2016

Effect of Vocabulary Test Preparation on Low-Income Black Middle School Students' Reading Scores

Ingrid Nadia Mitchell
Walden University

Walden University

COLLEGE OF EDUCATION

This is to certify that the doctoral study by

Ingrid Mitchell

has been found to be complete and satisfactory in all respects,
and that any and all revisions required by
the review committee have been made.

Review Committee
Dr. Nicolae Nistor, Committee Chairperson, Education Faculty
Dr. Beate Baltes, Committee Member, Education Faculty
Dr. Michelle Brown, University Reviewer, Education Faculty

Chief Academic Officer

Eric Riedel, Ph.D.

Walden University
2016

Abstract

Effect of Vocabulary Test Preparation on Low-Income Black Middle School Students' Reading Scores

by

Ingrid Nadia Mitchell

EdS, Union University, 2007

MA, University of Memphis, 2004

BS, Oral Roberts University, 1999

Doctoral Study Submitted in Partial Fulfillment

of the Requirements for the Degree of

Doctor of Education

Walden University

October 2016

Abstract

Black middle school students in the United States perform poorly on standardized reading achievement tests in comparison to other racial and ethnic groups. The purpose of this study was to examine the effectiveness of a vocabulary-focused test preparation program for underachieving Black middle school students. Thorndike's concept of test-wiseness, a test-taking capacity, provided the theoretical foundation for the study. Research questions investigated the difference in reading test scores on the Discovery Education Assessment of underachieving Black middle school students who participated in a key vocabulary test preparation program and those who did not. An intact-group comparison was used in which the research site, a large urban middle school in Tennessee, was matched with a similar middle school. Teachers at the research site were trained on a test preparation strategy meant to familiarize students with key vocabulary terms related to test items. Participants were Black students in Grades 6, 7, and 8, who were enrolled in Title 1 supplementary reading instruction and scored below proficiency, with 405 students in the treatment group and 249 students in the control group. The post-intervention reading test scores on the Discovery Education Assessment were compared between the groups, with a baseline test score used as the covariate. The adjusted mean scores for both 6th and 7th grade students were significantly greater for the intervention group ($p = .018$ and $p = .062$ respectively), whereas there was no significant difference in test scores for 8th grade students ($p = .246$). Implications for positive social change include providing research-based findings to the study site that support the vocabulary-focused test preparation program to improve the reading achievement of Black middle school students.

Effect of Vocabulary Test Preparation on Low-Income Black Middle School Students' Reading Scores

by

Ingrid Nadia Mitchell

EdS, Union University, 2007

MA, University of Memphis, 2004

BS, Oral Roberts University, 1999

Doctoral Study Submitted in Partial Fulfillment

of the Requirements for the Degree of

Doctor of Education

Walden University

October 2016

Dedication

This research study is dedicated to educators who strive to inspire learners to the aptitude track of educational accomplishment.

Acknowledgments

I acknowledge my parents Oliver W. and Angela M. Mitchell, who are teachers, for imparting the philosophy that learning is key to success in life.

Table of Contents

List of Tables

List of Figures

Section 1: Introduction to the Study

Introduction

At both state and national levels, Black students in the United States perform poorly on measures of reading achievement. According to the Nation's Report Card (2013), poor performance is steady with no improvements. This poor performance is disturbing. Despite racial and ethnic identities, students should show evidence of growth in assessment measures over time, particularly at the completion of each grade.

The local problem prompting this study was the poor performance of Black middle school students in Tennessee on standardized measures of reading achievement. In the 2011-2012 school year, Tennessee state results showed 20% of Grades 3-8 students scored below basic, 48% scored basic, 25% scored proficient, and 5% scored advanced on the Tennessee Comprehensive Achievement Program test (TCAP) (Tennessee State DOE, n.d.). To help improve the performance of Black students on standardized reading tests, teachers at one Tennessee middle school implemented a test preparation instructional program designed to familiarize students with formal tests that they are required to take annually. The test preparation program was a daily integration of 12 Powerful Words into core academic content (Bell, 2005) with Discovery Education Assessment (DEA) used as the measuring tool. This measuring tool is not a replacement for TCAP, but a formal assessment pre-determiner for students' performance rate of proficiency for the state TCAP test.

The purpose of this research study was to determine whether implementation of the instructional program impacts the reading achievement scores of Black students. In

this section, I present an overview of the problem at local, state, and national levels. The following sections feature discussions of the problem, nature of study, purpose, theoretical framework, definitions, assumptions, limitations, scope, and delimitations, significance, and summary.

Evidence of the Problem at the Local Level

Formative assessments are imperative in tracking student learning and skill-set progress for end of year cumulative testing. Therefore, Tennessee teachers in Grades 3-8 administer the DEA, three times a year, beginning (A), middle (B), end (C) (Discovery Education, 2008). The DEA is formatted as well as standardized to show student content knowledge and predict student performance outcomes for the TCAP test (Discovery Education, 2008). The TCAP does not publish its validity and reliability (Tennessee Comptroller DOE, n.d.). However, using the results of the formative DEA test, teachers reteach material that students have not yet mastered, to boost scores on the cumulative state exam (Discovery Education, 2008).

Public record from the Tennessee Department of Education website showed that in the participating district for the 2011-2012 school year, 27% of Grades 3-8 students scored below basic, 47% scored basic, 21% scored proficient, and 4% scored advanced on the TCAP. Table 1 shows how students' scale (or, cut) scores varied based on grade level across the state.

Table 1

Cut Scores for Middle School Students on the Tennessee Comprehensive Achievement Test, 2011-2012 School Year

Grade Level	Below Basic	Basic	Proficient	Advanced
6	0-33	34-51	52-68	69-75
7	0-36	37-54	55-66	67-75
8	0-31	32-54	55-66	67-75

Note: Any score above 75 is advanced placement.

Noted by school personnel on the publically viewed Tennessee School Improvement Plan, for the 2011-2012 school year, Grades 6 and 7 students at the test preparation site, School A, correctly answered a mean of 45% of test items in reading on the DEA, which is considered basic level achievement by the Tennessee Department of Education (DOE). For Grade 8 students, the mean score of correctly answered test items was 56%, which is also considered basic level achievement. Thus, students in all three grades scored at the second lowest of the four proficiency levels. Black males underperformed Black females. Of 241 Black males who completed the reading test, 205 (85%) scored below basic or basic, 30 (13%) scored proficient, and 6 (2%) scored advanced. In contrast, 62 (49%) of the Black females who completed the test scored below basic or basic, 35 (27%) scored proficient, and 30 (24%) scored advanced.

School A is being compared to School B with School B not offering the test preparation program. School B, did not provide a detailed account of 2011-2012 DEA scores on the Tennessee School Improvement Plan, which was public record. However,

School B was in *corrective action* status based on 2011-2012 TCAP scores, meaning that students in School B were low-performing (Tennessee State DOE, n.d).

Evidence of the Problem from the Professional Literature

A central focus of U.S. educational policy at both the state and national levels is ensuring that more Black students score at least proficient in reading (National Assessment of Educational Progress, [NAEP], n.d.). NAEP monitors trends in achievement tests, providing synthesis of state test results, and making comparisons of those results state to state (Buckendahl et al., 2009). Those results are used by districts to determine instructional strategies for their student bodies. Several researchers continue to publish evidence that Black students consistently do not read at grade level (see Condron, Tope, Steidl, & Freeman, 2013; Irizarry, 2015; Mosley, Wetzel, & Rogers, 2015; Reardon, Valentino, & Shores, 2012; West, 2012), and have determined that Black students are having academic difficulties in reading that lead to an overwhelming gap in reading achievement.

