'

perception of AAVE is established through the comparison of AAVE to Standard

American English (SAE), due to the habitual comparison of Black and white students

(Baker-Bell, 2020a; Harris & Schroeder, 2013; Inoue, 2015) in the education system,

educational research, and in society in general (Delgado & Stefancic, 2017; Mitri &

Terry, 2004). It is common to see SAE referred to by terms such as White Mainstream

English (WME), Mainstream American English (MAE), and Academic English (Baker-

Bell, 2020b; Mitri &Terry, 2014). These labels reinforce the notion that because SAE is

the language of white people, it is academic, neutral, and superior (Baker-Bell, 2020a).

Because AAVE is the language of Black people, and SAE is the language of

white people, SAE is held as the standard not only in society in general, but also in the

classroom. This expectation that SAE should be used in the classroom disadvantages

students color as these are students that come from multilingual or bilingual backgrounds

(Inoue, 2015). In the following section, I discuss the consequences of this expectation by

investigating the relation of AAVE use with academic performance.

AAVE and Academic Performance

Research on AAVE use has shown lower levels of classroom performance for schools in which AAVE is present (Charity et al., 2004; Mitri & Terry, 2014; Treiman, 2004; Treiman & Bowman, 2015; Wheeler, 2012). In this section, I will discuss the impact AAVE has on classroom performance in general prior to discussing its specific negative influence on perceived writing performance as evidenced by lower scores.

Charity et al. (2004) conducted a study of 27 African American students aged five to eight enrolled in kindergarten through second grade. These students attended low performing schools within low-income communities located in Cleveland, Ohio, New Orleans, Louisiana, and Washington, D.C. The goal of the study was to see if greater familiarity with SAE was positively correlated with reading achievement. Familiarity with SAE was measured by how readily students could repeat SAE sentences verbatim (sentence imitation) when instructed and recall a story that was initially told in SAE (story recall). Students' performance on these tasks were then quantified into phonological and grammatical scores. To measure reading achievement, each participant was given the following three subtests of the Woodcock Reading Mastery Tests-Revised: Word Identification, Word Attack, and Passage Comprehension. The authors found that greater familiarity with SAE was associated with increased performance on reading achievement ($p < .01$). This finding suggests that students not fluent or well-versed in SAE could have reading difficulties. By extension, because reading is involved in most classroom learning, lack of familiarity with SAE could negatively influence overall classroom performance.

        Treiman (2004) conducted a study that investigated spelling errors in college

students and concluded that errors attributed to language differences (i.e., final /d/

devoicing) were still present in adulthood. This study was conducted on students

attending Wayne State University located in Detroit, Michigan. The total number of

participants was 93 (White=47, Black=46). Participants were exposed to 40 words; 20 of

the words had unstressed syllables ending with /d/ and 20 had unstressed syllables with a

final /t/. Participants were first instructed to orally rate each word according to their level

of familiarity with it. The researchers then used the word in a sentence and repeated the

word. Finally, participants were instructed to spell the words and their responses were

recorded. Researchers found that a respondent's reliance on the oral production of

AAVE, as measured by the students' taped pronunciation in the previous task,

contributed to dialect-related misspellings. This study showed the tendency for

adult spellers to rely on AAVE, rather than SAE phonological characteristics of a word to

inform their spelling. The results support the influence AAVE use may have on test

performance at the college level, when certain components (e.g., writing and reading) of

tests require proficiency in SAE.

        Treiman and Bowman (2015) also investigated dialect-specific spelling errors in

Black students but studied children rather than adults. The researchers conducted a study

consisting of 29 Black children and 32 non-Black children in kindergarten through first

grade in Detroit, Michigan. The aim of the study was to see if phonological features of

AAVE impacted the way children spelled. They specifically focused on the "devoicing of

final /d/" (p. 1017). An example of this feature would be AAVE speakers having

difficulty spelling a word ending in [d], as they are more likely to pronounce [d] ending

words with a [t] (e.g., spelling salad as salat). Researchers assessed this using an

experimental spelling test consisting of 16 words, eight with a final consonant of [t] and

eight with a final consonant of [d]. Following the spelling test, children were instructed to

pronounce the words from the spelling test and their responses were recorded. Their

responses were coded by two raters as either ending in [d], [t], or not able to be scored.

They found that final /d/ words had a significant effect of group membership (β= -1.78, p

< .001). Black students had significantly lower performance on spelling words with the

final d in comparison to non-Black students. Specifically, Black students were more

likely to spell [d] ending words with a [t].

Mitri and Terry (2014) conducted a study on AAVE use and its influence on

phonological awareness as it related to reading ability/skills with Black children in

kindergarten through second grade**.** Phonological awareness was defined as "the ability to

attend to, think about, and manipulate phonological aspects of language" (p.558). The

frequency of AAVE use was assessed using a sentence imitation task where students

were presented with 15 sentences from a computer screen read by a SAE speaker. These

15 sentences were dialect-sensitive, which is defined as including words or phrases prone

to be wrongly identified by AAVE speakers. The students repeated the sentences, and

their responses were recorded and transcribed to identity potential features of AAVE.

After the transcription, phonological and grammatical difference scores were computed,

but only the phonological score was used in the study. These scores represented the

percentage of times in which AAVE phonological and grammatical characteristics were

used instead of SAE during the sentence imitation task. The students' phonological

awareness was assessed using the sound awareness subtest of the Woodcock-Johnson

Test of Achievement, third edition (WJ3) and an experimental measure that required

students to match final phonemes to words. The sound awareness subtest required

students to rhyme words and manipulate aspects of words to make them rhyme. In the

experimental measure, researchers provided participants with an image of ten objects and

instructed students to repeat the name back to the examiner. They were then shown three

more pictures, and they were instructed to choose the picture that rhymed with the first

picture. For example, if a student was shown a picture of a tooth, they were instructed to

identify the name, and were then shown three images of a booth (SAE), roof (dialect-

sensitive), and juice (neutral). Out of the three images they were supposed to choose the

one that rhymed with the target item, tooth. Reading and vocabulary skills were targeted

using the letter-word identification subtest from the WJ-III. This subtest targeted the

potential pronunciation differences between SAE and AAVE, and how these differences

influenced students' efficiency in identifying which words rhymed with a targeted word.

      Mitri and Terry (2014) found that students with high and low levels of AAVE use

significantly differed on their performance on phonological awareness tasks. Students

with high levels of AAVE use scored lower on achievement measures (i.e., letter word

identification, sound awareness, and picture vocabulary) in comparison to students with

low levels of AAVE. They also found that AAVE use was significantly related to

phonological performance. Students with high levels of AAVE were more likely to

choose dialect-sensitive words on the phonological rhyming task than students with low

levels of AAVE use. Researchers also found that there were significant negative

relationships between AAVE phonological difference scores and the three achievement

measures (*r*=-.39 to -.5). In other words, students with more AAVE use had lower

achievement scores.

        The Mitri et al. (2014) and the Treiman et al. (2015) studies introduced a potential

theoretical explanation for their findings, termed the linguistic mismatch hypothesis

(Mitri & Terry, 2014) or the linguistic interference hypothesis (Treiman et al., 2015).

Although the linguistic interference hypothesis and the linguistic mismatch hypothesis

are conceptually similar, the authors provided different terms for these. For simplicity, I

will refer to it as the linguistic interference hypothesis. Treiman and Bowman (2015)

defined this hypothesis with the following statement: "English aligns less closely with

spoken language for speakers of AAVE than for speakers of SAE, leading to certain

difficulties in learning to read and write for AAVE speakers" (p. 1015). The authors used

this hypothesis as a theoretical explanation for the negative relationship of AAVE use

with reading achievement and spelling ability (Mitri & Terry, 2014; Treiman & Bowman,

2015).

        Wheeler et al. (2012) provided a vignette that examined the assessment of reading

performance designed to determine promotion to the next grade and how scores differed

when AAVE usage was not penalized versus when it was. The first part of the vignette

introduced a teacher knowledgeable on patterns of AAVE usage who administered the

Developmental Reading assessment (DRA2) to a fourth-grade student. The DRA2

required students to read passages out loud while the administrator assessed their

performance for reading errors. Students were required to obtain at least a 91.6%

accuracy rate to pass. The teacher in the vignette acknowledged that some of the "errors"

the student exhibited were a result of dialect differences. When dialect differences were

present the teacher did not mark the pronunciation as incorrect, which lead to five out of

the student's 21 potential reading "errors" being attributed to dialect differences rather

than actual reading errors. The student was provided a 92.2% accuracy rate, which was

high enough for the student to advance to the next grade level. In the next vignette the

researchers observed the DRA2 administration and scoring of a teacher who was unaware

of AAVE patterns. This teacher marked all reading errors as incorrect regardless of

whether errors were related to dialect. The differences between the two scoring methods

were striking. When dialectic errors were penalized, the student was assumed to have

performed too poorly to advance to the next level/grade with an 89.76% accuracy rate.

This was an inaccurate representation of the student's academic standing due to the lack

of acknowledgement of the language barrier that contributed to the results. The failure to

consider AAVE use when scoring the assessment resulted in the student being labeled a

"struggling reader", whereas consideration of AAVE use in scoring painted a more

accurate picture of the student's actual reading ability, allowing the student to advance to

the next reading level.

       Considering the literature up until this point, it is apparent that the use of AAVE

in the classroom can have a negative impact on students' scores across multiple

performance domains (i.e., phonological awareness, reading achievement, and spelling

achievement). I want to transition now from the impact on overall academic performance

and focus on the influence of AAVE use on the perceived writing performance of Black

students.

AAVE and Writing Scores

One can consider how pronunciation differences at the phonological level might influence the writing performance of AAVE speakers just as it affected student's reading performance in the studies by Mitri and Terry (2014) and Wheeler (2012) and students' spelling performance in the studies by Treiman and Bowman (2004; 2015). Pronunciation/articulation differences could show up as potential spelling errors in students' writing due to students spelling words as they sound in AAVE. It would not be surprising to see students differentially penalized for this when language differences are not considered in the assessment of writing ability.

Ball (1996) provided insight on one teacher's attitude toward AAVE and how it resulted in penalization of students in the form of lower writing scores. The influence of negative attitudes toward AAVE was apparent in the story the author opened the article with. The story was about a fifth-grade teacher who provided feedback to a Black student about her language use in a creative essay. The teacher mentioned that the story was enjoyable, but that the language was unsatisfactory due to the use of "ain't", a marker of AAVE. The teacher stated that the student's choice of language is what "forced" the teacher to lower her grade to a B-.

Richardson (1997) conducted a study to determine how the presence of AAVE discourse patterns influenced the scores students' essays received. In the study, she introduced the following techniques and language that AAVE speakers were prone to using in their writing: signification, narrative sequencing, and testifying. Richardson examined the writing samples of three students enrolled in her freshman composition class at Cleveland State University. The rubric used to assess the essays was a variation

of the rubric used in the National Assessment of Educational Progress (NAEP).

Richardson compared the scores she gave the essays to the ratings of nine instructors

from her AAVE discourse-pattern study. She found that when students used more

features of AAVE or a style unique to AAVE oral tradition, their essays received lower

scores in comparison to writing that adhered to more "standard rhetorical paradigms"

(Richardson, 1997, p. 71) or features and styles associated with SAE. Richardson also

made the important point that AAVE speakers may use language not acknowledged by

standard grading practices or by their instructors. Instead, instructors often view non-

standard English as an invalid demonstration of writing skills, hence the lower scores.

    Johnson and VanBrackle (2012) conducted a study to highlight the bias raters had

toward AAVE on a writing assessment. Specifically, they assessed how raters scored

writing assessments when errors were due to either students' use of AAVE, English as a

second language (ESL), or SAE, and how raters' responses to these errors may differ

depending on the linguistic background the errors stemmed from. Three-hundred and

fifty-eight essays were rated and obtained from a college competency exam called the

Regents Writing Exam (RWE). Nine of these essays were manipulated for the study's

objective. These nine essays were constructed based on three sample essays from

previous students on the RWE. The three sample essays were classified as one of the

following: high pass (3), pass (2), or fail (1). For each type of sample essay an AAVE,

ESL, and SAE version was created which resulted in the total of nine essays (e.g., each

level of essay had an AAVE, ESL, and SAE version). To accurately measure the level of

potential bias towards each type of error, each essay that was rated had the same number

of errors (N=8) and only differed in the language used. When reporting the results, the

researchers referred to high passing, passing, and failed essays as high, intermediate, and low, respectively. The researchers found that the failure rate was highest for essays with errors related to AAVE in comparison to those with errors related to SAE or ESL. Specifically, AAVE intermediate essays were 6.1 times more likely to fail as SAE intermediate essays, and AAVE high essays were 9.1 times as likely to fail as a SAE high essay. These findings highlighted the bias raters held toward errors related to AAVE and how these biases resulted in differential probabilities of passing or failing, with the likelihood of failing being highest for AAVE speakers.

In addition to rater bias, it is also beneficial to look at potential bias in the components of the rubrics used to grade writing assessments. Inoue provided the six criteria used to assess writing ability on the EPT: 1) response to the topic; 2) understanding and use of the passage; 3) quality and clarity of thought; 4) organization, development, and support; 5) syntax and command of language; and 6) grammar, usage, and mechanics. With this grading rubric, more students of color were classified as remedial in terms of their English fluency in comparison to white students. Inoue reviewed studies assessing the amount of variance explained in writing scores by the six criteria, and it was concluded that a large percentage (57%) of variance was unexplained by the rubric criteria. Although the factors that might explain this unexplained variance were not named, I argue that one of these could be the biasing toward SAE on writing assessments. Although not explicitly stated by the authors, I offer speculation that what these criteria may actually be assessing are 1) whether the response to the topic is aligned with whiteness, 2) whether students can understand and utilize the passage written in SAE, 3) the quality and clarity of students' use of SAE, 4) students' organization,

development, and support articulated using proficient SAE, 5) students' SAE syntax

usage and command of SAE, and 6) students' grammar usage and mechanics as they

relate to the rules and regulations of SAE. The requirement of SAE is not included on

these rubrics, but it is implied at every stage. Randall (2021) offers a similar argument in

that the implicit construct of literacy (e.g., reading and writing) is the ability to

demonstrate SAE literacy.

Grading practices anchored in preference for SAE usage have likely contributed

to the historic inequities seen in the writing performance of students of color on the EPT

(Inoue, 2015). One reason for these inequities on the EPT that Inoue highlighted was the

requirement for students to read and respond to a passage written in SAE, which assumed

SAE proficiency and consequently disadvantaged non-SAE speakers (e,g., Hmong and

Black students). Additionally, the grading rubrics were designed to assess these six

criteria based on one dominant form of English, SAE. Because of the way rubrics are

currently constructed they do not acknowledge the wide variety of language backgrounds

of the individuals taking writing assessments but are instead used to maintain a dominant

white discourse through the preferencing of SAE. Inoue (2015) expounds on this with the

following statement to support this implicit argument:

> To put it bluntly, when the function of writing assessment is primarily to promote
>
> local SEAE [SAE] or dominant discourse, without regard to the literacies that
>
> various racial formations bring to the classroom, or the various ways that
>
> particular racialized linguistic structures are judged by the teacher, then many
>
> students are treated unfairly. (p. 7)

Writing Score Differences as Evidence of Construct Underrepresentation

The findings discussed in the previous sections provide evidence that Black students are perceived as having lower overall writing performance in the form of lower scores on writing assessments. But why are Black students receiving lower scores on writing assessments? Based on the research discussed in the previous section, it seems possible that the scoring methods currently used to assess writing do not adequately address the writing skills of Black students because they are only focused on the use of SAE and therefore count AAVE usage as incorrect. The result is that the construct of writing is underrepresented for Black students.

I use the results from the previous study (Inoue, 2015) in conjunction with the "racism is not an aberration" tenet of CRT to shape my argument that current grading rubrics result in lower scores for Black students because the components of rubrics align with SAE and whiteness, and therefore disadvantage Black students and their use of AAVE. Because racism is normal and present in daily life, it is an active contributor to inequities in academic performance due to grading practices, thus inherently perpetuating racism. More specifically, if methods of teaching and grading are not explicitly made to be anti-racist, they will disadvantage people of color and continue to center whiteness as the norm and the standard (Baker-Bell, 2020a; Inoue, 2015; Randall, 2021).