The number of intermediate Black students in Grades 4 and 8 who score below proficient in the U.S. has not declined (Nation's Report Card, 2013). According to the Nation's Report Card (2013), in 2011 Black students in 4th grade, who took their state exam, NAEP averaged a scale score of 205 out of a possible 500. White students averaged a scale score of 231. In 2013, Black students in 4th grade, who took their state exam, NAEP averaged a scale score of 206 out of a possible 500 (Nation's Report Card, 2013). White students averaged a scale score of 232. Taking a glimpse into upper intermediate or secondary level students, according to the Nation's Report Card, in 2011

Black students in 8th grade, who took their state exam, NAEP averaged a scale score of 249 out of a possible 500. White students averaged a scale score of 274. In 2013, Black students in 8th grade, who took their state exam, NAEP averaged a scale score of 250 out of a possible 500. White students averaged a scale score of 276. While in Tennessee, the low test performance of Black students is actually increasing (Tennessee State DOE, n.d.). This highlights that across the state of Tennessee, Black students, especially Black males are more likely than other racial or ethnic groups to be below proficient in reading (Johnson, 2014; Wood & Jocius, 2013).

Furthermore, only 11% of Black males attending large city schools are proficient in reading (Nation's Report Card, 2013). Similarly, according to Educational Testing Services only 8% of junior high Black males attending large city schools are proficient in reading as compared to 33% of Black females (Yaffe, 2012). Conferring to a 2012 report by the U.S. Department of Labor, young Black males, across the board, performed below other racial and ethnic groups in relation to graduation rates, literacy achievement, and college preparedness (see also, Enriquez, 2013; Milner, 2012; Piazza & Duncan, 2012). Unambiguously, because of the low achievement in reading, there are low percentage amounts of Black students who graduate from high school or enter college (Grier-Reed, Ehlert, & Dade, 2011; Keels, 2013; Lundy-Wagner & Gasman, 2011; Phillip, 2011; Wood & Palmer, 2013). Specifically, if Black males do not improve with reading growth achievement, graduation attainment, and become college or career ready, (Garo, 2013) reminds educators that there is a major correlation between reading achievement,

neighborhood poverty and violence. Nevertheless, with all stakeholders involved, success can be measured and accomplished (Purtell & McLoyd, 2013).

Problem Statement

According to the National Center for Educational Statistics (2009), from 1998-2007, low reading scores for Black students remained unchanged. Hence, a noteworthy strategy must be implemented to make a change in reading achievement. Therefore, there rests an obligation to give our students an opportunity to learn, live, and thrive as adults by being reading proficient.

As a result of the academic performance of Black students at the U.S. national level, the federal government has interceded to improve literacy among the disadvantaged populations. Since the 1960s, the federal government has funneled significant monies to lower-performing school districts in an effort to combat the effects of race and poverty. This funding began with the enactment of the Elementary and Secondary Education Act, (ESEA) in 1965 as part of the Johnson administration's War on Poverty (Kantor, 1991). Through the 1990s, federal policy resulted in little systemic change at the school level within reading in the U.S. (Kucan & Palinscar, 2011). However, with new laws, including the reauthorizations of ESEA as the Improving America's Schools Act of 1994, the No Child Left Behind Act of 2001, and Race to the Top of 2013, federal policymakers sought to stimulate reform through standards and accountability of core subjects and proficiency scores (Kucan & Palinscar, 2011; Lauen & Gaddis, 2012). The standards and the accountability annual audit through Federal Programs & Grants, FP&G are monitoring schools' assessment data, teacher professional development, community

engagement, English language acquisitions, and homeless education (No Child Left Behind [NCLB], 2002). In addition to the amended provisions plus standards, paramount among the initiatives under NCLB was Reading First, which focused on disadvantaged students in kindergarten through Grade 3.

Reading First provided students with a scientifically validated instruction in phonemic awareness, fluency, vocabulary, and reading comprehension strategies (Ashby, Burns, & Royle, 2014; Barone, 2013; Bean, Dole, Nelson, Belcastro, & Zigmond, 2015; NCLB, 2002). The program also mandated classroom screening, instructional and diagnostic assessments, and professional development for teachers (NCLB, 2002). In the participating district, many elementary schools were given Reading First funds and were named Reading First Schools by NCLB. Yet, Reading First did not show U.S. national growth in reading, and federally mandated investigations determined that targeted students were still short in achieving reading benchmarks (Burchinal et al., 2011; Gamse, Jacob, Horst, Boulay, & Unlu, 2009). Students who received Reading First instruction in Grades K-3 were the 7th graders who are subjects of this present study. Although there were gains in decoding or deciphering words in Grade 1, no significant effect on reading comprehension was found in Grades 1, 2, or 3 (Gamse et al., 2009).

Given the disappointing results of Reading First Schools, the district that participated in this research study began exploring alternate strategies for improving performance on reading tests. One such alternative pedagogical approach to improving reading test performance among struggling ethnic minorities was disaggregating data as the groundwork for test preparation (Tennessee State DOE, n.d.). Disaggregating the data

and determining area of need is vital, but after the data has been disaggregated, what specific strategy is efficient enough to streamline across curriculum? I suggest the 12 Powerful Words vocabulary program.

The local problem prompting this research study was the poor test performance of middle school Black students on standardized state measures of reading achievement. So, a possible lack of systematic test preparation may be a contributor to these students' test performance. Consequently, beginning in October 2012, the test preparation site (School A) began teaching test preparation skills through the use of the 12 Powerful Words created by Larry I. Bell (2005). Bell, an educational consultant, has identified 12 words that are most frequently used on standardized tests and are a source of students' difficulties when faced with these words on formal tests (see Appendix A). According to Bell, these 12 words obscure students' thinking and result in incorrect answers on state tests. Bell claimed that the daily integration of these words into core content and test preparation will boost students' academic achievement (Bell, 2005).

School A based the Powerful Words test preparation upon a specific goal to improve reading test performance of Black and other underperforming student populations (Bell, 2005). Unfortunately, the developer of this intervention only offers anecdotal data of program effectiveness. As such, no formal investigation of data is available.

I conducted this study within a large urban school district in the southern United States (Tennessee School District, n.d.). The participating district is comprised of 203 schools that served approximately 113,000 K-12 students, of whom 87% were Black, 9%

White, and 4% other races in [2012-2013] (Tennessee School District, n.d.). I conducted my research study at two sites. School A (the test preparation site) is a federally funded Title I middle school with an enrollment of 507 students, of whom 96% were Black in [2012-2013] (Tennessee School District, n.d.). All 17 core teachers at School A, met the highly qualified status (NCLB, 2002). Highly qualified status, given by NCLB, offers that teachers are current in licensure and endorsements to teach particular grade levels or subjects. The no test preparation site was School B, also a federally funded Title I middle school with an enrollment of 340 students, of whom 96% were Black (Tennessee School District, n.d.). All 21 core teachers at School B met the highly qualified status (NCLB, 2002). Schools are categorized by FP&G, as Title I based on low socioeconomic status (NCLB, 2002). The two locations are less than 12 miles apart in distance.

Research study participants included underachieving Black students in Grades 6, 7, and 8. I targeted these grades in this study because the middle school years are when students develop attitudes, values, and habits of mind that will largely direct their behaviors as adults (Cadeau, 2011; Coelho, Marchante, & Sousa, 2015; National Middle School Association [NMSA], 2010). Public record of a 3 year Reading average of academic achievement scores for School A, showed in 2010 a score of 35% below basic, 2011 a score of 33% below basic, and 2012 a score of 32% below basic. Thereby, School A fell short of the state benchmark for proficiency (Tennessee DOE, n.d.). For School B, public record of a 3 year Reading average of academic achievement scores showed in 2010 a score of 30% below basic, 2011 a score of 27% below basic, and 2012 a score of

25% below basic. School B also fell short of the state benchmark for proficiency (Tennessee DOE, n.d.).

As per federal guidelines under the No Child Left Behind Act (2002), when benchmarks are missed, the school district is mandated to provide specific supplementary instructional services. Supplemental educational services for reading improvement were implemented at both Schools A and B included Reading Plus, an online fluency program, which helps students read faster, for 90 minutes per week (NCLB, 2002); Stanford LAW, an online literacy program, which focuses on grammar, comprehension, and vocabulary, for 90 minutes per week (NCLB, 2002); and Title I after-school tutoring, a tutoring program that provides students with 40-70 hours of assistance per school year (NCLB, 2002). Title I assistance is contingent upon the amount of funds disseminated to individual school sites and time allotted for tutoring (NCLB, 2002). Students were identified for program placements based on how students performed on the state formative assessments.