In conducting my literature review, I did not find any empirical studies that have conducted research to specifically investigate how current writing rubrics preference SAE and disadvantage AAVE. But there are theoretical explanations for why this argument is valid when considering the writing performance of students that are non-native speakers of English. Paul Kei Matsuda (2006) provided a critical analysis on the historical development of current U.S. college composition classes. U.S. composition classes are

the English classes that students are required to take upon college entry to prepare them

for college-level writing. Matsuda argued that current composition classes perpetuate the

myth of linguistic homogeneity – "the tacit and widespread acceptance of the dominant

image of composition students as native speakers of a privileged variety of English" (p.

638). The assumption that English students are native speakers of SAE makes teachers

of writing, writing assessments, and grading practices ill-equipped to recognize writers

from different linguistic backgrounds. Instead, this assumption reinforces the notion that

good writers are those who can produce SAE proficiently. The following quote from

Matsuda (2006) illustrates how this assumption manifests in the grading practices of

writing assessments:

> Since any form of writing assessment – holistic, multiple-trait, or portfolio
>
> assessment—explicitly or implicitly included language as one of the criteria,
>
> writing teachers regularly and inevitably engage in what Bonny Norton and Sue
>
> Starfield have termed "covert language assessment."[2] (p. 640).

In psychometric terms, the construct of writing achievement is underrepresented

for Black students due to its overreliance on SAE use. Construct underrepresentation is

"the degree to which a test fails to capture important aspects of the construct. It implies a

narrowed meaning of test scores because the test does not adequately sample some types

of content, engage some psychological processes, or elicit some ways of responding that

are encompassed by the intended construct" (AERA, APA, NCME, 2014, p. 12).

Applying this definition to writing assessment, the exclusion of culture specific language

---

[2] Covert language assessment is defined as the "implicit assessment of non-native speaking and writing that takes place when the focus of the assessment is on the understanding of an academic subject such as geography, history, biology, or physics, and not on the examinee's language proficiency as such" (Norton & Starfield, 1997, p.292).

and syntax in the assessment of writing underrepresents the writing achievement of Black

and other minority students. The failure to represent these students' languages results in

construct underrepresentation and has been shown to result in lower writing achievement

scores (Ball, 1996; Inoue, 2015; Richardson, 1997), therefore disadvantaging students of

color while at the same time perpetuating racism through the continuous classification of

SAE as the standard for language use in education (Baker-Bell, 2020; 2021; Ball, 1996;

Inoue, 2015; Mitri & Terry, 2014; Randall, 2021; Randall et al., 2021; Wheeler et al.,

2012). Although often unspoken, it is nevertheless the case in the assessment of writing

that students are expected to align their writing directly with the rules of SAE. Any

language, culture-specific features, or vocabulary not adhering to the rules of SAE are

perceived as incorrect. Because of the exclusion of other forms of English, the construct

of writing is underrepresented for those who do not identify as SAE speakers. This failure

to represent the construct being measured in the same way across all groups of people is a

measurement issue; specifically, this represents a validity issue for the inferences made

from scores.

Randall (2021) provided a critical evaluation of current assessment practices, and

how they inherently exclude the experiences, cultural values, and languages of BIPOC

(Black, Indigenous, People of Color) students that supports the arguments made in the

previous paragraphs. She theorized that students were at a disadvantage on assessments

targeting literacy proficiency due to the construct of literacy being underrepresented for

them. This is because multilingual or nonstandard linguistic backgrounds are not

accommodated in the grading of writing assessments (Baker-Bell; Inoue, 2015; Randall,

2021). Because of this, culture specific words, phrases, or grammatical structures are

marked as incorrect. This results in differential penalizations in the writing scores of

Black students who are more likely to identify as multilingual or speak a form of

nonstandard English (e.g., AAVE). More specifically, Black students are demonstrating

writing ability, but because their use of AAVE does not align with SAE or with white

experiences or white culture, it is incorrect and is penalized in the form of lower scores.

This underrepresentation of writing achievement results in unfair inferences for students

from non-SAE backgrounds, but specifically for Black students as they are most likely to

be speakers of AAVE.

Quantification of AAVE Using Natural Language Processing

Race, gender, economic disadvantage, and writing score were already present in

my dataset. However, there was not a variable that captured the frequency of AAVE in

each writing sample. To create the AAVE variable in my path model I used a process

called Natural Language Processing (NLP) to investigate the amount of AAVE use in a

large number of essays. NLP, also known as computational linguistics, provides an

efficient way to code mass amounts of text data using automated extraction tools to "tag

words, sentences, and texts" (Crossley et al., 2014, p. 512). NLP makes it much easier to

quantify language features in order to conduct subsequent statistical analyses. The

principle aim of NLP is to "gather information on how humans understand and use

language through the development of computer programs created to process and

understand language in a human-like manner" (Crossley, 2013, p. 257). NLP was the

most efficient way to capture AAVE use, in comparison to other studies that tagged

features of AAVE by hand using multiple raters (Groenwold et al., 2020).

Additionally, NLP has not been widely used to assess the use of AAVE in the writing of Black students on standardized assessments. When NLP has previously been used to identify AAVE, it has focused on social media data (i.e., twitter through TwitterAAE dataset; Groenwold et al., 2020), which highlights the need for further investigation of AAVE in student essays.

Summary

Many consequential tests (e.g., SAT, ACT, GRE, EPT, etc.) require students to complete a writing component. Because some tests serve as a gateway to higher education, it is crucial that the tests are valid measures of writing ability so accurate inferences can be drawn from the results. As mentioned previously, there are unique features and structural properties of AAVE that differentiate it from SAE (Mitri & Terry, 2014; Ivy &Masterson, 2011). Because AAVE does not align to the rules and regulations of SAE, it is penalized on writing assessments. That lack of inclusivity on grading rubrics used to assess writing underrepresent the construct of writing for Black students as they only adhere to the demonstration of writing for SAE speakers. This violation of validity contributes to the writing scores potentially being an inaccurate representation of the writing ability for Black students or anyone who is not proficient in SAE. Because of this, it is important that research is done to investigate the effects of AAVE use on essay scores so grading rubrics can be developed to be more inclusive of individuals from diverse linguistic backgrounds. The goal of this research is to provide a potential answer for why inequities may exist for Black students and to contribute to the conversation of equity in assessment. Everyone deserves to have an equal chance of being scored in an equitable way, so the goal of this research is to highlight one area in which Black

students are not being scored equitably. This is one way to start the conversation towards

more equitable assessments of writing ability.

Additionally, penalizing and providing negative feedback towards AAVE has

adverse consequences for Black students. Thomas (2018) provided first-hand accounts of

negative experiences she endured with AAVE using an autoethnography depicting five

crucial periods in her linguistic development. She felt a decrease in sense of belonging,

silence, confusion in her identity, and internalization. Another student's account of the

penalization of their writing as a result of their use of AAVE led to them not wanting to

use writing to express themselves anymore out of fear of being penalized. The student

recounted the experience as being "bad" for her writing (Ball, 1996).

This differential penalization of AAVE and praise of SAE reinforces that the

former is wrong while the latter is right. One leads to academic success while the other

leads to failure. This contributes to a way in which white supremacy manifests in an

academic space. This differential penalization disproportionately impacts Black students.

The final consequence is possible mis-assessment. According to the Oakland Ebonics

Controversy, not acknowledging AAVE as the language of Black students contributed to

consequences such as their being overrepresented in special education classes, as their

ability was not adequately measured which led to incorrect inferences of academic

standing (Harris & Schroeder, 2013). Additionally, because the EPT assumed SAE

proficiency, it also contributed to the over representation of incoming college students as

needing remedial classes (Inoue, 2015).

The final purpose for this study is a reiteration of what I said in my introduction.

This study was motivated by my negative experiences in the education system as a Black

woman. I want to make classrooms more inclusive of Black students as it relates to their experiences, language, and performance because I want each generation to have a better academic experience than I did. I am committed to challenging the societal power structures that have painted the picture that Black knowledge is inferior because it looks and sounds different. Different is not equivalent to deficient.

In the section that follows, I detail the participants, procedures, and data analyses utilized in this study.

Chapter 3: Methods

Participants

The participants of this study were selected by The Learning Lab Agency, a nonprofit corporation in Arizona. The participants are students from the following two samples: Indiana Department of Education and The Georgia Virtual Academy. For this study I used only the students from the Indiana sample because it had the larger sample size.

The Indiana sample consisted of 22,536 students enrolled in sixth (N=3,116; 13.83%), eighth (N=4,087; 18.14%), or tenth grade (N= 15,332; 68.04%). Of the students, 11,597 identified as male (51.56%); 599 were English Language Learners (2.75%); 8,499 were identified as economically disadvantaged (37.71%) ;15 students identified as American Indian/Alaskan Native (.07%), 95 identified as Asian/Pacific Islander (.42%), 1,243 identified as Black/African American (5.52%), 486 identified as Hispanic/Latino (2.16%), 519 identified as two or more races/other (2.3%), and 20,177 identified as white (89.54%). All participants lived in the state of Indiana.

Exclusion Criteria

For this thesis, a secondary data analysis was conducted on the data obtained from The Learning Lab Agency. To align the data analyses with the purpose of the current study, the data were screened to only include individuals who identified as Black or white and who submitted an essay. An essay response was needed in order to identify any features of AAVE in the essay. This reduced the Indiana sample from 22,536 to a total of

21,420 students, where only one participant was removed for not responding to the essay

prompt, while the remainder were removed if they did not identify as Black or white.

As mentioned previously, a total of 21,420 students were included in the analyses

of the Indiana essays. Of the students, 94.2% identified as white, 51.26% identified as

male, and 63.68% were not identified as economically disadvantaged (Table 1).

Table 1

Demographics of the Reduced Indiana Sample

	Total	Percentage of Total
Economic Disadvantage		
Yes	7,779	36.32
No	13,641	63.68
Gender		
Female	10,441	48.74
Male	10,979	51.26
Race		
Black	1,243	5.8
white	20,177	94.2

Procedure

Data from students in Indiana were collected as a part of the 2015-16 to 2018-19

statewide standardized assessment referred to as the Indiana Statewide Testing for

Educational Progress-Plus (ISTEP+). These assessments were timed and administered on

a computer and aimed to measure students' level of mastery of skills such as reading,

writing, and mathematics. For the purpose of this study, the assessment of students'

writing was the focus. There were no inclusion or exclusion criteria followed during data

collection. If students were enrolled in school in Indiana, they were required to

participate in the ISTEP+ at the end of each school year from third grade to eighth grade

and one time in high school. The demographic variables were obtained via self-report.

Measures

Description of Writing Tasks

      Prompts. Students were instructed to select one of five prompts to create an

argumentative essay response. Before answering the prompt, students were instructed to

read the article corresponding to the question they would be answering. For each prompt,

students were told to state their stance on the topic, provide appropriate evidence to

support their claim, and format their argument into an introduction, a body, and a

conclusion. The following prompts were provided for students to choose from:

1. In the article "Making Mona Lisa Smile," the author describes how a new

   technology called the Facial Action Coding System enables computers to

   identify human emotions. Using details from the article, write an essay

   arguing whether the use of this technology to read the emotional expressions

   of students in a classroom is valuable.

2. In the article "Driverless Cars Are Coming," the author presents both positive

   and negative aspects of driverless cars. Using details from the article, create

   an argument for or against the development of these cars.

3. In "The Challenge of Exploring Venus," the author suggests that studying Venus is a worthy pursuit despite the dangers it presents. Using details from the article, write an essay evaluating how well the author supports this idea.

4. You have just read the article, "A Cowboy Who Rode the Waves." Luke's participation in the Seagoing Cowboys program allowed him to experience adventures and visit many unique places. Using information from the article, write an argument from Luke's point of view convincing others to participate in the Seagoing Cowboys program.

5. You have read the article "Unmasking the Face on Mars." Imagine you are a scientist at NASA discussing the Face with someone who thinks it was created by aliens. Using information in the article, write an argumentative essay by convincing someone that the Face is just a natural landform.

Students were provided a space to do their prewriting/planning of their response and were then provided a separate space for their final draft. After completing their response students were encouraged to use the following editing checklist to review their writing:

1. Check your capitalization and punctuation.

2. Spell all words correctly.

3. Check for sentence fragments or run-on sentences.

4. Keep verb tense consistent.

5. Make sure subject and verb agree.

6. Use words according to the rules of Standard English.

7. Remember to paragraph correctly.

Rubrics. A holistic rating form was used to provide each essay obtained from ISTEP+ with a score between one (minimum) and six (maximum), with one indicating little or no mastery and six indicating clear and consistent mastery (Appendix A). An essay with a score of one was described as:

> "…severely flawed by one or more of the following weaknesses: develops no viable point of view on the issue or provides little or no evidence to support its position; the essay is disorganized or unfocused, resulting in a disjointed or incoherent essay; the essay displays fundamental errors in vocabulary and/or demonstrates severe flaws in sentence structure; the essay contains pervasive errors in grammar, usage, or mechanics that persistently interfere with meaning."

An essay with a score of six was described as:

> " …effectively and insightfully develops a point of view on the issue and demonstrates outstanding critical thinking, using clearly appropriate examples, reasons, and other evidence to support its position; the essay is well organized and clearly focused, demonstrating clear coherence and smooth progression of ideas; the essays exhibits skillful use of language, using a varied, accurate, and apt vocabulary and demonstrates meaningful variety in sentence structure; the essay is free of most errors in grammar, usage, and mechanics."

AAVE Variable

To determine which AAVE features were most common, I consulted studies that looked at the writing behavior of AAVE speakers and identified which features of AAVE were present in their writing (Ball, 1997; Ivy & Masterson, 2011; Johnson, 2013;

Thompson, Craig, & Washington, 2004; Wheeler, 2019). All studies identified the 12

features listed in Table 1 as the ones most likely to be seen in an AAVE speaker's

writing. Once I identified the 12 commonly found AAVE markers in the writing of

AAVE speakers (Table 1), I developed a computer program to identify these features in

the essays. NLP was used to develop rules (or code) to quantify these features using the

Jupyter (version 6.3) and Spyder (version 4.2.5) interface of Python version 3.8.

The NLP program went through each essay and identified any usage of AAVE

that aligned with the 12 features from Table 1. Any time a feature of AAVE was used in

the essay, it was labeled using one of the 12 markers (i.e., multiple negation, ain't, etc.) to

create a variable that contained a list of the AAVE features used in each essay. For each

essay, the occurrences of the markers were then counted and a total AAVE variable was

computed. This variable was then used as the AAVE variable specified in the

hypothesized path model. Because the AAVE variable accounted for each occurrence of

AAVE, it was considered unbounded. However, in the data used for the current study, the

maximum number of occurrences of AAVE was 12.

Table 2

Markers of African American Vernacular English

AAVE Markers/Features	AAVE Example	SAE Equivalent
Copula Absence of regular be (is and are) (Ivy & Masterson,2011; Johnson, 2013)	She going	She is going
Completive done (Johnson, 2013)	He **done** ate all the food	He ate all the food
Multiple negation (Wheeler, 2019; Johnson, 2013)	I **don't** need **no** help on my homework	I don't need any help on my homework
Indefinite article -a	Do you have **a** umbrella?	Do you have an umbrella?
Habitual/Invariant "be" (Johnson, 2013; Mitri & Terry, 2014)	She **be** walking to school	She walks to school frequently
Absence of regular past tense -ed (Ivy & Masterson, 2011); Johnson, 2013)	They **play** yesterday	They played yesterday
Ain't (Mitri & Terry, 2014)	I **ain't** get your message	I didn't get your message
Absence of be auxiliary (Ivy &Masterson, 20ll)	He tired	He is tired
"g" dropping (Mitri & Terry, 2014)	She is **playin** (playing) outside	She is playing outside

Consonant cluster reduction (Mitri & Terry, 2014)	"The pirate has a treasure **[ches]** (chest)" (Mitri et at., 2014, p.57)	
Devoicing final consonants (Mitri & Terry, 2014)	Cold- "colt" Bed- "bet" Desk- "des"	
Zero possessive (deletion of -s) (Mitri & Terry, 2014)	We went to **grandma** house	We went to grandma's house

Data Analyses

Power Analysis

Prior to running any analyses on the essays, I conducted a power analysis using M*plus* 8.4 to determine the sample size needed to obtain adequate power (.8) for detecting the relation between AAVE use and writing scores. As sample size increases, power increases, so I expect to have high power given the large sample size (N= 21,420). However, because I am creating the AAVE variable and there has not been research regarding the magnitude of the parameter estimate between AAVE and writing score, I conducted a power analysis with a focus on that specific path. Because the magnitude of that parameter is unknown, but expected to be negative, I varied its unstandardized value between -.1 and -.9, in increments of .1. For each value of that specific parameter estimate, I started the sample size at 150 and increased it by 50 until I reached a power level of .8. A table for the sample size required for each value of the parameter estimate to obtain a power value of .8 or above is provided below (Table 3). As the magnitude of the parameter increased, the sample size requirement to obtain adequate power decreased.