Although the majority of both male and female students at School A scored below proficient, performance among Black males was lower than among Black females. Poor performance on standardized reading tests among Black males at the local level is consistent with the performance of Black males in the U.S. (Burchinal et al., 2011; Finkel, 2010; Johnson, 2014; West, 2012; Wood & Jocius, 2013). According to data from a 2013 NAEP report, male students in the 8th grade who were Black and eligible for free or reduced price lunch scored below the 25th percentile at a rate of 25%, which is

significantly different from students above the 75th percentile (Nation's Report Card, 2013).

Nature of the Study

The nature of this research study was quantitative. My purpose was to investigate the effectiveness of the Powerful Words test preparation program in order to inform curricular decision-making at a Tennessee middle school where teachers used a test preparation program. An intact-group comparison design was used to measure the effects of test preparation on reading test scores.

Purpose of the Study

Past research on test preparation for Black students has shown little or no growth in achievement due to lack in emphasis on particular test strategies and preparation methods. For example, Pan and Newfields (2011) showed that teachers who did not show efforts to prepare their students systematically for tests yielded no growth results for reading assessments. Witt (1994) determined that only insignificant score gains might be expected from test preparation. Further, Welsh, Eastwood, and D'Agostino (2014) found that student achievement benefits come from general instruction on the state standards. However, to refute (Pan & Newfields, 2011; Welsh et al., 2014; Witt, 1994) the current study investigated test conceptualization and a highly focused test preparation strategy that was taught to students. Therefore, it was anticipated that students receiving the treatment would score higher than equivalent students not receiving the treatment.

Research Question and Hypotheses

In this study, I sought to answer the following research question and test two corresponding hypotheses:

RQ: Will the Powerful Words test-preparation strategy that familiarizes students with key vocabulary terms related to test items improve reading test performance of underachieving Black middle school students?

H_0: There will be no significant differences in students' mean scores as a result of participating in the Larry Bell's 12 Powerful Words test preparation program for 7 weeks.

H_1: There will be significant differences in students' mean scores as a result of participating in the Larry Bell's 12 Powerful Words test preparation strategy for 7 weeks.

Theoretical Framework

The question investigated by the study was if the Powerful Words test-preparation strategy that familiarizes students with key vocabulary terms related to test items improves the reading test performance of underachieving Black students in middle school, then it is the foundations of test-wiseness that can improve the performance of urban Black students on standardized reading tests (Scruggs & Lifson, 1985). Black students were the focus for this research because out of 203 schools in the Tennessee research study district, approximately 113,000 K-12 students, 87% are Black (Tennessee School District, n.d.).

The influence of test-wiseness has been discussed in the literature for over 50 years. Test-wiseness was originally proposed as a probable effector, credible influence of reliability to measurements by Thorndike (1951). Thorndike identified test-wiseness as one of the grounds for variation observed in test score differences. Millman, Bishop, and Ebel (1965) described test-wiseness as the "capacity to utilize the characteristics and formats of the tests and/or the test-taking situation to receive a high score" (p. 87). Millman et al. suggested that test-wiseness is autonomous to the content area that is being tested. Similarly, Oakland and Weilart (1972) defined test-wiseness as a skill-set of practices test takers use in testing situations to achieve a proficient. Other terms for this trait include test sophistication and test-wisdom (Erickson, 1972; Preston, 1964)

Operational Definitions

The following terms are operationally defined as follows:

Reading achievement: Student performance on assessment questions intended to elicit varied cognitive processes and behaviors across a range of reading tasks (NAEP, n.d.).

Test preparation: A process by which students practice with different types of questions, vocabulary, or item formats that may be encountered on tests (Alon, 2010). Test preparation may also include learning how to spot clues for identifying the correct answer on a fixed-response test or learning when to guess (Alon, 2010).

Test-wiseness: A capacity used to receive a high score on tests (Millman et al., 1965). Test-wiseness may include a skill-set that has to be taught or learned (Ebel, 1965).

Underachievement: Underachievement in reading is defined as a student without a general learning disability performing below a proficient level on a standardized reading test. Although underachievement in reading may be due to decoding problems, reading difficulties may also arise from comprehension problems even when decoding skills are adequate (Allington, 2013).

Assumptions, Scope and Delimitations, and Limitations

Assumptions

An assumption of this study was that the measuring tool tested the reading objectives assigned by the district. According to Discovery Education, the DEA in reading used for Tennessee matches the eight reporting categories of the reading test of the Tennessee Comprehensive Assessment Program (TCAP): Language, Communications, Writing, Research, Logic, Information Text, Media, and Literature (Discovery Education, n.d.). Another assumption is that students attempted to do their best on the test based on instructional objectives to prepare for Discovery Assessment. The performance on the test affects students' course grade.

Scope and Delimitations

An important delimitation imposed by me was the decision to restrict test preparation to the Powerful Words test preparation program and not to include any other test preparation that might improve test performance. However, this delimitation was purposeful in that focusing on a particular aspect of test preparation will better inform our knowledge of one factor that may contribute to test-wiseness on standardized reading

tests. Another delimitation was providing students with only seven weeks to learn and practice the strategy.

The scope of this study was limited to underperforming urban Black students in Grades 6, 7, and 8 at a middle school in a large urban school district in the South of United States. Because Black students have for many years lagged behind in reading performance, underperforming Black students were identified as the target group of this study. Potentially, the results of this study may be replicated at other urban schools or with underperforming Black students, but generalization of the results of this study must be guarded. There were no inferences about other ethnic or racial subgroups, either at the site or elsewhere, from the present study.

Limitations

A potential limitation in this study is that participants at both schools were preselected based on low performance on previous formal reading tests. Additionally, student composition at both schools was limited in that both schools have above 96% Black students. Blacks made up about 96% of the sample, so there was no subgroup, other than Black, that had a large enough representation of participants. Thus, there was no opportunity to contrast the performance of Black students with peers from other ethnic groups.

Significance of the Study

In Tennessee, the low test performance is actually increasing (Tennessee DOE, n.d.). Reading aptitude is a fundamental skill that affects comprehensive learning and test performance. Across the state of Tennessee, Black students are more likely to be below

Proficient in reading (Tennessee DOE, n.d.). In relation to the local problem, 87% of the 113,000 students will be affected by programs such as Powerful Words test preparation or no test preparation. Therefore, application of this research study's concept to the local agencies must be planned, explicit, and consistent, such as Kontovourki and Campis's (2010) study in test genre course with a syllabus for urban students. Reading achievement on formative assessments is a strong indicator of graduating from high school, attending college, and being a productive citizen equipped with the skills for personal advancement (see, also, Grier-Reed et al., 2011; Keels, 2013; Lundy-Wagner & Gasman, 2011; Phillip, 2011; Wood & Palmer, 2013).

Summary

The present study tracked the effectiveness of a Powerful Words test preparation program, with the anticipation that it would improve the reading test performance of Black urban students. If the results show statistically significant or gained mean scores, then educational practice, particularly within the supplementary service curriculum of Title I schools could be impacted. In addition, this line of research may provide some insight into the degree to which standardized reading tests measure pure reading achievement or some combination of reading achievement and test-wiseness. Section 2 will review the literature on directed test preparation as a strategy to improve the test performance. Section 3 presents the research methods.

Section 2: Literature Review

Introduction

Test-wiseness as a theoretical construct is not a general pedagogy. First introduced by Thordike (1951) as a variance to assessment data, it was then modified by Ebel (1951) who stated that test-wiseness was a taught skill-set. Additionally, Oakland and Weilart (1972) expanded the concept to a skill-set that promoted test proficiency. Finally, Scruggs and Lifson (1985) applied the thought to Black students as a way to close the racial gap with reading proficiency. The following review of the literature shows evidence of test preparation, and how it is effective. I focused on the effectiveness of test preparation for students at all learning levels, minority students, and low-achieving students.