The largest sample size needed was 800 for a parameter estimate of -.1. Given this, I

anticipate adequate power for the parameter given the large sample size even if the

parameter estimate is not large in magnitude.

Table 3

Power Analysis for Unstandardized Parameter Estimates of -.1 to -.9 for AAVE to

Writing Score Path

Parameter Estimate: AAVE | Writing Score	N	Power	Average Standard Error
-.1	800	.802	.004
-.2	250	.891	.063
-.3	100	.837	.100
-.4	65	.880	.120
-.5	40	.860	.160
-.6	25	.810	.20
-.7	20	.825	.224
-.8	15	.803	.260

-.9	15	.863	.320

Path Model Estimation

To examine the hypothesized relationships between race, gender, economic disadvantage, AAVE use, and writing score (Figure 2), I conducted a path model analysis in M*plus* Version 8.4 (Muthén & Muthén, 1998-2017). I used a path model because all the variables could be directly observed in the data. One requirement of a path model to be identified is that the degrees of freedom are greater than or equal to zero. This is a necessary, but not sufficient, condition for identification of the model. Sufficient conditions for identification of path models are that the model is recursive (e.g., arrows are unidirectional) and there are no correlated residuals. The path model in Figure 1 is recursive, has no correlated residuals, and is just identified with zero degrees of freedom. The exogenous variables (race, gender, and economic disadvantage) were categorical and cannot be assumed to be normally distributed. However, there is no assumption of normality for exogenous variables in path models, only endogenous variables. The endogenous variables, AAVE and writing scores, were treated as continuous variables in the path model. Univariate normality of the two endogenous variables was assessed based on values of skewness and kurtosis. Of the two endogenous variables, AAVE use was not normally distributed (skewness=4.43, kurtosis=30.63), but writing scores were (skewness=-.11, kurtosis=-.11)). Outliers were assessed using Mahalanobis distance and are discussed further in the results section. Due to the presence of non-normality in the endogenous variables, I used Maximum Likelihood estimation with robust standard

errors (termed MLR in M*plus*) to estimate parameters as this method adjusts estimates of

standard errors and values of fit statistics for non-normality.

Because my model is saturated, I did not evaluate fit, as saturated models fit the

data perfectly. It is hypothesized that the relationship between race and writing score is

partially mediated by AAVE use. A mediated effect is also referred to as an indirect

effect from the independent variable to the dependent variable by way of a third variable

(the mediator; Chaplin, 2007). In a basic mediation model as implemented in regression,

the simple relationship between the independent variable (IV) and the dependent variable

(DV) without the presence of a mediator is analyzed first and the path is referred to as (a)

or (β1). Then, the path between the IV and the DV while controlling for the mediator is

(a`) or (β3), the path from the IV to the mediator is (b) or (β2), and the path from the

mediator to the DV is (c) or (β4) are analyzed next (Chaplin, 2007, p. 622). I provide a

visual of this relationship in Figure 2. Mediation can be either full or partial. A partially

mediated relation indicates that the relation between the IV and the DV is only partially

explained by the mediator through paths b and c; the remaining part of the relation must

be represented by a direct from the IV to DV (a'). If [a'] is zero, then the relation between

the IV and DV is explained by the indirect effect via [b] and [c]; this is said to be a fully

mediated effect. I did not anticipate that the relationship between race and writing would

be fully explained (or fully mediated) by AAVE use, so AAVE was hypothesized to

partially mediate the relationship between race and writing score. The fully and partially

mediated models are shown in Figures 3 and 4, respectively. Paths shown in red indicate

those involved in the mediated relationship between race and writing via AAVE.

Figure 2

Basic Mediation Model
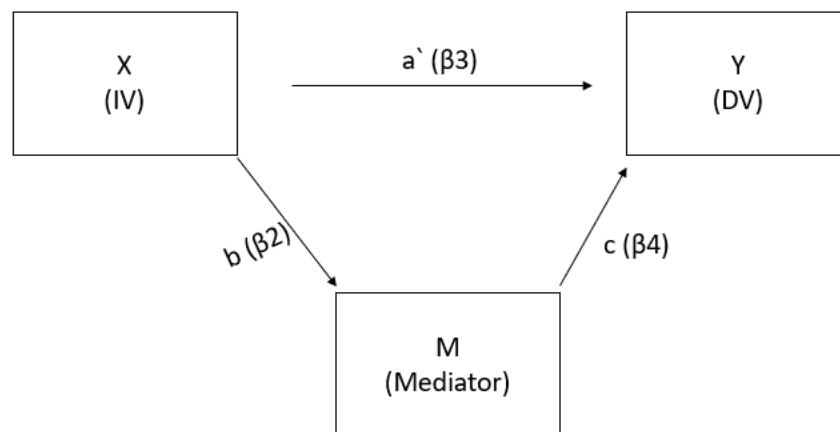
Because the current study is using methods of SEM, testing for the mediated

effect will be done by statistically comparing the chi-square values obtained from the

partially and fully mediated models. The most complex model, the partially mediated

model, is just-identified and must fit perfectly. For the fully mediated model, the direct

path from race to writing was fixed to zero, resulting in one degree of freedom. If the chi-

square for this model is significant then the partially mediated model fits better than the

fully mediated model, indicating the direct path from race to writing is needed to

reproduce the observed relations.

Figure 3

Hypothesized Partially Mediated Relationship of Race with Writing Scores
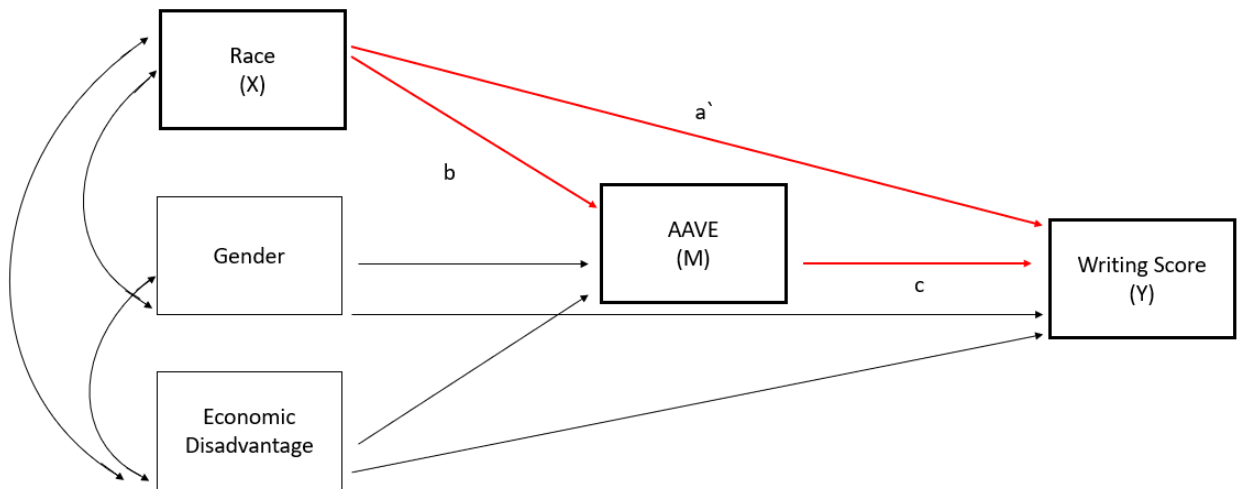
Figure 4

Hypothesized Fully Mediated Relationship of Race with Writing Scores
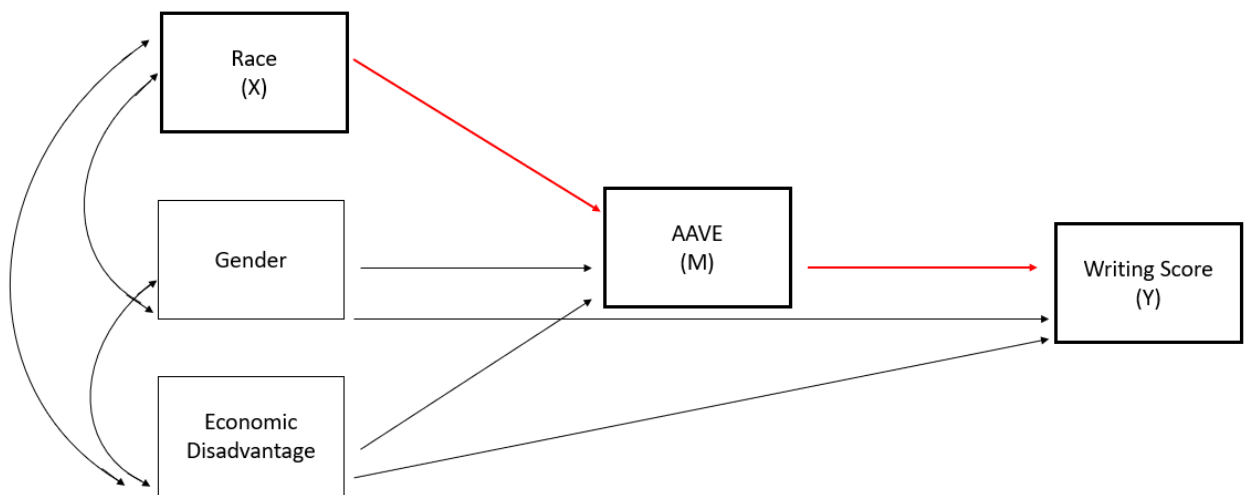
Chapter 4: Results

Descriptive Statistics

Prior to analyzing the model, descriptive statistics were obtained using *SAS*

Version 9.4 and are reported in this section. Tables 4 and 5 show the differences between

the descriptive statistics associated with AAVE frequency, writing score, gender, and

economic disadvantage for Black and white students. The sample consisted of mostly

white participants (N=20,177; 94.2%) and 1,243 (5.8%) Black participants. Of the

sample, 10,441 (48.74 %) identified as female and 10,979 (51.26 %) identified as male.

The sample had 7,779 (36.32%) participants who were identified as economically

disadvantaged, and the remainder were not (N=13,641; 63.68%). Overall, the sample had

a mean score of 3.32 (SD=1.03; Minimum: 1, Maximum=6, skewness=-.11, kurtosis=-

.11) on the Indiana holistic writing rubric. On the computed AAVE variable the mean for

all participants was .219 (SD= .668; Minimum= 0, Maximum=12; skewness=4.43,

kurtosis=30.63). Based on the graph in Figure 5, it is worth noting that the distribution of

the AAVE variable did not drastically differ between groups and most of participants had

zero occurrence of features of AAVE.

Table 4

Mean Writing Scores and AAVE Differences for Black and White Students

	Black	White
Holistic writing score	2.34	3.38
AAVE	.28	.22

Note. Holistic writing scores can range from 1 to 6. AAVE scores were computed by

NLP and can range from 0 to 12. Both groups significantly differed on each variable.

Table 5

Percentage of Gender and Economic Disadvantage by Race

	Black (N= 1,243)	White (N= 20,177)
Economically Disadvantaged		
Yes	73.61%	34.02%
No	26.39%	65.98%
Gender		
Male	54.95%	51.03%
Female	45.05%	48.97%

Figure 5

Distribution of AAVE Variable by Race
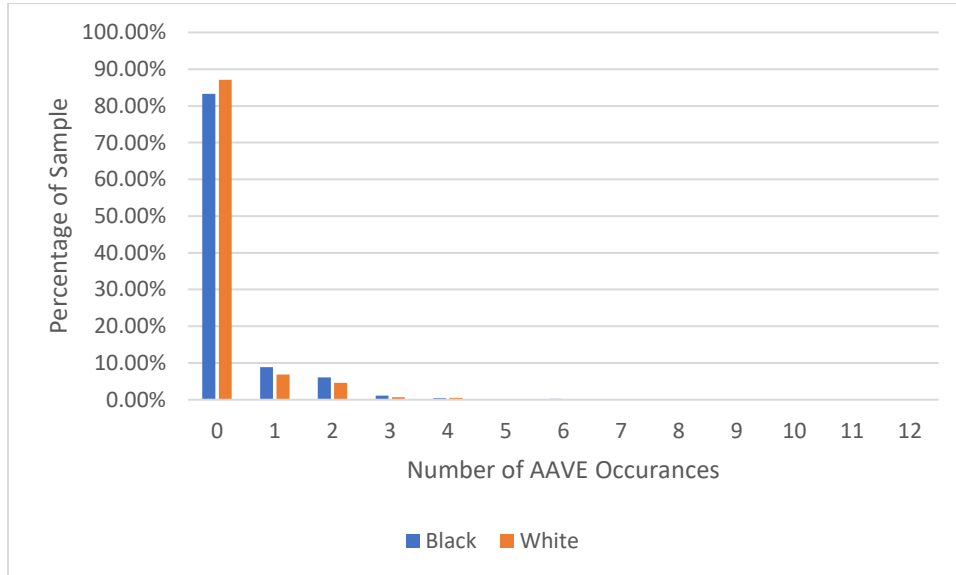
Data Screening

The interactions between race, gender, economic disadvantage when predicting

AAVE and writing scores were negligible, so it was assumed the relationships among the

variables would be best represented using a main effects only model (e.g., path model).

Based on the correlations depicted in Table 6, there appeared to be no issues of

multicollinearity among the predictors as no correlations exceeded .8. Univariate

normality was assessed using values of skewness and kurtosis for the two endogenous

variables: AAVE (skewness=4.43, kurtosis=30.629) and writing score (skewness=-.11,

kurtosis=-.11). Based on the values for skewness and kurtosis, AAVE violated the

assumption of normality due to exceeding a skewness value of |3| and a kurtosis value of

|10|. The data were also screened for outliers using Mahalanobis distance. Of the 21,420

students, there were 307 outliers (1.4%) that were significant at the .001 level. Because

theses outliers made up such a small percentage of the data, I did not remove them.

Table 6

Correlations Among Study Variables

	Holistic Writing Score	Race	Gender	AAVE	Economically Disadvantaged
Holistic Writing Score	1				
Race	-.24***	1			
Gender	-.17***	.018**	1		
AAVE	.02**	.02**	.02***	1	
Economically Disadvantaged	-.25***	.19***	.01	.02**	1

* <.05, ** <.01 , *** <.001

Model Estimation

The path model based on the study's hypotheses was analyzed using M*plus*

Version 8.4 (Muthen & Muthen, 1998-2017). Because of the nonnormality of the AAVE

variable (see Figure 6), the path model was initially analyzed using Maximum Likelihood

estimation with robust standard errors (MLR). I compared the parameter estimates and

standard errors obtained from MLR to those from Maximum Likelihood (ML) estimation.

They were no differences in parameter estimates, standard errors, and significance levels

between the two estimation methods, so here I only describe the results from the ML

estimation method.

Additionally, in the M*plus* program the default specification for path models is to

set all exogenous variable covariances to zero. To determine whether this would affect

values of the other model parameters, I ran models both with and without these

covariances. Results indicated the covariance among the three variables was negligible,

as can be seen in Table 6, and including these covariances did not affect other parameter

estimates. I therefore report estimates obtained without inclusion of the exogenous

variable covariances.

Model Results

In the following sections I will describe the results of the partially and fully

mediated models separately and conclude with a comparison of the two models using a

Chi-Square Difference test. For both models, I report the unstandardized parameter

estimates along with their standard errors.

Partial Mediation Model Results

Values of Direct Paths

All the path values in the hypothesized partially mediated model were statistically significant ($\chi^2(0) = 0$; Figure 6). Parameter estimates and standard errors (in parentheses) are shown in Figures 6 and 7. Given the small magnitude of some of the parameter estimates, it is important to note that this significance may be due to the large sample size (N=21,420). Race (b=.048, 95% CI [.009, .087], p =.015), gender (b=.031, 95% CI [.013, .048], p = .001), and economic disadvantage (b=.026, 95% CI [.007, .044], p =.008) positively predicted frequency of AAVE use. Black students, those who identified as males, and those who were economically disadvantaged had higher frequencies of AAVE use in comparison to white students, those who identified as females, and those who were not economically disadvantaged, respectively. Race (b=-.850, 95% CI [-.906, -.794], $p <$ .001), gender (b=-.341, 95% CI [-.367, -.316], $p <$ .001), and economic disadvantage (b= -.453, 95% CI[-.481, -.426], $p <$ .001) also negatively predicted writing scores. Black students, those who identified as males, and those who were economically disadvantaged received overall lower writing scores. Contrary to the hypothesized negative effect, AAVE positively predicted writing scores (b=.05, 95% CI [.030,.069], $p <$ .001). According to these results, as AAVE increased writing score increased.

Values of Indirect Paths

Race (b=.002, 95% CI [.000,.005], p = .028), gender (b=.002, 95% CI [.000,.003], p = .005), and economic disadvantage (b=.001, 95% CI [.000,.002), p = .0195) had significant positive indirect effects on writing scores through AAVE. Based on these

results, Black students had higher levels of AAVE which led to higher writing scores.

Similarly, males and students who were economically disadvantaged had higher levels of

AAVE, leading to higher writing scores.

R^2 values

In terms of practical significance, the model did not explain a substantial amount

of variance in the endogenous variables. The explained variance in writing score was

12.7% and .01% in AAVE use**.**

Figure 6

Partially Mediated Model Showing Parameter Estimates and Standard Errors
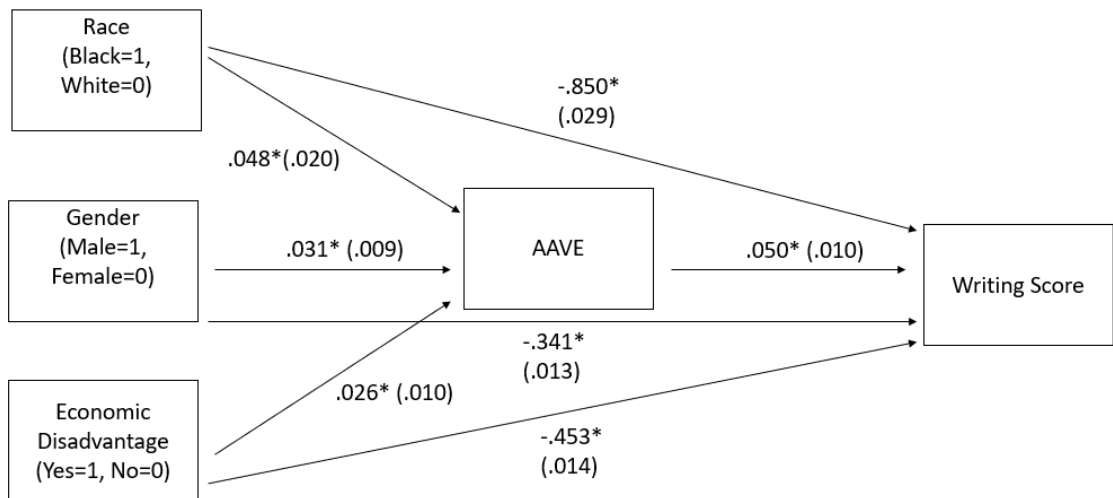
Full Mediation Model Results

Values of Direct Paths

As in the partial mediation model, the path values of the fully mediated model were all statistically significant ($\chi 2(1) = 860.985$, $p < .01$; Figure 7). However, the significant chi-square value indicated that this model was misspecified due to fixing the relationship between race and writing to zero. Due to this misspecification, the parameter estimates were biased and are therefore not interpreted.

R_2 Values

Although the R-square value of AAVE use did not change from the partially to the fully mediated model, the R-square value for the amount of variance explained in writing score decreased when the fully mediated model was specified. The fully mediated model explained about 9.1 % of the variance in writing score (compared to 12.7% for the partially mediated model) and .01% of the variance in AAVE use.

Figure 7

Fully Mediated Model Showing Parameter Estimates and Standard Errors
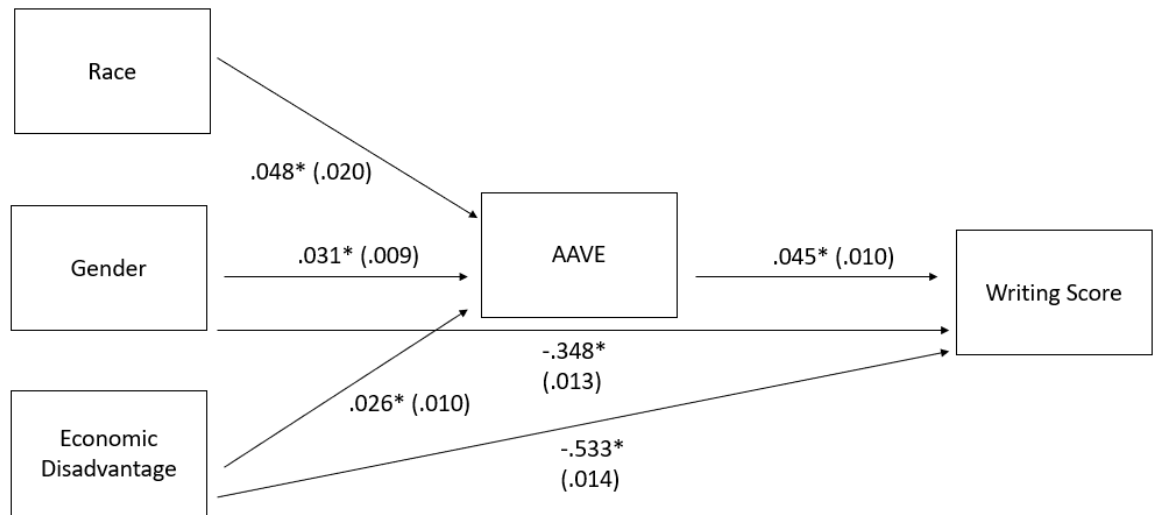
Test of Partial v. Full Mediation

To test whether the relationship between race and writing score was fully or partially mediated by AAVE use, a chi-square difference test was used to determine whether the partially mediated model provided a significantly better fit to the data than the full mediated model. The chi-square difference test determined that the overall fit of the partially mediated model with one additional path was statistically better than that of the fully mediated model, χ^2_D (1) = 860.985, $p < .001$. Given the results, the hypothesized partial mediation of race to writing score through AAVE use was supported.

Chapter 5: Discussion

The focus of this study was on one facet of the education system—writing assessment. I used CRT as the framework to understand how race and language influence writing performance, and in a broad sense, the education system. I further explored how AAVE speakers interact with writing assessments and how their language may impact how their performance is perceived through the scores received. To do so, I used NLP to get a better understanding of the prevalence of language differences in students' writing by tagging features of AAVE in essays retrieved from a standardized assessment. NLP tagged the features of AAVE I specified in the essays and created the AAVE variable depicted in my model. I then hypothesized a path model to investigate how these different levels of AAVE affected students' writing scores through methods of SEM. The primary relationships of interest in this study were those between race, AAVE, and writing scores.

The results of the Johnson and VanBrackle (2012), Kynard (2008), and Richardson (1997) studies led me to hypothesize that essays with more frequent occurrences of AAVE would result in lower writing scores. At the conclusion of my literature search there was no specific literature linking students' use of AAVE to their lower writing scores on standardized writing assessments, but it was evident that AAVE had a negative relationship with writing performance (Johnson & VanBrackle, 2012; Kynard, 2008; Richardson, 1997) in elementary and middle school classrooms, and in U.S. college English composition courses (Baker-Bell, 2020a).This emphasized a need for the current study as a means to look into how AAVE may impact perceived writing performance on standardized assessments. The hypothesized negative relationship between AAVE and

writing scores was not supported, however. Although the relationship was positive, the estimated relationship between AAVE and writing can be considered ignorable due to the small magnitude of the parameter (b=.05). A potential explanation for this difference could be that students in the current study were taking a standardized assessment rather than participating in classroom tasks. Because of this, it is possible that students were more likely to use AAVE in the classroom rather than on a standardized assessment. Additionally, the essays in the current study were obtained from students in grades six, eight, and ten. It can be assumed that students at this age have been in the education system for a sufficient amount of time that they have encountered penalization of AAVE in their English classes prior to taking this assessment. Thus, the students in this study were not likely to use features of AAVE in their essays, resulting in restriction of range in the AAVE values. This restriction of range in turn produced a small total relation between AAVE and writing. In the following paragraphs, I describe potential explanations for the restriction of range in the AAVE variable.

One "solution" AAVE speakers are exposed to as a means to "fix" their language expression is a concept called code-switching. Teachers encourage students to code-switch as a way "to assess the needs of the setting (the time, place, audience, and communicative purpose) and intentionally choose the appropriate language style for that setting" (Wheeler, 2008, p. 57). Students are taught that the appropriate language for academia and formal settings is SAE. Because the writing assessment in this study was part of the statewide standardized assessment, it is quite likely that students viewed this standardized writing assessment as formal writing which therefore required them to switch from AAVE to SAE when articulating their argument. In fact, this idea was

emphasized by the instructions accompanying the assessment, which included an

additional reminder via an editing checklist that students should "use words that align

with Standard English". I would like to note that code-switching is a problematic

approach to dealing with AAVE, but it is a potential explanation for the restriction of

range in the AAVE variable and the unexpected path value. Because of students'

probable use of code-switching, there was little use of AAVE, as evidenced by the low

amount of AAVE features found in this study. Based on the frequencies of the AAVE

variable, 83.27% of Black participants and 87.10% of white students had zero

occurrences of AAVE (M=.22, SD= .67). The low variance of the AAVE variable likely

attenuated its correlation with writing scores, which could have contributed to the low

path value.

Because Black students are more likely to receive lower writing scores and are

most likely to be speakers of AAVE (Lippi-Green, 1997), I hypothesized that one

explanation for the relationship between race and writing scores is Black students' use of

AAVE in their writing. As mentioned previously, it was not evident that previous

empirical studies linked the inequities in writing assessments to students' use of AAVE,

but based on the negative perception of AAVE (Harris & Schroeder, 2013; Baker-Bell,

2020b;2020a; Wheeler, 2010) and penalization of AAVE in the classroom (Ball, 1996;

Charity et al., 2004; Wheeler, 2012), it was a possible explanation for inequities seen in

writing performance. Based on this, I expected the relationship between race and writing

scores to be partially mediated by use of AAVE. Based on the results of the chi-square

difference test, the partial mediation model resulted in better fit to the data in comparison

to the fully mediated model. In other words, the model that best represented the data was

one in which I allowed race to relate directly to writing while simultaneously allowing

race to be indirectly related to writing through AAVE use. When a partially mediated

model is supported, it represents that that there are other potential mediators that

contribute to the relationship between race and writing scores. I believed that the relation

between race and writing was explained by AAVE to some extent, but it is likely that

writing could be further explained by other variables not in my model or measured in the

dataset. It is worth noting that future studies should continue to investigate other

mediating factors of the relationship between race and writing scores. However, the

argument that the AAVE does not affect writing scores is weak one because a) literature

supports the relationship between AAVE and writing scores, b) the small parameter

estimates are likely due to sampling error around zero, and c) it is possible that the

AAVE variable did not capture all features of AAVE.

It is possible that NLP did not measure AAVE as effectively as expected. The

distribution of AAVE was almost identical between Black and white students. The most

frequent feature of AAVE was the indefinite article -a (e.g., using "a" instead of "an"

before a word that begins with a vowel). It is possible that this was a common mistake

that students made in their writing rather than a distinctive marker of AAVE. So, students

that were tagged to have used features of AAVE may not have actually been using

AAVE in their writing. It is possible that more distinctive features of AAVE need to be

identified to further distinguish between SAE and AAVE as these two languages have

some overlap. To check the accuracy in AAVE identification, I went through 100 essays

to see if essays tagged for using AAVE were actually using features of AAVE. Of the

100 essays, 73% were tagged correctly for containing AAVE which indicated the

possibility of construct underrepresentation and construct irrelevant variance in the AAVE variable.

In addition to the primary hypotheses discussed above, I also investigated a few secondary hypotheses related to the other variables in the model. Based on the studies by Cronley et al. (2017), Inoue (2015), Mitri and Terry (2004), and Wheeler (2019; 2012), I expected that Black students would have lower writing scores than white students. In alignment with the previous studies, it was found that, in comparison to white participants (M= 3.38), Blacks students obtained lower writing scores (M= 2.34). Based on studies by Lippi-Green (1997), Fitton et al. (2012), and Wainwright (2018) it was expected that Black students, those who identified as males, and those who were economically disadvantaged would have higher frequencies of AAVE use. The results aligned with the previous research and these hypotheses were supported.

Based on studies by Cronley et al. (2017), Reynold et al. (2015), and Reilly et al. (2019), it was hypothesized the those who identified as female would have higher writing scores. The current study concluded that gender was negatively related to writing score, with those who identified as female having higher writing scores in comparison to those who identified as male. Additionally, studies by Charity et al. (2004) and Kim et al. (2001) supported the hypothesized negative relationship between economic disadvantage and writing scores. In the current study, those who were economically disadvantaged had lower writing scores than those who were not economically disadvantaged. Finally, race, gender, and economic disadvantage were all significantly related to writing scores indirectly through AAVE use.

Implications

Overall, Black students obtained lower scores on the writing assessment of the ISTEP+. The result in this study was consistent with the literature (Cronley et al., 2017; Inoue 2015; Mitri &Terry 2004; Wheeler 2010; 2012). It is important to consider how these inequities on writing assessments are problematic due to their consequences for students' performance on writing assessments as many consequential tests require a written component (GRE, SAT, ACT, EPT, etc.). Lower scores on these tests are a likely reason for Black students' higher rejection rates to programs in higher education. There was a total of 16.3 million students enrolled in undergraduate institutions in 2016 and of this total, 2.2 million were Black and 9.1 million were white (de Brey et al., 2019). According to the National Center of Education Statistics, of the students awarded doctoral degrees in 2015-16, only 9% were Black and 68% were white (de Brey et al., 2019). It is critical to continue to question why these disparities exist in undergraduate enrollment and doctoral degree attainment. A common factor contributing to access to these forms of education is the presence of standardized assessments, most of which include a written component. Black students are required to complete these consequential writing assessments in a language that is not intuitive to them and does not align with their linguistic background, which may likely affect their performance.

Although the findings did not support the hypothesized negative relationship between AAVE and writing scores, there was evidence that students were using some features of AAVE in their writing (M=.219, SD= .668). There was also evidence that SAE was the preferred language in this setting. This was evident in the editing checklist where students were instructed to use "Standard English." Another way in which SAE

preference was evident was in the requirement for students to read a passage in SAE prior
to completing the prompt. The prompt then instructed participants to center their
argument around and utilize evidence from the passage to strengthen their argument. The
fact that this passage aligned with the words, style, and grammatical structure of SAE
provided a cue to students that this language was required in this setting.