Literature Search Strategy

I used three databases to conduct the search: ERIC, Educational Research Complete, and SAGE. I used the search terms *test-wiseness*, *test preparation*, *test prep*, *testing in language arts*, and *test familiarization* and restricted my search to the years 1960-2015. A wide-spread database search was needed because much of the research on test-wiseness was conducted prior to 1980 (Oakland & Weilart, 1972). My searches initially yielded a total of 5,557 results for both peer-reviewed and non-peer-reviewed articles. The topic test-wiseness presented (950 articles), test preparation displayed (1026), test prep showed (184), testing in language arts exhibited (3,374) and test familiarization revealed (23). I eliminated non-peer-reviewed articles and articles whose authors did not discuss specific strategies to prepare students for engaging test-taking

behaviors. Thus, in this review, I incorporated 35 articles, emphasizing research conducted within the past 6 years, but also including earlier, seminal research. The review is organized into six parts: effects of test preparation on test performance; strategic approaches to test-taking; preparation of test-taking strategies; the role of technology as test preparation strategy; the role of motivation for test preparation; test-as-genre studies; and a brief overview of research methodologies used in the test preparation literature.

Review of the Literature

Effects of Test Preparation on Test Performance

Researchers have suggested that test preparation may improve students' test performance. The literature in this section discussed the effects of test preparation in training for high school standardized examinations for low-achieving minority students. The minority subgroup includes those students who are too considered either traditional or nontraditional status.

In a longitudinal research study, Alon (2010) tracked and compared racial minority students to White students who were coached separately on SAT test strategies through private tutors and classes. The test preparation included vocabulary and test formatting practice. Subsequently, the study focused on Black students and found that there were notable inequalities in test preparation for particular subgroups of Black student participants. Overall, Alon (2010) concluded that test preparation boosts test proficiency on formative assessments of content areas. However, there must be equity in

motivation for learning test-wiseness, and in the delivery of the test preparation practice or program.

Ebanks et al., (2011) developed an equalizing test preparation program titled Project 2011, which was tailored to help students pass the New York City's Specialized High School Admissions Test for Black students. As participants in this program, students were offered test preparation through vocabulary and motivation, similar to this current study. Results showed that students dramatically improved in their test performance. Because of Project 2011, students were able to enroll and participate in the city's most elite educational programs. In essence, this allowed for students to be immersed in culturally diverse learning communities. Diversity is imperative, and permits for Black students that are gifted to reach their maximum potential (Winsler, Gupta, Karkhanis, Deepti, & Levitt, 2013). Consequently, the goal of teachers or schools providing test preparation is to close the proficiency gap in reading.

Similarly, Bicak (2013) developed a test preparation course with a Test Taking Strategies Scale (TPTTS) to evaluate test preparation and the test-taking strategies used. The participants of the study were students from various public high schools who wanted to pass college/university entrance exams. Student exam scores increased in relation to consistently answering questions correctly from 39% to 78%, dependent upon the number of items tested (Bicak, 2013).

Misco (2010) focused on 10 ethically mixed students who were high school juniors and seniors. These students had failed an exit exam that preceded participation in Misco's test preparation course. Due to the failure of the exam, prior to the test

preparation course, teachers remediated the course curriculum by tailoring student learning to reasoning, test-taking skills, and verification of answers with text. Results of the test preparation course yielded that all 10 students passed the exit exam. Misco thereby concluded that test preparation should be part of the general curriculum, and could be applied to subgroups such as minority students who were also English language learners.

Jiang and Elen (2011) sought to find if the instructional effectiveness of test preparation could be accomplished with scaffolding questioning techniques within a text. Out of the 42 student participants, they found that 82% of students activated their prior knowledge when reading text, and 76% of students stated that the use of the instructional devices helped in answering the questions from the text (Jiang & Elen, 2011). Therefore, students' scores increased based on the method of combining test preparation and prior knowledge.

Overall, research (Ebanks et al., 2011) indicated that test preparation can improve performance of minorities, but also traditional and nontraditional status students when a targeted and explicit approach is followed (see, also, Alon, 2010; Misco, 2010; Mizumoto & Takeuchi, 2009; Jiang & Elen, 2011). Beyond direct studies of test performance, the literature on test preparation has also expanded to include three emerging lines of research: test-taking strategies, the role of technology, the role of motivation for test preparation, and tests-as-genre studies.

Strategic Approaches to Test-Taking

Test preparation may improve assessment performance by modifying the students' strategic approach to a test. The following studies used various applied deliberate methods for students who were learning English as a second language. As a conclusion to tactical implementations, researchers offered that oral language can too benefit.

Mizumoto and Takeuchi (2009) examined 146 students at two Japanese universities that studied English as a Foreign Language. The authors investigated the effectiveness of 10 weeks of explicit vocabulary strategies by administering a pretest and posttest for a treatment and a control group. The findings showed that the experimental group had a higher score gain than the control group with the vocabulary test.

Yang (2000) compared test-taking strategies and exam performance with experienced and inexperienced test-takers. Yang probed these students to employ test-wiseness in responding to a Test of English as a Foreign Language Practice (TOEFL) Test B. Twenty-three out of 390 Chinese TOEFL students were identified as test-wise and 17 were identified as inexperienced test-takers. The 23 students who were considered to be experienced in test-wiseness performed better on test items than the 17 inexperienced students. Qualitative interviews later showed that experienced test-takers were test-wise in making meaning from questions. They were more analytical, logical, and systematic in answering test questions than inexperienced students.

Continuing with English language learners, Zhengdong (2009) conducted a study of 146 students who took an International English Language Testing System (IELTS) test

preparation course. Participants joined in the study for a total of 5-10 weeks and received 10-20 hours of test preparation. Of the 146 participants, 56 stated that they had already taken the course at the University Language Center. After the test preparation course, data from the study were collected through a questionnaire and interviews. The results showed that all students who took the IELTS test preparation course scored higher than the students who did not take any test preparation course. This strategy can then be applied to language specific assessments.

Farnsworth (2013) observed test preparation in relation to language assessment and instructional practices; oral English proficiency in a second language program. The tests used in the study were the Basic English Skills Test Plus (BEST Plus) and the Versant English Test (VET). The results revealed that students in the VET coached group displayed an increase in VET scores by 4.56 points on an 80-point scale while the BEST Plus coached group revealed an increase in scores by 63.24 points (Farnsworth, 2013).

Test Preparation Strategies

Many different learning institutions or programs use methods of test preparation strategy. Some programs focus on markings or grammatical clues found in the question. Others emphasize potential performance theory. However, what is common is that teachers must be inspired to deliver the strategy with fidelity to students of all learning levels, adult to Grades K-12. The role of motivation for test preparation is discussed in detail (see page 29).

Hoffman and Nottis (2008) conducted a mixed methods study that investigated adolescents' perceptions of strategies with high stakes tests. Participants included 215

Grade 8 students. Out of the 215, there were two major groups of students: those who received test preparation through self-efficacy efforts and those that participated in test preparation with rewards as motivation. Results showed that students were more motivated by extrinsic rewards for participating and being proficient on tests than those who were relying on self-efficacy to be proficient on tests.

Stough (1993) focused on how students responded on answer sheets to determine test-wiseness. Participants consisted of 90 college freshmen who were members of a learning strategies course. As a pre-identifier, students answer sheets were marked for wrong answers. Test-wise strategies included teaching students to avoid words in tests such as always and never, avoid unrelated alternatives, and pay attention to grammatical clues (Stough, 1993). After participating in the course, students' markings were found to be more capable, which thereby increased scores. The author concluded that as a result of the learning strategies course, students learned to process test information, which made them test-wise.

In Rice, Geels, Trafimow, and Hackett's (2011) study, there were 346 undergraduate students from a large Southwestern university. The authors applied potential performance theory (PPT), developed by Trafimow and Rice (2008), to analyze student scores. The results indicated that consistent use of strategies over time affects students' total testing performance. Thereby, most likely enhancing students' confidence and motivation to be proficient. So, how can working adults use test preparation strategies to improve test-wiseness for career advancement?