I urge readers to ponder why SAE is the language preferred in academia. This
preference contributes to the belief that AAVE is wrong, inappropriate, discouraged, and
unacceptable in academia. This was illustrated specifically in this study in the written
portion of the ISTEP+, but is also the case for virtually any standardized assessment (see
Fisher & Lapp, 2013). The preference for SAE is representative of how AAVE is viewed
as a language problem that needs to be fixed (i.e., by code-switching) for students to be
academically successful (see *Learning to Talk Like the Test: Guiding Speakers of AAVE*;
Fisher & Lapp, 2013). This requirement to leave one's language at the door of the
classroom is an erasure and rejection of a group of people because language is indicative
of someone's entire identity. Because of this, it is important that we use the results of this
and similar studies to continue to justify the instatement of cultural sustaining pedagogies
and antiracist assessment practices (see Baker-Bell, 2020a; Inoue 2015; Ladson-Billings,
1992; Poe, Inoue, & Elliot, 2018; Randall et al., 2021; Randall, 2021).

Limitations

There were a few limitations that are important to bring to the reader's attention.
First, there were unbalanced samples of Black (*N*=1,243) and white participants (*N*=
20,177). Because of this, there is a possibility that the tested model is more representative

of how AAVE relates to white participants due to the majority of participants identifying as white. My second limitation was the non-normality of the AAVE variable. The most frequent number of AAVE features in an essay was 0, which resulted in a positively skewed distribution. This restriction of range could have influenced the parameter value obtained for the AAVE to writing path. Another potential contributor to the distribution of this variable was the editing checklist participants were encouraged to use after completing their essay. One of the components on the checklist instructed participants to check whether they had "use[d] words that align with Standard English", which likely decreased the amount of AAVE features found in the essay. Another limitation was that the path model assumed a linear relationship between AAVE and writing score, but it is possible that this relationship was nonlinear. Finally, there was the geographical restriction of participants. All participants were from Indiana, which limits the generalizability of the results.

Future Directions

Because AAVE may vary across informal and formal settings, it would be worth investigating not only how AAVE operates on standardized writing assessments but also how it operates in classroom writing tasks to compare the frequency of AAVE use across settings. Additionally, as mirrored in the results, Black students are performing significantly lower on writing assessments. Future research should continue to investigate these inequities to determine what factors are contributing to these lower scores. I hope this thesis fuels the ongoing conversation surrounding equity in assessment and encourages mores studies to utilize NLP as well as SEM in investigating the language and experiences of Black students. Finally, The AAVE variable used in this study was

limited to grammatical instances of AAVE. Future studies could expand the AAVE

variable to encompass dialect-specific spelling errors and use of personal pronouns in

storytelling to further investigate how linguistic racism may operate on writing

assessments.

References

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

American Psychological Association. (2021). *Inclusive language guidelines.* https://www.apa.org/about/apa/equitydiversity-inclusion/language-guidelines.pdf

Baker-Bell, A. (2020a). *Linguistic justice: Black language, literacy, identity, and pedagogy*. Routledge.

Baker-Bell, A. (2020b). Dismantling anti-black linguistic racism in English language arts classrooms: Toward an anti-racist black language pedagogy. *Theory Into Practice, 59*:1, 8-21D.OI: 10.1080/00405841.2019.1665415

Ball, A., & Lardner, T. (1997). Dispositions toward Language: Teacher Constructs of Knowledge and the Ann Arbor Black English Case. *College Composition and Communication*, *48*(4), 469. https://doi.org/10.2307/358453

Baldwin, J. (1979, July 29*). If Black English isn't a language, then tell me, what is?* The New York Times. https://archive.nytimes.com/www.nytimes.com/books/98/03/29/specials/baldwin-english.html?_r=1

Ball,A.F. (1996). Expository writing patterns of African American students. *English Journal,* 27-36.

Bell, D. (1990). Brown and the interest convergence dilemma. *In D. Bell (Ed.), Shades of Brown: New perspectives on school desegregation* (pp. 90- 106). New York: Teachers College Press.

Bell, D. (1992). *Faces at the bottom of the well*. New York: Basic Books.

Bridges, B.(n.d). African Americans and college education by the numbers.

https://uncf.org/the-latest/african-americans-and-college-education-by-the-

numbers

Chaplin, W.F. (2007). Moderator and mediator models in personality research. In R.W.

Robins, R.C. Fraley & R.F. Krueger (Eds.), *Handbook of research in personality*

psychology. The Guilford Press. New York.

Charity, A.H., Scarborought, H.S., & Griffin, D.M. (2004). Familiarity with school

English in African American children and its relation to early reading

achievement. *Child Development, 75*(5), 1340-1356.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd Edition).

Hillsdale, NJ: Lawrence Earlbaum Associates.

Cronley, C., Kilgore, C., & Daniels-Lerberg, T. (2017). A multivariate analysis of writing

is BSW case study papers. *Journal of Baccalaurete Social Work, 22*(1), 181-205.

DOI:10.18084/1084-7219.22.1.181

Crossley, S. A. (2013). Advancing research in second language writing through

computational tools and machine learning techniques: A research agenda.

Language Teaching, *46*(2), 256–271.

https://doi.org/10.1017/S0261444812000547

Crossley, S. A., Allen, L. K., Kyle, K., & McNamara, D. S. (2014). Analyzing discourse

processing using a simple natural language processing tool. *Discourse Processes*,

51(5–6), 511–534. https://doi.org/10.1080/0163853X.2014.910723

de Brey, C., Musu, L., McFarland, J., Wilkinson-Flicker, S., Diliberti, M., Zhang, A., Branstetter, C., & Wang, X. (2019). *Status and Trends in the Education of Racial and Ethnic Groups 2018* (NCES 2019-038). U.S. Department of Education. Washington, DC: National Center for Education Statistics. Retrieved [2022] from https://nces.ed.gov/ pubsearch/.

Delgado, R., & Stefancic, J. (2017). *Critical race theory: An introduction* (3rd ed.,). New York: New York University Press.

Dorn, R. (2019). Dialect-specific models for automatic speech recognition of African American Vernacular English. *Proceedings of the Student Research Workshop Associated with RANLP 2019*, 16–20. https://doi.org/10.26615/issn.2603-2821.2019_003

Fisher, D., & Lapp, D. (2013). Learning to talk like the test: Guiding speakers of African American Vernacular English. *Journal of Adolescent and Adult Literacy, 56*(8). doi:10.1002/JAAL.198

Fitton, L., Johnson, L., Wood, C., Schatschneider, C., & Hart, S. A. (2021). Language variation in the writing of African American students: Factors predicting reading achievement. *American Journal of Speech - Language Pathology (Online), 30*(6), 2653–2667. http://dx.doi.org/10.1044/2021_AJSLP-20-00263

Groenwold, S., Ou, L., Parekh, A., Honnavalli, S., Levy, S., Mirza, D., & Wang, W. Y. (2020). Investigating African-American Vernacular English in transformer-based text generation. *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 5877–5883. https://doi.org/10.18653/v1/2020.emnlp-main.473

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),193-204.

Inoue, A.B.(2015). The function of race in writing assessments. In A.B. Inoue, *Antiracist

writing assessment ecologies:Teaching and assessing writing for a socially just

future* (pp. 25 – 75). The WAC Clearinghouse. Parlor Press. DOI 10.37514/PER-

B.2015.0698

Ivy, L. J., & Masterson, J. J. (2011). A comparison of oral and written English styles in

African American students at different stages of writing development. *Language,

Speech, and Hearing Services in Schools*, *42*(1), 31–40.

https://doi.org/10.1044/0161-1461(2010/09-0069)

Jayakumar, U. M., & Adamian, A. S. (2017). The fifth frame of colorblind ideology:

Maintaining the comforts of colorblindness in the context of white fragility.

Sociological Perspectives, *60*(5), 912–936.

https://doi.org/10.1177/0731121417721910

Johnson, D.V. (2013). Effects of African American Vernacular English on English

language learners' academic writing [Master's Thesis, State University of New

York]. http://hdl.handle.net/20.500.12648/494

Johnson, D., & VanBrackle, L. (2012). Linguistic discrimination in writing assessment:

How raters react to African American "errors," ESL errors, and standard English

errors on a state-mandated writing exam. *Assessing Writing*, *17*(1), 35–54.

https://doi.org/10.1016/j.asw.2011.10.001

Kellogg, R.T.(2016). Language. In *Fundamentals of cognitive psychology* (pp.213-243).

Thousand Oaks: SAGE Publications

Kim, Y.S., Puranik, C., & Otaiba, S. A. (2015). Developmental trajectories of writing

skills in first Grade: Examining the effects of SES and language and/or speech

Impairments. *The Elementary School Journal, 115*(4), 593–613.

https://doi.org/10.1086/681971

Kline, R. B. (2016). *Principles and Practice of Structural Equation Modeling* (4th ed.)

(D.A. Wright & T.D. Little, Eds.). The Guilford Press.

Kynard, C.(2008). Writing while Black: The colour line, Black discourses and

assessment in the institutionalization of writing instruction. *English Teaching:*

Practice and Critique, 7(2), 4-34.

Ladson-Billings, G.(1992). Reading between the lines and beyond the pages: A culturally

relevant approach to literacy teaching. *Theory into Practice, XXXI* (4).

Ladson-Billings, G. (1998). Just what is critical race theory and what's it doing in a nice

field like education? *International Journal of Qualitative Studies in Education*,

11(1), 7-24. DOI: 10.1080/0951839982368638

Lippi-Green,R. (1997). What we talk about when we talk about Ebonics:Why definitions

matter. *The Black Scholar, 27*(2).

Matusda, P.K.(2006). The myth of linguistic homogeneity in U.S. college composition.

College English, 68(6), 637-651.

McClendon, G. (2014). High school principals perceptions on ebonics: A framework for

school leaders. *International Journal of Education and Social Science, 1*(4).

Mitri,S.M., & Terry, N.P. (2014). Phonological awareness skills in young African

    American English speakers. *Read Writ, 27*, 555-569. DOI 10.1007/s11145-013-

    9458-z.

Muthén, L.K. and Muthén, B.O. (1998-2017). *Mplus User's Guide. Eighth Edition*.

    Los Angeles, CA: Muthén & Muthén

Norton, B., & Starfield, S. (1997). Covert language assessment in academic writing.

    Language Testing, 14(3), 278-294. https://doi.org/10.1177/026553229701400305

Poe, M., Inoue, A.B., & Elliot, N. (2018). *Writing assessment, social justice, and the*

    advancement of opportunity. The WAC Clearinghouse; University Press of Colorado.

Pollock,K.E., & Meredith, L.H. (2001). Phonetic transcription of African American

    Vernacular English. *Communication Disorders Quarterly, 23*(1).

Randall, J.(2021). "Color-Neutral" is not a thing: Redefining construct definition and

    representation through a justice-oriented critical antiracist lens. *Educational*

    Measurement: Issues and Practice,0(0),1-9

Randall, J., Poe, M., & Slompe,D.(2021).Ain't oughta be in the dictionary: Getting to

    justice by dismantling anti-black literacy Assessment practices. *Journal of*

    Adolescent & Adult Literacy, 64(5), 594-599. doi: 10.1002/jaal.1142

Rashid, K. (2011). 'To break asunder along the lesions of race'. The Critical race theory

    of W.E.B. Du Bois. *Race Ethnicity and Education*, *14*(5), 585–602.

    https://doi.org/10.1080/13613324.2011.578127

Reilly, D., Neumann, D. L., & Andrews, G. (2019). Gender differences in reading and

    writing achievement: Evidence from the National Assessment of Educational

Progress (NAEP). *American Psychologist, 74*(4), 445–458.

https://doi.org/10.1037/amp0000356

Reynolds, M. R., Scheiber, C., Hajovsky, D. B., Schwartz, B., & Kaufman, A. S. (2015).

Gender differences in academic achievement: Is writing an exception to the

gender similarities Hypothesis? *The Journal of Genetic Psychology, 176*(4), 211–

234. https://doi.org/10.1080/00221325.2015.1036833

Richardson, E. (1997). Can you be black and write and right. *Alternatives to grading

student writing*, 63-75.

Rickford, J. (1999). The Ebonics controversy in my backyard: a sociolinguist's

experiences and reflections. *Journal of Sociolinguistics, 3*(2), pp.267-266.

Sautter, J. M., Thomas, P. A., Dupre, M. E., & George, L. K. (2012). Socioeconomic

status and the Black-White mortality crossover. *American Journal of Public

Health*, *102*(8), 1566–1571. https://doi.org/10.2105/AJPH.2011.300518

Sealey-Ruiz, Y.(2005).Spoken soul: The language of Black imagination and reality. *The

Educational Forum,20*, 37-46.

Smitherman, G. (2006). *Word From the Mother: Language and African Americans*. New

York, NY: Routledge.

Thomas, C. (2018) Negotiating words and worlds: an autoethnography of linguistic

identity development. *International Journal of Qualitative Studies in Education,

31*(7), 612-625, DOI: 10.1080/09518398.2018.1468044

Thompson, C. A., Craig, H. K., & Washington, J. A. (2004). Variable production of

African American English across oracy and literacy contexts. *Language, Speech,*

and Hearing Services in Schools, *35*(3), 269–282. https://doi.org/10.1044/0161-1461(2004/025)

Treiman, R., & Bowman, M.(2015). Spelling in African American children: the case of the final consonant devoicing. *Read Writ, 28*, 1012-1028. DOI 10.1007/s11145-015-9559.

Treiman, R.(2004). Spelling and dialect: Comparisons between speakers of African American Vernacular English and white speakers. *Psychonomic Bulletin & Review, 11*(2), 338-342.

Wainwright, A.B.(2019). Gender differences in the narrative productions of African American adults. *American Journal of Speech-Language Pathology, 28*, 623-638. https://doi.org/10.1044/2018_AJSLP-18-0153

Wheeler, R.S.(2008). Becoming adept to code-switching. *Educational Leadership*, 54-58.

Wheeler,R.S. (2019). Attitude change is not enough: Disrupting deficit grading practices to disrupt dialect prejudice*. Proc Ling Soc Amer, 4* (10), 1-12. https://doi.org/10.3765/plsa.v4i1.4505

Wheeler, R.S., Cartwright, K.B., & Swords, R. (2012). Factoring AAVE into reading assessment and instruction. *The Reading Teacher, 65*(5), 416-425.

Appendix A: Writing Rubric
Holistic Rating Form

After reading each essay and completing the analytical rating form, assign a holistic score based on the rubric below. For the following evaluations you will need to use a grading scale between 1 (minimum) and 6 (maximum). As with the analytical rating form, the distance between each grade (e.g., 1-2, 3-4, 4-5) should be considered equal.

SCORE OF 6: An essay in this category **demonstrates clear and consistent mastery**, although it may have a few minor errors. A typical essay effectively and insightfully develops a point of view on the issue and demonstrates outstanding critical thinking, using clearly appropriate examples, reasons, and other evidence to support its position; the essay is well organized and clearly focused, demonstrating clear coherence and smooth progression of ideas; the essay exhibits skillful use of language, using a varied, accurate, and apt vocabulary and demonstrates meaningful variety in sentence structure; the essay is free of most errors in grammar, usage, and mechanics.

SCORE OF 5: An essay in this category **demonstrates reasonably consistent mastery**, although it will have occasional errors or lapses in quality. A typical essay effectively develops a point of view on the issue and demonstrates strong critical thinking, generally using appropriate examples, reasons, and other evidence to support its position; the essay is well organized and focused, demonstrating coherence and progression of ideas; the essay exhibits facility in the use of language, using appropriate vocabulary demonstrates variety in sentence structure; the essay is generally free of most errors in grammar, usage, and mechanics.

SCORE OF 4: An essay in this category **demonstrates adequate mastery**, although it will have lapses in quality. A typical essay develops a point of view on the issue and demonstrates competent critical thinking, using adequate examples, reasons, and other evidence to support its position; the essay is generally organized and focused, demonstrating some coherence and progression of ideas exhibits adequate; the essay may demonstrate inconsistent facility in the use of language, using generally appropriate vocabulary demonstrates some variety in sentence structure; the essay may have some errors in grammar, usage, and mechanics.

SCORE OF 3: An essay in this category **demonstrates developing mastery**, and is marked by ONE OR MORE of the following weaknesses: develops a point of view on the issue, demonstrating some critical thinking, but may do so inconsistently or use inadequate examples, reasons, or other evidence to support its position; the essay is limited in its organization or focus, or may demonstrate some lapses in coherence or progression of ideas displays; the essay may demonstrate facility in the use of language, but sometimes uses weak vocabulary or inappropriate word choice and/or lacks variety or

demonstrates problems in sentence structure; the essay may contain an accumulation of errors in grammar, usage, and mechanics.

SCORE OF 2: An essay in this category **demonstrates little mastery**, and is flawed by ONE OR MORE of the following weaknesses: develops a point of view on the issue that is vague or seriously limited, and demonstrates weak critical thinking, providing inappropriate or insufficient examples, reasons, or other evidence to support its position; the essay is poorly organized and/or focused, or demonstrates serious problems with coherence or progression of ideas; the essay displays very little facility in the use of language, using very limited vocabulary or incorrect word choice and/or demonstrates frequent problems in sentence structure; the essay contains errors in grammar, usage, and mechanics so serious that meaning is somewhat obscured.

SCORE OF 1: An essay in this category **demonstrates very little or no mastery**, and is severely flawed by ONE OR MORE of the following weaknesses: develops no viable point of view on the issue, or provides little or no evidence to support its position; the essay is disorganized or unfocused, resulting in a disjointed or incoherent essay; the essay displays fundamental errors in vocabulary and/or demonstrates severe flaws in sentence structure; the essay contains pervasive errors in grammar, usage, or mechanics that persistently interfere with meaning.

Appendix B: Data Screening Code

```
3
4  ******************* ANALYSES POST NLP*******************;
5
6  proc import out=Indiana
7     datafile="C:\Users\nesbi\indiana_with_AAVE_totals.xlsx"
8     replace dbms=XLSX;
9     Getnames=yes;
10 run;
11
12 proc contents data=Indiana; run;
13
14 proc freq data=Indiana;
15  tables sum_aave_flags;
16 run;
17
18 proc freq data=Indiana;
19  tables race_ethnicity;
20 run;
21
22 proc print data=indiana(obs=10); run;
23
24 data Indiana_AAVE; set Indiana;
25  keep contains_aave economically_disadvantaged essay_word_count gender
26  grade indiana_grammar_score indiana_holistic_score prompt_name race_ethnicity sum_aave_flags;
27 run;
28
29 proc print data=indiana_aave(obs=10); run;
30

--
31 data Indiana_AAVE; set Indiana_AAVE;
32  if race_ethnicity= "American Indian/Alaskan Native" then delete;
33  if race_ethnicity= "Asian/Pacific Islander" then delete;
34  if race_ethnicity= "Two or more races/Other" then delete;
35  if race_ethnicity= "Hispanic/Latino" then delete;
36 run;
37
38 data Indiana_AAVE; set Indiana_AAVE;
39  if race_ethnicity= "Black/African American" then race=1;
40  if race_ethnicity= "White" then race=0;
41 run;
42
43 proc freq data=Indiana_AAVE; tables race_ethnicity * race; run;
44
45 data Indiana_AAVE; set Indiana_AAVE;
46  if economically_disadvantaged= "Not economically disadvantaged" then ED=0;
47  if economically_disadvantaged= "Economically disadvantaged" then ED=1;
48 run;
49
50 proc freq data=Indiana_AAVEb; tables economically_disadvantaged * ED; run;
51
52 data Indiana_AAVE; set Indiana_AAVE;
53  if gender="M" then new_gender= 1;
54  if gender="F" then new_gender=0;
55 run;
56
57 proc freq data=Indiana_AAVE; tables gender * new_gender; run;
```

```
76
77  *AAVE by race;
78
79⊟proc sort data=Indiana_AAVE; by race;
80⊟proc means data=Indiana_AAVE;
81  var sum_aave_flags;
82  by race;
83  run;
84
85⊟proc freq data=Indiana_AAVE;
86  tables Indiana_holistic_score Indiana_grammar_score;
87  run;
88
89⊟proc univariate data=Indiana_AAVE;
90  var sum_aave_flags;
91  histogram/;
92  inset mean std n skewness kurtosis;
93  run;
94
95⊟proc univariate data=Indiana_AAVE;
96  var Indiana_holistic_score;
97  histogram/;
98  inset mean std n skewness kurtosis;
99  run;
100
101⊟proc freq data=Indiana_AAVE;
102  tables sum_aave_flags;
103  run;
104
```

```
105  *overall participant descriptive;
106
107 proc contents data=Indiana; run;
108
109 proc freq data=Indiana;
110  tables city economically_disadvantaged ell grade gender race_ethnicity;
111  run;
112
113 proc freq data=Indiana;
114  tables state;
115  run;
116
117 proc means data=Indiana_AAVE;
118  var Indiana_grammar_score Indiana_holistic_score sum_aave_flags;
119  run;
120
121  *Participant descriptives for my sample;
122
123 proc sort data=Indiana_AAVE; by race;
124 proc freq data=Indiana_AAVE;
125  tables new_gender ED;
126  by race;
127  run;
128
129 proc sort data=Indiana_AAVE; by race;
130 proc freq data= Indiana_AAVE;
131  tables sum_aave_flags;
132  by race;
133  run;
134
```

```
134
135 proc means data=Indiana_AAVE;
136  var Indiana_grammar_score Indiana_holistic_score sum_aave_flags;
137  run;
138
139 proc corr data=Indiana_AAVE;
140  var Indiana_grammar_score Indiana_holistic_score sum_aave_flags Race new_Gender ED;
141  run;
142
143 proc sort data=Indiana_AAVE; by grade; run;
144 Proc freq data= Indiana_AAVE; tables contains_aave; by grade; run;
145
146  *Reading out file for mplus;
147
148 proc print data= indiana_AAVE;
149  run;
150
151 data Indiana_mplus; set Indiana_AAVE;
152  keep Indiana_grammar_score sum_aave_flags race ED new_gender;
153  run;
154
155 proc print data= Indiana_mplus; run;
156
157 data indiana_mplus_out2; set indiana_mplus;
158  file "C:\Users\nesbi\OneDrive\Desktop\Indiana_mplus2.dat"
159  dlm= '09'x;
160  put Indiana_grammar_score sum_aave_flags race ED new_gender;
161  run;
162
```

Appendix C: Partial Mediation Mplus Syntax and Output

```
INPUT INSTRUCTIONS

  TITLE:Full model;
  DATA: FILE IS "C:\Users\nesbi\OneDrive\Desktop\Indiana_mplus.dat";
        nobservations = 21420;
    VARIABLE: NAMES ARE writing AAVE race ED gender;
    analysis: estimator = ml;
    MODEL:writing on race gender ED AAVE;
      AAVE on race gender ED;
      !race with gender;
      !gender with ED;
      !race with ED;
    MODEL INDIRECT:
    writing IND AAVE race;
  writing IND AAVE gender;
  writing IND AAVE ED;


    OUTPUT: sampstat standardized residual cinterval;



INPUT READING TERMINATED NORMALLY



Full model;

SUMMARY OF ANALYSIS

Number of groups                                                1
Number of observations                                      21420

Number of dependent variables                                   2
Number of independent variables                                 3
Number of continuous latent variables                           0

Observed dependent variables

  Continuous
   WRITING      AAVE

Observed independent variables
   RACE         ED          GENDER


Estimator                                                      ML
Information matrix                                        OBSERVED
Maximum number of iterations                                 1000
Convergence criterion                                   0.500D-04
Maximum number of steepest descent iterations                  20
```

```
Input data file(s)
  C:\Users\nesbi\OneDrive\Desktop\Indiana_mplus.dat

Input data format   FREE


SAMPLE STATISTICS


      SAMPLE STATISTICS



         Means
             WRITING         AAVE          RACE           ED
GENDER

         _____       _____       _____       _____
_____
      1        3.321          0.219          0.058          0.363
0.513



         Covariances
             WRITING         AAVE          RACE           ED
GENDER

         _____       _____       _____       _____
_____
 WRITING       1.064
 AAVE          0.014          0.446
 RACE         -0.057          0.003          0.055
 ED           -0.124          0.007          0.022          0.231
 GENDER       -0.088          0.008          0.002          0.003
0.250



         Correlations
             WRITING         AAVE          RACE           ED
GENDER

         _____       _____       _____       _____
_____
 WRITING       1.000
 AAVE          0.020          1.000
 RACE         -0.236          0.021          1.000
 ED           -0.250          0.022          0.192          1.000
 GENDER       -0.171          0.023          0.018          0.012
1.000


THE MODEL ESTIMATION TERMINATED NORMALLY



MODEL FIT INFORMATION
```

```
Number of Free Parameters                       11

Loglikelihood

        H0 Value                        -51336.079
        H1 Value                        -51336.079

Information Criteria

        Akaike (AIC)                    102694.159
        Bayesian (BIC)                  102781.852
        Sample-Size Adjusted BIC        102746.894
          (n* = (n + 2) / 24)

Chi-Square Test of Model Fit

        Value                               0.001
        Degrees of Freedom                      0
        P-Value                            0.0000

RMSEA (Root Mean Square Error Of Approximation)

        Estimate                            0.000
        90 Percent C.I.                     0.000  0.000
        Probability RMSEA <= .05            0.000

CFI/TLI

        CFI                                 1.000
        TLI                                 1.000

Chi-Square Test of Model Fit for the Baseline Model

        Value                            2938.965
        Degrees of Freedom                      7
        P-Value                            0.0000

SRMR (Standardized Root Mean Square Residual)

        Value                               0.000




MODEL RESULTS

                                              Two-Tailed
                  Estimate      S.E.  Est./S.E.    P-Value

 WRITING  ON
    RACE            -0.850     0.029   -29.609      0.000
    GENDER          -0.341     0.013   -25.898      0.000
    ED              -0.453     0.014   -32.471      0.000
    AAVE             0.050     0.010     5.040      0.000
```

```
AAVE      ON
   RACE                 0.048       0.020       2.432       0.015
   GENDER               0.031       0.009       3.344       0.001
   ED                   0.026       0.010       2.642       0.008


Intercepts
   WRITING              3.699       0.011     342.193       0.000
   AAVE                 0.191       0.007      25.952       0.000


Residual Variances
   WRITING              0.929       0.009     103.489       0.000
   AAVE                 0.445       0.004     103.489       0.000
```

STANDARDIZED MODEL RESULTS

STDYX Standardization

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
WRITING ON				
RACE	-0.193	0.006	-30.106	0.000
GENDER	-0.165	0.006	-26.218	0.000
ED	-0.211	0.006	-33.137	0.000
AAVE	0.032	0.006	5.042	0.000
AAVE ON				
RACE	0.017	0.007	2.432	0.015
GENDER	0.023	0.007	3.345	0.001
ED	0.018	0.007	2.642	0.008
Intercepts				
WRITING	3.585	0.017	206.068	0.000
AAVE	0.286	0.011	25.635	0.000
Residual Variances				
WRITING	0.873	0.004	205.316	0.000
AAVE	0.999	0.000	2037.408	0.000

STDY Standardization

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
WRITING ON				
RACE	-0.824	0.027	-30.393	0.000
GENDER	-0.331	0.013	-26.419	0.000
ED	-0.439	0.013	-33.515	0.000
AAVE	0.032	0.006	5.042	0.000
AAVE ON				

```
    RACE                  0.072     0.030     2.433     0.015
    GENDER                0.046     0.014     3.346     0.001
    ED                    0.038     0.014     2.643     0.008

 Intercepts
    WRITING               3.585     0.017   206.068     0.000
    AAVE                  0.286     0.011    25.635     0.000

 Residual Variances
    WRITING               0.873     0.004   205.316     0.000
    AAVE                  0.999     0.000  2037.408     0.000
```

STD Standardization

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
WRITING ON				
RACE	-0.850	0.029	-29.609	0.000
GENDER	-0.341	0.013	-25.898	0.000
ED	-0.453	0.014	-32.471	0.000
AAVE	0.050	0.010	5.040	0.000
AAVE ON				
RACE	0.048	0.020	2.432	0.015
GENDER	0.031	0.009	3.344	0.001
ED	0.026	0.010	2.642	0.008
Intercepts				
WRITING	3.699	0.011	342.193	0.000
AAVE	0.191	0.007	25.952	0.000
Residual Variances				
WRITING	0.929	0.009	103.489	0.000
AAVE	0.445	0.004	103.489	0.000

R-SQUARE

Observed Variable	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
WRITING	0.127	0.004	29.875	0.000
AAVE	0.001	0.000	2.632	0.008

QUALITY OF NUMERICAL RESULTS

```
    Condition Number for the Information Matrix
0.133E-01
       (ratio of smallest to largest eigenvalue)
```

```
TOTAL, TOTAL INDIRECT, SPECIFIC INDIRECT, AND DIRECT EFFECTS


                                                     Two-Tailed
                       Estimate      S.E.   Est./S.E.   P-Value

Effects from RACE to WRITING

  Sum of indirect       0.002       0.001     2.190      0.028

  Specific indirect

    WRITING
    AAVE
    RACE               0.002       0.001     2.190      0.028


Effects from GENDER to WRITING

  Sum of indirect       0.002       0.001     2.787      0.005

  Specific indirect

    WRITING
    AAVE
    GENDER             0.002       0.001     2.787      0.005


Effects from ED to WRITING

  Sum of indirect       0.001       0.001     2.340      0.019

  Specific indirect

    WRITING
    AAVE
    ED                 0.001       0.001     2.340      0.019


STANDARDIZED TOTAL, TOTAL INDIRECT, SPECIFIC INDIRECT, AND DIRECT
EFFECTS


STDYX Standardization

                                                     Two-Tailed
                       Estimate      S.E.   Est./S.E.   P-Value

Effects from RACE to WRITING

  Sum of indirect       0.001       0.000     2.190      0.028

  Specific indirect
```

```
WRITING
AAVE
RACE                     0.001      0.000      2.190        0.028


Effects from GENDER to WRITING

  Sum of indirect        0.001      0.000      2.787        0.005

  Specific indirect

    WRITING
    AAVE
    GENDER               0.001      0.000      2.787        0.005


Effects from ED to WRITING

  Sum of indirect        0.001      0.000      2.340        0.019

  Specific indirect

    WRITING
    AAVE
    ED                   0.001      0.000      2.340        0.019


STDY Standardization

                                                   Two-Tailed
                      Estimate    S.E.   Est./S.E.   P-Value

Effects from RACE to WRITING

  Sum of indirect        0.002      0.001      2.191        0.028

  Specific indirect

    WRITING
    AAVE
    RACE                 0.002      0.001      2.191        0.028


Effects from GENDER to WRITING

  Sum of indirect        0.001      0.001      2.787        0.005

  Specific indirect

    WRITING
    AAVE
    GENDER               0.001      0.001      2.787        0.005
```

```
Effects from ED to WRITING

  Sum of indirect       0.001      0.001      2.340       0.019

  Specific indirect

    WRITING
    AAVE
    ED                  0.001      0.001      2.340       0.019


STD Standardization

                                                    Two-Tailed
                       Estimate      S.E.   Est./S.E.   P-Value
Effects from RACE to WRITING

  Sum of indirect       0.002      0.001      2.190       0.028

  Specific indirect

    WRITING
    AAVE
    RACE                0.002      0.001      2.190       0.028


Effects from GENDER to WRITING

  Sum of indirect       0.002      0.001      2.787       0.005

  Specific indirect

    WRITING
    AAVE
    GENDER              0.002      0.001      2.787       0.005


Effects from ED to WRITING

  Sum of indirect       0.001      0.001      2.340       0.019

  Specific indirect

    WRITING
    AAVE
    ED                  0.001      0.001      2.340       0.019



CONFIDENCE INTERVALS OF MODEL RESULTS

                  Lower .5%  Lower 2.5%    Lower 5%    Estimate
Upper 5%  Upper 2.5%   Upper .5%
```

```
WRITING   ON
    RACE                -0.924      -0.906      -0.897      -0.850
-0.803      -0.794      -0.776
    GENDER              -0.375      -0.367      -0.363      -0.341
-0.320      -0.316      -0.307
    ED                  -0.489      -0.481      -0.476      -0.453
-0.430      -0.426      -0.417
    AAVE                 0.024       0.030       0.034       0.050
0.066       0.069       0.075


 AAVE     ON
    RACE                -0.003       0.009       0.016       0.048
0.081       0.087       0.100
    GENDER               0.007       0.013       0.016       0.031
0.046       0.048       0.054
    ED                   0.001       0.007       0.010       0.026
0.041       0.044       0.050


 Intercepts
    WRITING              3.671       3.678       3.681       3.699
3.717       3.720       3.727
    AAVE                 0.172       0.177       0.179       0.191
0.203       0.206       0.210

 Residual Variances
    WRITING              0.906       0.912       0.914       0.929
0.944       0.947       0.952
    AAVE                 0.434       0.437       0.438       0.445
0.452       0.454       0.456


CONFIDENCE INTERVALS OF STANDARDIZED MODEL RESULTS


STDYX Standardization

                  Lower .5%  Lower 2.5%   Lower 5%    Estimate
Upper 5%   Upper 2.5%   Upper .5%

 WRITING   ON
    RACE                -0.209      -0.205      -0.203      -0.193
-0.182      -0.180      -0.176
    GENDER              -0.182      -0.178      -0.176      -0.165
-0.155      -0.153      -0.149
    ED                  -0.228      -0.224      -0.222      -0.211
-0.201      -0.199      -0.195
    AAVE                 0.016       0.020       0.022       0.032
0.043       0.045       0.049


 AAVE     ON
    RACE                -0.001       0.003       0.005       0.017
0.028       0.031       0.035
```

```
    GENDER              0.005      0.009      0.012      0.023
0.034        0.036        0.040
    ED                  0.000      0.005      0.007      0.018
0.030        0.032        0.036


 Intercepts
    WRITING             3.541      3.551      3.557      3.585
3.614        3.619        3.630
    AAVE                0.258      0.264      0.268      0.286
0.305        0.308        0.315


 Residual Variances
    WRITING             0.862      0.865      0.866      0.873
0.880        0.881        0.884
    AAVE                0.997      0.998      0.998      0.999
1.000        1.000        1.000


STDY Standardization

                  Lower .5%  Lower 2.5%   Lower 5%    Estimate
Upper 5%   Upper 2.5%   Upper .5%

 WRITING  ON
    RACE               -0.894     -0.877     -0.869     -0.824
-0.780       -0.771       -0.754
    GENDER             -0.363     -0.355     -0.352     -0.331
-0.310       -0.306       -0.299
    ED                 -0.473     -0.465     -0.461     -0.439
-0.418       -0.414       -0.406
    AAVE                0.016      0.020      0.022      0.032
0.043        0.045        0.049


 AAVE     ON
    RACE               -0.004      0.014      0.023      0.072
0.121        0.131        0.149
    GENDER              0.011      0.019      0.023      0.046
0.068        0.072        0.081
    ED                  0.001      0.010      0.014      0.038
0.062        0.067        0.075


 Intercepts
    WRITING             3.541      3.551      3.557      3.585
3.614        3.619        3.630
    AAVE                0.258      0.264      0.268      0.286
0.305        0.308        0.315


 Residual Variances
    WRITING             0.862      0.865      0.866      0.873
0.880        0.881        0.884
    AAVE                0.997      0.998      0.998      0.999
1.000        1.000        1.000
```