Chung-Herrera, Ehrhart, Ehrhart, Solamon, and Kilian (2009) examined test preparation strategies for Black and White adults in police and fire departments. The authors found that surveyed Blacks were more motivated with test preparation. The self-initiation was seen in tutorial attendance and enthusiasm to perform well in preparation for career advancement tests. There were 149 participants included in this study. Overall, 87 or 58% of the participants were Black. Within jobs and based on results for those who took the tutorial, 34 Blacks became candidates for police sergeant, 23 Blacks for fire lieutenant, 10 Blacks for fire captain, and 17 Blacks for police lieutenant. Chung-Herrera et al. reported that 84 out of the 87 Black participants were placed in desired promotional jobs.

Conversely, Matsumura, Garnier, and Resnick (2010) investigated a test preparation program named content-focused coaching (CFC). This test preparation program was designed to coach teachers, who then coached students for test preparation. Fifteen elementary schools were randomly chosen and trained for CFC in a study that lasted for one school year. The participants included 96 fourth and fifth grade teachers, and 63 students. At end of year, all 63 student participants were assessed as proficient. Matsumura et al. administered a supplementary survey to the teachers. Results from the survey indicated that teachers perceived that the test preparation training benefitted their students and their pedagogical practice (Lam, 2013; Matsumura et al., 2010).

Writing too has been noted as an aspect of test preparation strategy. Bicer, Capraro, and Capraro (2013) studied writing as a process for test preparation in helping students understand and build with problem solving skills. 96 students participated in a 6-

week study, randomly assigned into two groups. One group focused solely on writing with mathematical problem solving. The other group focused on homework and high-stakes test preparation with mathematical problem solving. The results from the study indicated that 85% of the students in the writing mathematical problem-solving group performed better than those in the high stakes test preparation class (Bicer et al., 2013).

The Role of Technology as a Test Preparation Strategy

Another link of research in test preparation strategies are the roles of technology in test preparation. Technology can be seen with use of online paced courses with designed test preparation syllabus. Other technology tools are clickers for adult learners and well as school aged students. The studies below show evidence of designed technology learning programs that work.

Kawasaki and Tominaga (2010) developed an engineering college course, using a tProgrEss computer program for beginning students. This program scored student responses, and determined the rate to which students were proficient on the test given. When in test preparation mode, tProgrEss notified students of their answer responses immediately. Results showed that all students were able to progress to intermediate course levels.

Gadbury-Amyot, Austin, and Overman (2013) studied test preparation for the U.S. National Board Dental Examination (NBDE) which was taught online by faculty content experts. A total of 592 students from 32 states participated in the test preparation course. Students could review the test preparation material at their pace. At the end of the online course, Gadbury-Amyot et al. reported that the 592 student participants were

driven to spend more time in preparation for the NBDE, due to the readily available online technology.

Dietz-Uhler and Lanter (2009) used a Web-based test-preparation technique that directed students in an introductory psychology class to answer four questions requiring them to engage, analyze, reflect, and apply reading material before taking a quiz on content. Participants were separated into two groups. One group consisted of those who answered the four questions using the technology before the quiz. The other group were those who took the quiz, without using the technology. Students who responded to all the questions, using technology prior to the quiz, performed better than those who did not (Dietz et al., 2009).

Chen and Lan (2013) studied the use of technology based clickers for 151 undergraduate students taking basic chemistry classes at the National Chung Hsing University in Taiwan. This study focused students' perceptions of the use of the Student Response System technology (SRS) clickers as an assessment tool in class. Chen and Lan (2013) reported that perceptions focused on how clickers affect the understanding of course content, engagement, and test preparation. Results showed that learning by way of clickers rendered positive benefits to test preparation and instructional effectiveness. All 151 students showed an increase as a result of using the SRS (Chen & Lan, 2013).

Fike, Fike, and Lucio (2012) conducted an intervention study for pre-service teachers to determine if a methods class had an impact on their ability to learn what was needed to pass a Texas state exam, and earn a professional licensure for their teaching careers. The intervention study like Chen & Lan (2013) used SRS, clickers, as a test

preparation tool. Participants took a mid-term and end-term exam for data results. Fike et al. (2012) reported that one third of the exam questions were reviewed in class using clickers; the second third of the questions were reviewed without the use of clickers; and the final third of the questions were not reviewed in class. Students stated that they were more engaged when clickers were used for review. Hence, the authors concluded that the use of clickers for pre-service teacher education classes provided engagement, and test preparation with the SRS tool was beneficial.

Espey and Brindle (2010) also researched the effects of SRS. Clickers were used in this study as an instructional tool to increase student engagement in preparation for exams. The clickers provided quick feedback for students and teachers to see data results, immediately. Espey and Brindle (2010) suggested that there are positive benefits from the interactive uses of clickers at college levels and Grades K-12. Even though clickers are very popular in many learning setting, designed technology programs are great tools for test preparation as well.

Ruberg (2001) sought to determine if the BioBlast CD-ROM computer program was an effective test-preparation tool for students. The study incorporated twenty-one schools from the United States and Singapore that utilized BioBlast CD-ROM as a way to coach students for test preparation as well as proficiency on tests. Ruberg's analysis of student performance on pre- and posttests showed that the implementation of BioBlast CD-ROM computer program was associated with student gains with particular schools and classes. A year later, Ruberg compared scores of the same schools and students with an additional year of pre- and posttests (Ruberg, 2001). However, there were changes in

the implementation strategies of the BioBlast CD-ROM computer program. Nevertheless, Ruberg concluded that, overall, students that were able to use the BioBlast software, showed performance gains appropriate to their knowledge level and experience. It is implied that the classes were heterogeneous. Therefore, students with disabilities can too benefit from test preparation (see also Dietz-Uhler & Lanter, 2009; Gadbury-Amyot et al., 2013; Hughes, Schumaker, Deshler, & Mercer, 1988).

Hughes, Schumaker, Deshler, and Mercer (1988), and Hughes and Schumaker (1991) developed a test-taking strategy specifically for secondary students with disabilities. Lancaster,et al.(2009) then adapted the test-taking strategy to a computerized program to teach test preparation to secondary students with disabilities. Classes were isolated to determine the effects of the program. The use of test-taking strategies, including familiarization, vocabulary, and ability to think or engage aloud with tests, helped students in being test-wise. A report of results from the study showed that there were statistical differences found in the computerized program, and it was effective in teaching students to use test-taking strategy (Lancaster et al., 2009).

Fede, Pierce, Mathews, and Wells (2013) comparably investigated the effects of a computer-assisted test preparation instructional program that focused on word problem-solving skills for 5th grade students that perform below proficiency. A total of 32 students participated in a randomized controlled study. Sixteen students were given a 12-week computer-assisted test preparation instructional program. The other 16 students received a 12-week general state test preparation. All students completed the Massachusetts Comprehensive Assessment Test. Results showed that the computer-

assisted test preparation instructional program group showed greater gains than the control group on general state test items from the Massachusetts Comprehensive Assessment Test (Fede et al., 2013).

As a conclusion to technology, Bayer-Hummel followed Mangold's (2007) fundamental concept of gaming for research study participants. Bayer-Hummel (2009) suggested that the electronic game Jeopardy was a fun test preparation strategy. Therefore, it motivated students in test preparation and improved their performance.

The Role of Motivation for Test Preparation

Only limited attention has been given to the preparation of test motivation. Bayer-Hummel (2009) proposed that educators generally should strive to motivate and inspire students to participate actively in the learning process. If teachers' attitudes towards test preparation were positive, then learning could be self-regulated (Lam, 2013).

Lam (2013) investigated the perspectives of teachers and students regarding test preparation strategic approaches, as well as how test preparation could impact performance on formative assessments. Participants consisted of 2 teachers and 14 students in Grades 2 and 3. Interviews were conducted in which questions were geared towards teacher attitudes on utilizing test preparation approaches. Results showed that participants did believe that attitudes regarding test preparation could enhance student performance and promote self-regulated learning (Lam, 2013). Therefore, perspectives on test preparation, delivery, and motivation matters.

Turner (2009) incorporated motivation into a test preparation program for middle school students. After a year of tiered lessons for test preparation combined with attention

to test motivation, the pass rate on standardized reading tests increased from 46% to 74% (Turner, 2009).

More recently, Xie and Andrews (2013) introduced expectancy-value motivation theory to explain influence of motivation on test design and student test preparation. In Xie and Andrews's study there were 870 test takers in a college English class in China. At the beginning of research, a perception assessment and questionnaire were administered to gauge motivation. The study lasted 10 weeks. At the end of 8-weeks, the test preparation questionnaire was re-administered to reassess motivation. Results showed that test takers who perceived test preparation positively placed high importance on the assessment, were more confident, and increased test performance by 36% (Xie & Andrews, 2013). As a result, there is an expectation that test preparation with motivation is effective and contributes to higher test score achievement (see, also, Alon, 2010; Bayer-Hummel, 2009; Chung-Herrera et al., 2009; Ebanks et al., 2011; Hoffman & Nottis, 2008; Kawasaki & Tominaga, 2010; Mizumoto & Takeuchi, 2009; Turner, 2009; Xie & Andrews, 2013).

Test-as-Genre Studies

Other researchers have focused on improving test scores by introducing students to the concept of tests as a genre. Therefore, assessments could be considered a core class categorized by test-wiseness. This reiterates that standardized tests require a specific knowledge and performance of thinking skill-sets (Hornof, 2008).

Doe and Fox (2011) studied three second language (L2) test takers' test preparation genre course strategies that were employed for 6-weeks across testing

contexts. Participants completed a practice test, standardized test, posttest, and subsequent interview. Doe and Fox (2011) reported that the experience of test preparation classes enabled students to acquire test-wise strategies that would make them efficient test takers.

Hornof (2008) similarly theorized that standardized reading tests should be taught as a genre, and developed a two-week unit for 4th grade students to prepare for standardized reading tests. The author taught selection of answers through modeling elimination of wrong answer choices. Charts were created and used to help students visualized how to eliminate. Then, Hornof emphasized stamina, making practice assessments longer over time to increase student engagement with tests. As well, Hornof proposed that only studying test preparation materials help students be better test takers. The results showed that test takers scored more efficiently on each progressive test.

Kontovourki and Campis (2010) studied an urban 3rd grade classroom in which test preparation was approached by means of genre study. Students received test preparation that emphasized test skills, test language, and building stamina as seen in Hornof (2008). The authors used observations, interviews, and artifacts to analyze how students evolved as test-takers. Kontovourki and Campis (2010) concluded that test preparation can become meaningful to students and can reposition student self-concept from one of failure to success.

Inclusively, the foregoing research on test preparation and motivation suggests that direct attention to preparing students for tests can improve test performance across a wide range of student populations. However, not all research on test preparation has

reported positive results. For example, Pan and Newfields (2011) studied English as a foreign language (EFL) proficiency tests and found that both students and teachers may undermine efforts to prepare for tests. The authors collected data from 160 teachers and 1,415 student surveys from various school locations. Not all school locations required English proficiency to complete the learning program. Pan and Newfields (2011) also interviewed 25 teachers and 38 students. Amongst the schools participating in the study, only interviews from the English speaking proficiency schools showed that test preparation instruction was more readily available. Therefore, Pan and Newfields (2011) found that minimal impact was made on overall student learning of test preparation for English proficiency, because there was insignificant teacher willingness to add test preparation to class instruction.

Welsh et al. (2014) observed that accountability with student proficiency on high-stakes tests are universal within the United States. Therefore, due to high stakes tests, the assumption is that there is more test preparation time allotted during the instructional day than traditional instructional time. Accordingly, the authors investigated the relationship between student test performance and test preparation activities in 32 third and fifth-grade classrooms (Welsh et al., 2014). After using test preparation items similar to the state test, test format practice, and teaching test-taking skills, Welsh et al. (2014) found that student achievement would be greater from general instruction on the state standards rather than from isolated test preparation.

Witt (1994) attempted to resolve contradictory results of researcher analyses regarding the effectiveness of test preparation. The author determined that only

insignificant score gains might be expected from test preparation. Further, consistent with the findings of Kulik et al. (1990), Witt proposed that test preparation might be more effective when a pretest is administered as part of the test preparation program. The pretest would gauge the levels of students' ability before test preparation begins. Pearlman (1984), who was cited in Witt's (1994) study, too concluded that much of the variance reported by Kulik et al. was insignificant score gains, as an expectation from test preparation.

Previously Employed Methodologies to Measure Test-Wiseness

Throughout the literature on test preparation, four research methodologies are most common: (1) single group pretest/posttest, (2) control group versus experimental group, (3) mixed method with survey plus pretest/posttest, and (4) longitudinal to measure performance over time. Single group pretest/posttest has the advantage of measuring change in performance, although it may be difficult to infer that change is attributable to a treatment as opposed to maturation (Creswell, 2012). For example, single group pretest/posttest designs have been used to measure changes in test performance after test preparation (Chen & Lan, 2013; Chung-Herrera et al., 2009; Dietz-Uhler & Lanter, 2009; Ebanks et al., 2011; Espey & Brindle, 2010; Farnsworth, 2013; Fede et al., 2013; Fike et al., 2012; Gadbury-Amyot et al., 2013; Hornof, 2008; Jiang & Elen, 2011; Kawasaki & Tominaga, 2010; Kulik et al., 1990; Lancaster et al., 2009; Misco & Thomas, 2010; Rice et al., 2011; Ruberg, 2001; Stough, 1993; Turner, 2009; Welsh et al., 2014). Although control group designs can provide a much stronger inference, results may be attributed to treatment (Bicer et al., 2013; Yang, 2000). Control group designs

have been much less frequently used in studies of test preparation (Bicer et al., 2013; Yang, 2000). More often, researchers included pretest/posttest results as part of a mixed methods design in which the quantitative test results were augmented by qualitative survey data (Hoffman & Nottis, 2008; Kontovourki & Campis, 2010; Lam, 2013; Matsumura et al., 2010; Mizumoto & Takeuchi, 2009; Pan & Newfields, 2011; Xie & Andrews, 2013; Zhendong, 2009). Seldom have researchers of test preparation used longitudinal data to measure change over extended time periods (Alon, 2010; Doe & Fox, 2011).

The present study attempted to overcome the shortcoming of a single group pretest/posttest design by introducing a matched control group. This matched control group provided a higher level of inference. Inferences are attributed to treatment. In this study, I used an analysis of covariance to compare treatment and non-treatment scores on the December administration of the DEA reading test, after 7-weeks of treatment, with initial status in reading functioning as a covariate. Results of this study could lead to informed decisions regarding test preparation at the study site and study sites similar in demographics. This study also suggests approaches for reducing low test scores on standardized reading tests.

Summary

The present study investigated the effectiveness of preparing urban Black students on a test preparation strategy intended to improve test performance on a standardized reading test. The literature on test-wiseness reviewed in this section suggests that it is an attribute that all learners can acquire. Collected works on test preparation suggests that

with targeted approach, students can improve in test performance. However, and although few, discrepancies in the literature on test preparation strategies appear to be the result of deficiency of fidelity in implementation of the test-preparation strategy. For example, past research on test preparation for Black students has shown little or no growth in achievement due to lack of emphasis on a particular test strategy or preparation method (Pan & Newfields, 2011). Though, what is understood is that strategies must be planned, explicit, and consistent, such as in Kontovourki and Campis's (2010) study in which researchers developed a test genre course with a syllabus for urban students. In another example of explicit test preparation, Doe and Fox (2011) prearranged and implemented a genre test course for students with non-native speakers of English. Similarly, Gadbury-Amyot et al. (2013) implemented an adult distance-learning program for explicit test preparation, and Lancaster et al. developed a well-organized computer-based test preparation program for students with disabilities. While research is moving beyond direct studies of test performance, conclusions regarding test-taking strategies, the role of technology, the role of motivation for test preparation, and tests-as-genre are still tentative.

Section 3 presents the study's research design, sample, setting, and methods by which data will be collected and analyzed.

Section 3: Research Method

Introduction

The local problem prompting this study was the poor performance of Black middle school students in Tennessee on standardized measures of reading achievement. I employed an intact-group comparison design to measure the effects of test preparation on reading test scores during the 2012-2013 academic year at two neighboring middle schools in Tennessee, both of which are in the same district. Educators at both schools participated in a Title I supplementary instructional program. School A served as the test preparation site, and School B served as the no test preparation site.