STD Standardization

	Lower .5%	Lower 2.5%	Lower 5%	Estimate	Upper 5%	Upper 2.5%	Upper .5%
WRITING ON							
RACE	-0.924	-0.906	-0.897	-0.850	-0.803	-0.794	-0.776
GENDER	-0.375	-0.367	-0.363	-0.341	-0.320	-0.316	-0.307
ED	-0.489	-0.481	-0.476	-0.453	-0.430	-0.426	-0.417
AAVE	0.024	0.030	0.034	0.050	0.066	0.069	0.075
AAVE ON							
RACE	-0.003	0.009	0.016	0.048	0.081	0.087	0.100
GENDER	0.007	0.013	0.016	0.031	0.046	0.048	0.054
ED	0.001	0.007	0.010	0.026	0.041	0.044	0.050
Intercepts							
WRITING	3.671	3.678	3.681	3.699	3.717	3.720	3.727
AAVE	0.172	0.177	0.179	0.191	0.203	0.206	0.210
Residual Variances							
WRITING	0.906	0.912	0.914	0.929	0.944	0.947	0.952
AAVE	0.434	0.437	0.438	0.445	0.452	0.454	0.456

CONFIDENCE INTERVALS OF TOTAL, TOTAL INDIRECT, SPECIFIC INDIRECT, AND DIRECT EFFECTS

	Lower .5%	Lower 2.5%	Lower 5%	Estimate	Upper 5%	Upper 2.5%	Upper .5%
Effects from RACE to WRITING							
Sum of indirect	0.000	0.000	0.001	0.002	0.004	0.005	0.005
Specific indirect							
WRITING AAVE RACE	0.000	0.000	0.001	0.002	0.004	0.005	0.005

Effects from GENDER to WRITING

  Sum of indirect   0.000   0.000   0.001   0.002
0.002   0.003   0.003

  Specific indirect

   WRITING
   AAVE
   GENDER   0.000   0.000   0.001   0.002
0.002   0.003   0.003

Effects from ED to WRITING

  Sum of indirect   0.000   0.000   0.000   0.001
0.002   0.002   0.003

  Specific indirect

   WRITING
   AAVE
   ED   0.000   0.000   0.000   0.001
0.002   0.002   0.003

CONFIDENCE INTERVALS OF STANDARDIZED TOTAL, TOTAL INDIRECT, SPECIFIC INDIRECT,
AND DIRECT EFFECTS

STDYX Standardization

               Lower .5% Lower 2.5%   Lower 5%   Estimate
Upper 5% Upper 2.5%   Upper .5%

Effects from RACE to WRITING

  Sum of indirect   0.000   0.000   0.000   0.001
0.001   0.001   0.001

  Specific indirect

   WRITING
   AAVE
   RACE   0.000   0.000   0.000   0.001
0.001   0.001   0.001

Effects from GENDER to WRITING

```
   Sum of indirect      0.000        0.000        0.000        0.001
0.001          0.001        0.001
```

   Specific indirect

     WRITING
     AAVE
```
     GENDER            0.000        0.000        0.000        0.001
0.001          0.001        0.001
```

Effects from ED to WRITING

```
   Sum of indirect      0.000        0.000        0.000        0.001
0.001          0.001        0.001
```

   Specific indirect

     WRITING
     AAVE
```
     ED                0.000        0.000        0.000        0.001
0.001          0.001        0.001
```

STDY Standardization

```
                    Lower .5%  Lower 2.5%    Lower 5%     Estimate
Upper 5%  Upper 2.5%   Upper .5%
```

Effects from RACE to WRITING

```
   Sum of indirect      0.000        0.000        0.001        0.002
0.004          0.004        0.005
```

   Specific indirect

     WRITING
     AAVE
```
     RACE              0.000        0.000        0.001        0.002
0.004          0.004        0.005
```

Effects from GENDER to WRITING

```
   Sum of indirect      0.000        0.000        0.001        0.001
0.002          0.003        0.003
```

   Specific indirect

     WRITING
     AAVE
```
     GENDER            0.000        0.000        0.001        0.001
0.002          0.003        0.003
```

Effects from ED to WRITING

  Sum of indirect   0.000   0.000   0.000   0.001
0.002   0.002   0.003

   Specific indirect

     WRITING
     AAVE
     ED   0.000   0.000   0.000   0.001
0.002   0.002   0.003

STD Standardization

                  Lower .5% Lower 2.5%   Lower 5%   Estimate
Upper 5% Upper 2.5%   Upper .5%

Effects from RACE to WRITING

  Sum of indirect   0.000   0.000   0.001   0.002
0.004   0.005   0.005

   Specific indirect

     WRITING
     AAVE
     RACE   0.000   0.000   0.001   0.002
0.004   0.005   0.005

Effects from GENDER to WRITING

  Sum of indirect   0.000   0.000   0.001   0.002
0.002   0.003   0.003

   Specific indirect

     WRITING
     AAVE
     GENDER   0.000   0.000   0.001   0.002
0.002   0.003   0.003

Effects from ED to WRITING

  Sum of indirect   0.000   0.000   0.000   0.001
0.002   0.002   0.003

   Specific indirect

     WRITING
     AAVE

```
     ED                  0.000        0.000        0.000        0.001
0.002        0.002        0.003
```

RESIDUAL OUTPUT

ESTIMATED MODEL AND RESIDUALS (OBSERVED - ESTIMATED)

Model Estimated Means/Intercepts/Thresholds

	WRITING	AAVE	RACE	ED
GENDER				
_____	_____	_____	_____	_____
1	3.321	0.219	0.058	0.363
0.513				

Residuals for Means/Intercepts/Thresholds

	WRITING	AAVE	RACE	ED
GENDER				
_____	_____	_____	_____	_____
1	0.000	0.000	0.000	0.000
0.000				

Standardized Residuals (z-scores) for Means/Intercepts/Thresholds

	WRITING	AAVE	RACE	ED
GENDER				
_____	_____	_____	_____	_____
1	0.000	5.800	0.000	0.000
0.000				

Normalized Residuals for Means/Intercepts/Thresholds

	WRITING	AAVE	RACE	ED
GENDER				
_____	_____	_____	_____	_____
1	0.000	0.004	0.000	0.000
0.000				

Model Estimated Covariances/Correlations/Residual Correlations

	WRITING	AAVE	RACE	ED
GENDER				
_____	_____	_____	_____	_____

```
 WRITING          1.064
 AAVE             0.014          0.446
 RACE            -0.057          0.003          0.055
 ED              -0.124          0.007          0.022          0.231
 GENDER          -0.088          0.008          0.002          0.003
0.250
```

```
           Residuals for Covariances/Correlations/Residual
Correlations
                WRITING         AAVE           RACE           ED
GENDER
                _____       _____       _____       _____
_____
 WRITING          0.000
 AAVE             0.000          0.000
 RACE             0.000          0.000          0.000
 ED               0.000          0.000          0.000          0.000
 GENDER           0.000          0.000          0.000          0.000
0.000
```

```
           Standardized Residuals (z-scores) for
Covariances/Correlations/Residual Corr
                WRITING         AAVE           RACE           ED
GENDER
                _____       _____       _____       _____
_____
 WRITING          0.442
 AAVE             0.000          0.000
 RACE            -2.218          0.000          0.000
 ED               0.000          0.000          0.000          0.000
 GENDER           0.000          0.000          0.000          0.000
0.000
```

```
           Normalized Residuals for
Covariances/Correlations/Residual Correlations
                WRITING         AAVE           RACE           ED
GENDER
                _____       _____       _____       _____
_____
 WRITING          0.007
 AAVE             0.000          0.000
 RACE            -0.027          0.000          0.000
 ED               0.000          0.000          0.000          0.000
 GENDER           0.000          0.000          0.000          0.000
0.000
```

```
DIAGRAM INFORMATION
```

Appendix D: Full Mediation Mplus Syntax and Output

```
   TITLE:Full model;
   DATA: FILE IS "C:\Users\nesbi\OneDrive\Desktop\Indiana_mplus.dat";
          nobservations = 21420;
     VARIABLE: NAMES ARE writing AAVE race ED gender;
     analysis: estimator = ml;
     MODEL:writing on gender ED AAVE;
       AAVE on race gender ED;
       !race with gender;
       !gender with ED;
       !race with ED;
     MODEL INDIRECT:
     writing IND AAVE race;
   writing IND AAVE gender;
   writing IND AAVE ED;


     OUTPUT: sampstat standardized residual cinterval;



INPUT READING TERMINATED NORMALLY



Full model;

SUMMARY OF ANALYSIS

Number of groups                                                  1
Number of observations                                       21420

Number of dependent variables                                    2
Number of independent variables                                  3
Number of continuous latent variables                            0

Observed dependent variables

  Continuous
   WRITING     AAVE

Observed independent variables
   RACE          ED         GENDER


Estimator                                                       ML
Information matrix                                         OBSERVED
Maximum number of iterations                                  1000
Convergence criterion                                   0.500D-04
Maximum number of steepest descent iterations                   20
```

```
Input data file(s)
  C:\Users\nesbi\OneDrive\Desktop\Indiana_mplus.dat

Input data format  FREE


SAMPLE STATISTICS


     SAMPLE STATISTICS


        Means
            WRITING        AAVE         RACE          ED
GENDER


_____
       1       3.321        0.219        0.058        0.363
0.513



        Covariances
            WRITING        AAVE         RACE          ED
GENDER


_____
 WRITING       1.064
 AAVE          0.014        0.446
 RACE         -0.057        0.003        0.055
 ED           -0.124        0.007        0.022        0.231
 GENDER       -0.088        0.008        0.002        0.003
0.250



        Correlations
            WRITING        AAVE         RACE          ED
GENDER


_____
 WRITING       1.000
 AAVE          0.020        1.000
 RACE         -0.236        0.021        1.000
 ED           -0.250        0.022        0.192        1.000
 GENDER       -0.171        0.023        0.018        0.012
1.000



THE MODEL ESTIMATION TERMINATED NORMALLY



MODEL FIT INFORMATION

Number of Free Parameters                        10
```

```
Loglikelihood

        H0 Value                          -51766.571
        H1 Value                          -51336.079

Information Criteria

        Akaike (AIC)                      103553.143
        Bayesian (BIC)                    103632.864
        Sample-Size Adjusted BIC          103601.084
          (n* = (n + 2) / 24)

Chi-Square Test of Model Fit

        Value                               860.985
        Degrees of Freedom                        1
        P-Value                              0.0000

RMSEA (Root Mean Square Error Of Approximation)

        Estimate                              0.200
        90 Percent C.I.                       0.189  0.212
        Probability RMSEA <= .05              0.000

CFI/TLI

        CFI                                   0.707
        TLI                                  -1.053

Chi-Square Test of Model Fit for the Baseline Model

        Value                              2938.965
        Degrees of Freedom                        7
        P-Value                              0.0000

SRMR (Standardized Root Mean Square Residual)

        Value                                 0.042



MODEL RESULTS

                                                        Two-Tailed
                        Estimate      S.E.  Est./S.E.    P-Value

  WRITING   ON
     GENDER             -0.348       0.013   -25.851      0.000
     ED                 -0.533       0.014   -38.102      0.000
     AAVE                0.045       0.010     4.460      0.000

  AAVE     ON
     RACE                0.048       0.020     2.431      0.015
```

```
     GENDER            0.031      0.009      3.345      0.001
     ED                0.026      0.010      2.643      0.008

 Intercepts
     WRITING           3.683      0.011    334.342      0.000
     AAVE              0.191      0.007     25.953      0.000

 Residual Variances
     WRITING           0.967      0.009    103.489      0.000
     AAVE              0.445      0.004    103.490      0.000


STANDARDIZED MODEL RESULTS


STDYX Standardization

                                                     Two-Tailed
                    Estimate       S.E.   Est./S.E.    P-Value

 WRITING   ON
     GENDER           -0.168      0.006    -26.199      0.000
     ED               -0.248      0.006    -39.292      0.000
     AAVE              0.029      0.007      4.461      0.000

 AAVE      ON
     RACE              0.017      0.007      2.431      0.015
     GENDER            0.023      0.007      3.346      0.001
     ED                0.018      0.007      2.643      0.008

 Intercepts
     WRITING           3.569      0.017    204.323      0.000
     AAVE              0.286      0.011     25.637      0.000

 Residual Variances
     WRITING           0.909      0.004    242.205      0.000
     AAVE              0.999      0.000   2037.230      0.000


STDY Standardization

                                                     Two-Tailed
                    Estimate       S.E.   Est./S.E.    P-Value

 WRITING   ON
     GENDER           -0.337      0.013    -26.399      0.000
     ED               -0.516      0.013    -39.926      0.000
     AAVE              0.029      0.007      4.461      0.000

 AAVE      ON
     RACE              0.072      0.030      2.432      0.015
     GENDER            0.046      0.014      3.346      0.001
     ED                0.038      0.014      2.644      0.008
```

```
Intercepts
   WRITING                   3.569      0.017     204.323       0.000
   AAVE                      0.286      0.011      25.637       0.000

Residual Variances
   WRITING                   0.909      0.004     242.205       0.000
   AAVE                      0.999      0.000    2037.230       0.000
```