The school district which participated in the study was zoned by the local zoning board into four regions. Both schools are in the same urban region and served similar students. There is robust familiarity with the participating district, because the school district selected me as Executive Leader in 2011-2013. As part of my executive leader training, I was required to visit all schools in the district, and provide feedback about their academic learning communities.

Research Design and Approach

I used an intact-group comparison design in which the test scores of students receiving the test preparation treatment in School A was compared to test scores from equivalent students in School B who did not receive the specific treatment. It was not possible to randomly assign students to groups, because students were pre-identified by the school district as non-proficient, and the great majority of the student population was classified as non-proficient. Therefore, a true experimental design could not be used as is

common in social and educational research (Lodico, Spaulding, & Voegtle, 2010). Customarily, the intact-group comparison design features two groups, as seen in this research study. One group is exposed to the treatment while the other group is not exposed to the treatment (Lodico et al., 2010). Therefore, use of an intact-group comparison design enables a researcher to compare posttest scores of two matched groups (Lodico et al., 2010), which is pertinent to this current study.

In order to assure equivalence of intact groups, I used two approaches. First, I matched the two schools (School A and School B) for location, neighborhood, ethnicity, socioeconomic status, student achievement, and teacher experience. I have confidence that I selected schools that had similar student compositions. Secondly, I compared the initial status in reading, which I statistically controlled by using an ANCOVA (Lodico et al., 2010; Mertler & Vannatta, 2005). Public record of the state report card of academic achievement scores showed that in 2010 School A received an overall reading score of 35%. While in 2011, School A earned an overall reading score of 33%, and in 2012, an overall reading score of 32% (Tennessee DOE, n.d). Reading state report card of academic achievement scores showed that in 2010 School B earned an overall reading score of 30%. In 2011, School B earned an overall reading score of 27%, and in 2012, an overall reading score of 25% (Tennessee DOE, n.d.).

The school district's regional zone, in which this study was done, had six school closures due to poor academic achievement at the end of the 2013 school year (Tennessee School District, n.d.). This was followed by another six school closures due to poor academic achievement at the end of the 2014 school year (Tennessee DOE, n.d.). As of

the 2015-2016 school year, the only district schools still operating are Schools A and B. Therefore, the selection of the two schools to compare for this study was accurate.

Setting and Sample

I selected School A for the test preparation treatment based on a district completed analysis of underachievement in reading, as measured by DEA and state reading tests. To determine the minimum number of student participants required, I conducted a G* power analysis (see Mayr, Buchner, Erdfelder, & Faul, 2007). The input parameters were: *F* tests, mean difference between two groups of approximately equal size, a priori, and two-tailed. For a medium effect size ($d = 0.5$), an α error probability of 0.05 and a power (1-β error probability) of 0.95. The resulting minimum sample size was 210.

The School A initial sample comprised 501 out of 507 total students. Six students are not a part of the district's Excel archival data that tracks non proficient students. Those students comprised 1% of the sample, were proficient or advanced in reading. At School B, the initial sample comprised 282 out of 340 total students. Fifty-eight students were not a part of the district's Excel archival data that tracks non-proficient students. Those 17% of students were proficient or advance in reading. Participants from both sites included students with no special needs and students with special needs or needing inclusion.

The total student aggregate for this study was 785. However, some students' pretest or posttest scores were missing. Students who had a missing DEA scores, I

excluded. After identifying students with missing scores, the total sample size was 680 students, thus exceeding the minimum sample size required for G* Power analysis.

The sample included 654 Black students in Grades 6, 7, and 8. All students in the sample received free or reduced priced lunch. Of the selected study participants, 104 (15%) were Black students with disabilities. In addition, 1% of students in this study were labeled as other, and 0% of the students were labeled as gifted and talented. Detailed demographic data for research participants show gender, race and ethnicity, English language learners, low socioeconomic status, and students with disabilities are provided in Table 2 for School A and Table 3 for School B.

Table 2

School A: Demographic Results For Test Preparation Participants

		Grade 6		Grade 7		Grade 8	
Demographic Variable		n	%	n	%	n	%
Gender	Male	86	.57	79	.59	71	.47
	Female	64	.42	53	.40	77	.52
	Total	150	100	132	100	148	100
Race/ethnicity	Black	142	.94	122	.92	141	.95
	Hispanic	8	.05	10	.07	7	.04
	Other	0	0	0	0	0	0
	Total	150	100	132	100	148	100
English language	Yes	0	0	1	.007	2	.01
learner	No	150	100	131	.99	146	.98
	Total	150	100	132	100	148	100
Low socioeconomic	Yes	150	100.0	132	100.0	148	100.0
status	No	0	0	0	0	0	0
	Total	150	100	132	100	148	100
Disability	Yes	23	.15	19	.14	23	.15
	No	127	.84	113	.85	125	.84
	Total	150	100	132	100	148	100

Table 3

School B: Demographic Results For No Test Preparation Participants

		Grade 6		Grade 7		Grade 8	
Demographic Variable		n	%	n	%	n	%
Gender	Male	39	.51	49	.56	46	.52
	Female	37	.48	37	.43	41	.47
	Total	76	100	86	100	87	100
Race/ethnicity	Black	75	.98	85	.98	87	100
	Hispanic	0	0	0	0	0	0
	Other	1	.01	1	.01	0	0
	Total	76	100	86	100	87	100
English language	Yes	0	0	0	0	0	0
learner	No	76	100	86	100	87	100
	Total	76	100	86	100	87	100
Low	Yes	76	100	86	100	87	100
socioeconomic	No	0	0	0	0	0	0
status	Total	76	100	86	100	87	100
Disability	Yes	14	.18	10	.11	15	.17
	No	62	.81	76	.88	72	.82
	Total	76	100	86	100	87	100

Teachers instructed both groups using methodologies already approved by the building principals, who had autonomy for curricular decisions, with the exception of mandated Title 1 supplemental academic programs, see page 10 (NCLB, 2002; Tennessee State DOE, n.d.). Moreover, as per Title I requirements, student or teacher absenteeism cannot fall below 93% of the daily instructional hours (NCLB, 2002). So, absenteeism did not affect the data analysis.

Treatment

Students in Grades 6, 7, and 8 at both School A and School B sites received the same district-mandated reading curriculum throughout the school year and the same district-mandated supplementary reading remediation as part of the Title I program. In addition, students at School A, the test preparation site, received an auxiliary instruction for test preparation, called the Powerful Words test preparation program (see Appendix A). Other districts with schools having similar demographics and performance scores have used the Powerful Words test preparation program, but only anecdotal evidence is provided (Bell, 2005).

To complement the in-class mini-lessons, the treatment also involved me providing 15 minutes of performance on the school's intercom at the start of instructional day for motivation that included the Powerful Word song, the spelling of the words, and reciting the definition of words. To assure fidelity of implementation of the Powerful Words program, I observed each classroom daily as part of monitoring for district quality checks for instruction. I sat in the back of the room and did not participate in the instruction nor did I interact with any students. In addition, the Instructional Coach and

administrative team checked the lesson plans from all 17 core content teachers weekly to ensure focus on lesson delivery aimed to include the 12 Powerful Words test preparation strategy, and to incorporate reading for non-reading classes.

In contrast, School B did not complement the district mandate for core instruction and participated only in the district's mandated curriculum scope and sequence. In addition to core instruction at both sites, qualifying students at the treatment and non-treatment schools participated in remedial Title 1 instruction. Details of the Title 1 supplemental instruction is explained on pages 10 and 11 of this study.

Treatment occurred for seven weeks during the 2012-2013 school year from the beginning of the year to approximately mid-year. After the seven-week treatment ended, students completed the DEA and their test scores were stored in the districts database. Because the students completed the DEA at the beginning of the year, there are DEA scores for two points during the year (before treatment and after treatment).

Variables

The midyear DEA reading test scores were the dependent variable. The independent variable was the participation in the Powerful Words test preparation program. The independent variable (participation in test preparation program) was a dichotomous variable measured on a nominal scale, while the dependent variable and covariate (DEA) were on an interval scale.