STD Standardization

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
WRITING ON				
GENDER	-0.348	0.013	-25.851	0.000
ED	-0.533	0.014	-38.102	0.000
AAVE	0.045	0.010	4.460	0.000
AAVE ON				
RACE	0.048	0.020	2.431	0.015
GENDER	0.031	0.009	3.345	0.001
ED	0.026	0.010	2.643	0.008
Intercepts				
WRITING	3.683	0.011	334.342	0.000
AAVE	0.191	0.007	25.953	0.000
Residual Variances				
WRITING	0.967	0.009	103.489	0.000
AAVE	0.445	0.004	103.490	0.000

R-SQUARE

Observed Variable	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
WRITING	0.091	0.004	24.330	0.000
AAVE	0.001	0.000	2.632	0.008

QUALITY OF NUMERICAL RESULTS

```
     Condition Number for the Information Matrix
0.292E-01
        (ratio of smallest to largest eigenvalue)
```

TOTAL, TOTAL INDIRECT, SPECIFIC INDIRECT, AND DIRECT EFFECTS

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value

```
Effects from RACE to WRITING

  Sum of indirect        0.002      0.001      2.135      0.033

  Specific indirect

    WRITING
    AAVE
    RACE                 0.002      0.001      2.135      0.033


Effects from GENDER to WRITING

  Sum of indirect        0.001      0.001      2.676      0.007

  Specific indirect

    WRITING
    AAVE
    GENDER               0.001      0.001      2.676      0.007


Effects from ED to WRITING

  Sum of indirect        0.001      0.001      2.274      0.023

  Specific indirect

    WRITING
    AAVE
    ED                   0.001      0.001      2.274      0.023


STANDARDIZED TOTAL, TOTAL INDIRECT, SPECIFIC INDIRECT, AND DIRECT
EFFECTS


STDYX Standardization

                                                  Two-Tailed
                     Estimate      S.E.  Est./S.E.  P-Value

Effects from RACE to WRITING

  Sum of indirect        0.000      0.000      2.135      0.033

  Specific indirect

    WRITING
    AAVE
    RACE                 0.000      0.000      2.135      0.033
```

```
Effects from GENDER to WRITING

   Sum of indirect      0.001      0.000      2.676      0.007

   Specific indirect

     WRITING
     AAVE
     GENDER            0.001      0.000      2.676      0.007


Effects from ED to WRITING

   Sum of indirect      0.001      0.000      2.274      0.023

   Specific indirect

     WRITING
     AAVE
     ED                0.001      0.000      2.274      0.023


STDY Standardization
```

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value

```
Effects from RACE to WRITING

   Sum of indirect      0.002      0.001      2.135      0.033

   Specific indirect

     WRITING
     AAVE
     RACE              0.002      0.001      2.135      0.033


Effects from GENDER to WRITING

   Sum of indirect      0.001      0.000      2.676      0.007

   Specific indirect

     WRITING
     AAVE
     GENDER            0.001      0.000      2.676      0.007


Effects from ED to WRITING

   Sum of indirect      0.001      0.000      2.274      0.023

   Specific indirect
```

```
  WRITING
  AAVE
  ED                      0.001       0.000       2.274       0.023
```

STD Standardization

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value

Effects from RACE to WRITING

```
  Sum of indirect        0.002       0.001       2.135       0.033

  Specific indirect

    WRITING
    AAVE
    RACE                 0.002       0.001       2.135       0.033
```

Effects from GENDER to WRITING

```
  Sum of indirect        0.001       0.001       2.676       0.007

  Specific indirect

    WRITING
    AAVE
    GENDER               0.001       0.001       2.676       0.007
```

Effects from ED to WRITING

```
  Sum of indirect        0.001       0.001       2.274       0.023

  Specific indirect

    WRITING
    AAVE
    ED                   0.001       0.001       2.274       0.023
```

CONFIDENCE INTERVALS OF MODEL RESULTS

	Lower .5%	Lower 2.5%	Lower 5%	Estimate	Upper 5%	Upper 2.5%	Upper .5%
WRITING ON							
GENDER	-0.382	-0.374	-0.370	-0.348	-0.326	-0.321	-0.313

	Lower .5%	Lower 2.5%	Lower 5%	Estimate	Upper 5%	Upper 2.5%	Upper .5%
ED	-0.569	-0.560	-0.556	-0.533	-0.510	-0.505	-0.497
AAVE	0.019	0.025	0.028	0.045	0.061	0.065	0.071
AAVE ON							
RACE	-0.003	0.009	0.016	0.048	0.081	0.087	0.100
GENDER	0.007	0.013	0.016	0.031	0.046	0.048	0.054
ED	0.001	0.007	0.010	0.026	0.041	0.044	0.050
Intercepts							
WRITING	3.654	3.661	3.664	3.683	3.701	3.704	3.711
AAVE	0.172	0.177	0.179	0.191	0.203	0.206	0.210
Residual Variances							
WRITING	0.943	0.949	0.952	0.967	0.983	0.986	0.991
AAVE	0.434	0.437	0.438	0.445	0.452	0.454	0.456

CONFIDENCE INTERVALS OF STANDARDIZED MODEL RESULTS

STDYX Standardization

	Lower .5%	Lower 2.5%	Lower 5%	Estimate	Upper 5%	Upper 2.5%	Upper .5%
WRITING ON							
GENDER	-0.185	-0.181	-0.179	-0.168	-0.158	-0.156	-0.152
ED	-0.265	-0.261	-0.259	-0.248	-0.238	-0.236	-0.232
AAVE	0.012	0.016	0.018	0.029	0.040	0.042	0.046
AAVE ON							
RACE	-0.001	0.003	0.005	0.017	0.028	0.031	0.035
GENDER	0.005	0.009	0.012	0.023	0.034	0.036	0.040
ED	0.000	0.005	0.007	0.018	0.030	0.032	0.036
Intercepts							
WRITING	3.524	3.535	3.541	3.569	3.598	3.604	3.614

```
     AAVE                 0.258        0.264        0.268        0.286
0.305        0.308        0.315

  Residual Variances
     WRITING              0.899        0.901        0.903        0.909
0.915        0.916        0.918
     AAVE                 0.997        0.998        0.998        0.999
1.000        1.000        1.000
```

STDY Standardization

```
                    Lower .5%  Lower 2.5%   Lower 5%     Estimate
Upper 5%   Upper 2.5%   Upper .5%

  WRITING   ON
     GENDER              -0.370       -0.362       -0.358       -0.337
-0.316       -0.312       -0.304
     ED                  -0.550       -0.542       -0.537       -0.516
-0.495       -0.491       -0.483
     AAVE                 0.012        0.016        0.018        0.029
0.040        0.042        0.046

  AAVE      ON
     RACE                -0.004        0.014        0.023        0.072
0.121        0.131        0.149
     GENDER               0.011        0.019        0.023        0.046
0.068        0.072        0.081
     ED                   0.001        0.010        0.014        0.038
0.062        0.067        0.076

  Intercepts
     WRITING              3.524        3.535        3.541        3.569
3.598        3.604        3.614
     AAVE                 0.258        0.264        0.268        0.286
0.305        0.308        0.315

  Residual Variances
     WRITING              0.899        0.901        0.903        0.909
0.915        0.916        0.918
     AAVE                 0.997        0.998        0.998        0.999
1.000        1.000        1.000
```

STD Standardization

```
                    Lower .5%  Lower 2.5%   Lower 5%     Estimate
Upper 5%   Upper 2.5%   Upper .5%

  WRITING   ON
     GENDER              -0.382       -0.374       -0.370       -0.348
-0.326       -0.321       -0.313
     ED                  -0.569       -0.560       -0.556       -0.533
-0.510       -0.505       -0.497
```

	Lower .5%	Lower 2.5%	Lower 5%	Estimate	Upper 5%	Upper 2.5%	Upper .5%
AAVE	0.019	0.025	0.028	0.045	0.061	0.065	0.071

AAVE ON

	Lower .5%	Lower 2.5%	Lower 5%	Estimate	Upper 5%	Upper 2.5%	Upper .5%
RACE	-0.003	0.009	0.016	0.048	0.081	0.087	0.100
GENDER	0.007	0.013	0.016	0.031	0.046	0.048	0.054
ED	0.001	0.007	0.010	0.026	0.041	0.044	0.050

Intercepts

	Lower .5%	Lower 2.5%	Lower 5%	Estimate	Upper 5%	Upper 2.5%	Upper .5%
WRITING	3.654	3.661	3.664	3.683	3.701	3.704	3.711
AAVE	0.172	0.177	0.179	0.191	0.203	0.206	0.210

Residual Variances

	Lower .5%	Lower 2.5%	Lower 5%	Estimate	Upper 5%	Upper 2.5%	Upper .5%
WRITING	0.943	0.949	0.952	0.967	0.983	0.986	0.991
AAVE	0.434	0.437	0.438	0.445	0.452	0.454	0.456

CONFIDENCE INTERVALS OF TOTAL, TOTAL INDIRECT, SPECIFIC INDIRECT, AND DIRECT EFFECTS

	Lower .5%	Lower 2.5%	Lower 5%	Estimate	Upper 5%	Upper 2.5%	Upper .5%

Effects from RACE to WRITING

	Lower .5%	Lower 2.5%	Lower 5%	Estimate	Upper 5%	Upper 2.5%	Upper .5%
Sum of indirect	0.000	0.000	0.000	0.002	0.004	0.004	0.005

Specific indirect

WRITING
AAVE

	Lower .5%	Lower 2.5%	Lower 5%	Estimate	Upper 5%	Upper 2.5%	Upper .5%
RACE	0.000	0.000	0.000	0.002	0.004	0.004	0.005

Effects from GENDER to WRITING

	Lower .5%	Lower 2.5%	Lower 5%	Estimate	Upper 5%	Upper 2.5%	Upper .5%
Sum of indirect	0.000	0.000	0.001	0.001	0.002	0.002	0.003

Specific indirect

WRITING
AAVE

```
     GENDER            0.000        0.000        0.001        0.001
0.002       0.002        0.003


Effects from ED to WRITING

  Sum of indirect      0.000        0.000        0.000        0.001
0.002       0.002        0.002

   Specific indirect

     WRITING
     AAVE
     ED                0.000        0.000        0.000        0.001
0.002       0.002        0.002
```

CONFIDENCE INTERVALS OF STANDARDIZED TOTAL, TOTAL INDIRECT, SPECIFIC INDIRECT,
AND DIRECT EFFECTS

STDYX Standardization

```
                  Lower .5%  Lower 2.5%   Lower 5%     Estimate
Upper 5%  Upper 2.5%   Upper .5%
```

Effects from RACE to WRITING

```
  Sum of indirect      0.000        0.000        0.000        0.000
0.001       0.001        0.001

   Specific indirect

     WRITING
     AAVE
     RACE              0.000        0.000        0.000        0.000
0.001       0.001        0.001


Effects from GENDER to WRITING

  Sum of indirect      0.000        0.000        0.000        0.001
0.001       0.001        0.001

   Specific indirect

     WRITING
     AAVE
     GENDER            0.000        0.000        0.000        0.001
0.001       0.001        0.001


Effects from ED to WRITING
```

```
   Sum of indirect     0.000        0.000        0.000        0.001
0.001       0.001      0.001


   Specific indirect

     WRITING
     AAVE
     ED                0.000        0.000        0.000        0.001
0.001       0.001      0.001


STDY Standardization

                    Lower .5%  Lower 2.5%   Lower 5%    Estimate
Upper 5%  Upper 2.5%   Upper .5%

Effects from RACE to WRITING

   Sum of indirect     0.000        0.000        0.000        0.002
0.004       0.004      0.005


   Specific indirect

     WRITING
     AAVE
     RACE              0.000        0.000        0.000        0.002
0.004       0.004      0.005



Effects from GENDER to WRITING

   Sum of indirect     0.000        0.000        0.001        0.001
0.002       0.002      0.003


   Specific indirect

     WRITING
     AAVE
     GENDER            0.000        0.000        0.001        0.001
0.002       0.002      0.003



Effects from ED to WRITING

   Sum of indirect     0.000        0.000        0.000        0.001
0.002       0.002      0.002


   Specific indirect

     WRITING
     AAVE
     ED                0.000        0.000        0.000        0.001
0.002       0.002      0.002
```

```
STD Standardization

                    Lower .5%  Lower 2.5%    Lower 5%     Estimate
Upper 5%   Upper 2.5%   Upper .5%

Effects from RACE to WRITING

  Sum of indirect      0.000        0.000        0.000        0.002
0.004        0.004       0.005

  Specific indirect

    WRITING
    AAVE
    RACE               0.000        0.000        0.000        0.002
0.004        0.004       0.005


Effects from GENDER to WRITING

  Sum of indirect      0.000        0.000        0.001        0.001
0.002        0.002       0.003

  Specific indirect

    WRITING
    AAVE
    GENDER             0.000        0.000        0.001        0.001
0.002        0.002       0.003


Effects from ED to WRITING

  Sum of indirect      0.000        0.000        0.000        0.001
0.002        0.002       0.002

  Specific indirect

    WRITING
    AAVE
    ED                 0.000        0.000        0.000        0.001
0.002        0.002       0.002



RESIDUAL OUTPUT


      ESTIMATED MODEL AND RESIDUALS (OBSERVED - ESTIMATED)


          Model Estimated Means/Intercepts/Thresholds
```

```
                   WRITING       AAVE         RACE          ED
GENDER

                  _____     _____     _____     _____
_____
         1         3.321        0.219        0.058        0.363
0.513
```

```
         Residuals for Means/Intercepts/Thresholds
                   WRITING       AAVE         RACE          ED
GENDER

                  _____     _____     _____     _____
_____
         1         0.000        0.000        0.000        0.000
0.000
```

```
         Standardized Residuals (z-scores) for
Means/Intercepts/Thresholds
                   WRITING       AAVE         RACE          ED
GENDER

                  _____     _____     _____     _____
_____
         1         0.000        0.000        0.000        0.000
0.000
```

```
         Normalized Residuals for Means/Intercepts/Thresholds
                   WRITING       AAVE         RACE          ED
GENDER

                  _____     _____     _____     _____
_____
         1         0.000        0.000        0.000        0.000
0.000
```

```
         Model Estimated Covariances/Correlations/Residual
Correlations
                   WRITING       AAVE         RACE          ED
GENDER

                  _____     _____     _____     _____
_____
  WRITING          1.064
  AAVE             0.014        0.446
  RACE            -0.012        0.003        0.055
  ED              -0.124        0.007        0.022        0.231
  GENDER          -0.088        0.008        0.002        0.003
0.250
```

```
         Residuals for Covariances/Correlations/Residual
Correlations
                   WRITING       AAVE         RACE          ED
GENDER
```

```
                 _____    _____    _____    _____
_____
 WRITING         0.000
 AAVE            0.000       0.000
 RACE           -0.045       0.000       0.000
 ED              0.000       0.000       0.000       0.000
 GENDER          0.000       0.000       0.000       0.000
0.000
```

```
             Standardized Residuals (z-scores) for
Covariances/Correlations/Residual Corr
                 WRITING     AAVE        RACE        ED
GENDER

                 _____    _____    _____    _____
_____
 WRITING         0.000
 AAVE            0.000       0.000
 RACE          -28.206       0.000       0.000
 ED              0.000       0.000       0.000       0.000
 GENDER          0.000       0.000       0.000       0.000
0.000
```

```
             Normalized Residuals for
Covariances/Correlations/Residual Correlations
                 WRITING     AAVE        RACE        ED
GENDER

                 _____    _____    _____    _____
_____
 WRITING         0.000
 AAVE            0.000       0.000
 RACE          -26.442       0.000       0.000
 ED              0.000       0.000       0.000       0.000
 GENDER          0.000       0.000       0.000       0.000
0.000
```