Instrumentation and Materials

Reading test performance was measured by the reading subtest of the DEA, a timed, multiple-choice test that measures skills in reading, mathematics, and science

(Discovery Education, n.d.). The DEA is administered to all students in the school district three times during the year as both a tool for remediated instruction and as a predictive test for the TCAP. All schools must adhere to district's testing times and schedule.

The reading subtest is administered in two sections of 75 minutes each. The two sections contained test items categorized into seven areas: language, vocabulary, writing and research, communication and media, logic, informational text, and literature. Each student received a raw score for the seven reading categories and a reading composite raw score. The number of reading passages presented could change from year to year, but in general there are four to five long passages per section. Students responded to test items in a multiple-choice format (Discovery Education, n.d.). The test is timed and students are credited for the number of correct answers with no penalty for guessing. Raw score data are thus on an interval scale based on number correct. Discovery Education, a subsidiary of Discovery Communications, is responsible for test construction, scoring, and reporting student performance. Students at both sites took the district required DEA Reading Test at the beginning and middle of the school year. The test results were available and released as archival data.

Content Validity

All DEA tests are state-specific (Discovery Education, n.d.). To ensure content validity, the publisher aligns test content to the state's content standards and the content sampled by the state assessment test (Discovery Education, n.d.). The publisher also uses the Webb Alignment Tool to support alignment with state-specific content standards in both the number of standards objectives sampled; and the cognitive complexity of

standards and objectives (Discovery Education, n.d.). According to Discovery Education, the DEA in reading used for Tennessee matches the eight reporting categories of the reading test of the Tennessee Comprehensive Assessment Program (TCAP): language, communications, writing, research, logic, information text, media, and literature.

Criterion Validity

Although the Discovery Education Assessment (DEA) is designed as a formative assessment tool to guide classroom instruction, it has been shown to significantly predict scores on the reading test of the TCAP. In a validity study conducted with 40,000 students in the participating School District in 2009, Discovery Education (n.d.) significantly predicted scores on the TCAP reading test at the .01 level. For Grade 6, the correlations between three administrations of the DEA and the spring TCAP reading test were .73 (Form A), .72 (Form B), and .69 (Form C) (Discovery Education, n.d.). For Grade 7, the correlations between three administrations of the DEA and the spring TCAP reading test were .74 (Form A), .71 (Form B), and .71 (Form C) (Discovery Education, n.d.). For Grade 8, the correlations between three administrations of the DEA and the spring TCAP reading test were .71 (Form A), .72 (Form B), and .71 (Form C) (Discovery Education, n.d.). The ethnicity frequency for middle school Black students is 46.43% (Discovery Education, n.d.).

Reliability

A reliability study of the DEA for Grade 6 Reading was conducted in 2009 with a sample size of 29,253 students. The Cronbach's alpha, a measure of internal consistency, a reliability coefficient of .81 was determined (Discovery Education, n.d.). The same

reliability study of the DEA for Grade 7 Reading was conducted in 2009 with a sample size of 29,182 students. The Cronbach's alpha reliability coefficient was .76 (Discovery Education, n.d.). Again for Grade 8, a reliability study of the DEA for Reading was conducted in 2009 with a sample size of 29,043 students. The Cronbach's alpha reliability coefficient was .84 (Discovery Education, n.d.). Discovery Education reported no data on test-retest or parallel-forms reliability.

Data Collection Procedure

All students at School A and School B sites were administered the DEA by their certificated homeroom teacher, in accordance with procedures directed by the Tennessee DOE. The pretest was given on October 2-5, 2012, and the posttest was given on December 12-14, 2012. Student answer sheets were forwarded to DEA for scoring and reporting of results. Discovery Assessment books and students' scores are discarded at the end of every school year by school site (Discovery Education, 2008). Discovery Assessment is only used to gauge student growth in preparation for state test for current school terms.

The overall results of this study were taken from the district's Excel report for Discovery Tests A and B. A data file request form was submitted by me to the local school district on March 5, 2015. The approval of the data file was returned to me from district on March 10, 2015. Results were given by Discovery Assessment Education to the local district, and subsequently to me. In addition, I signed a data usage agreement on March 14, 2015 and on March 17, 2015 was signed, then returned by the participating district's planning and accountability representative. All data used in this research study

was archival data. As of the end of March 2015, I am the only person with this data. The district and Discovery Assessment disposed of results after three school terms of initial reporting.

Role of the Researcher

I was an Instructional Coach at School A. My main roles as an Instructional Coach were to provide teacher professional development and monitor instruction, predominantly reading. I was then selected among the administrative group to manage the Powerful Words test preparation program initiative, and to provide a test preparation strategy that would familiarize students with key vocabulary terms related to reading test items thus improve reading test performance. Specifically, I supervised the reading program and intervention at the test preparation site. My role was to train teachers to use 12 Powerful Words in daily lessons, so that students could learn to recognize, use, and apply the targeted word in class discussions, daily lessons, and practice test passages. In addition, I started the instructional day with a performance of the 12 Powerful Words song on the school's intercom for motivation, spelling the word, reciting the word's definition, and encouraging students to focus on the word throughout the day. I also served as an instructional monitor by checking lesson plans and observing classes on a daily basis. Further, I was available as a resource to teachers for assistance in the instruction of 12 Powerful Words. I had moderate contact with the students only through my role as an Instructional Coach, and not because of this research study. This treatment was a separate enterprise to boost test preparation for students.

Protection of Participants' Rights

The treatment provided to School A group was made a part of the regular classroom instruction in a public school setting taught by qualified teachers, and there was minimal to no risk associated with participation in the study. The district mandated participation in Title 1 supplementary services based on state test score results. In addition, assignment to School A and School B was based on the district's normal pattern of assignment to schools, so there should have been no emotional risk of being assigned to either group. Nothing in the classroom placed students at physical risk as the learning activities in both groups constituted traditional classroom work. Mental risk was negligible because both groups received instruction, including the treatment that would be considered appropriate and necessary.

A parental consent form was not required for participation in the study because student participation in Title 1 supplementary services had already been approved by the school district. Both groups were instructed using methodologies already approved by the building principals, who have autonomy for curricular decisions using district guidelines for time allotment, student: teacher ratios, and teacher qualifications.

To comply with ethical standards and U.S. federal regulations regarding protection of human subjects, I submitted an application to the Walden University Institutional Review Board (IRB) prior to collection of any data.

Data Analysis

To test the null hypothesis that there would be no improvement in reading test scores after participation in a test preparation strategy that familiarizes students with key

vocabulary terms, I used raw scores for the Discovery Education Assessment reading composite in an ANCOVA (Lodico et al., 2010; Mertler & Vannatta, 2005). The independent variable was the type of instruction (test preparation versus no test preparation), the dependent variable was the reading achievement score as measured by Discovery B Education Assessment, and the covariate was initial score in reading as measured by Discovery A Education Assessment.

In contrast to a *t* test or analysis of variance, which do not control for initial differences between pre-existing groups, the analysis of covariance (ANCOVA) was appropriate for this study, because it permits a level of experimental control even in the absence of randomized groups (Lodico et al., 2010; Mertler & Vannatta, 2005). Although the use of pre-existing groups may reduce statistical power, ANCOVA permits variations that lead to applicable analysis (Lodico et al., 2010; Mertler & Vannatta, 2005).

The two schools were matched soundly. The sample size was large enough to show effects of treatment and logically cause educators or researchers to take notice of this auxiliary test preparation program. Section 4 display results in detail.

Section 4: Results

Introduction

The purpose of this quantitative study was to determine if there was an effect of the test preparation programs at two middle schools in the United States, Tennessee, with a 96% Black student population. The research question explored in this study was, Will the Powerful Words test-preparation strategy that familiarizes students with key vocabulary terms related to test items improve reading test performance of underachieving Black students in middle school? Overall, there was a significant effect for Grad 6. However, for Grades 7 and 8 there was marginal to no effect. Dialogue of the results is discussed in the interpretation of findings section. In this section, the results of the descriptive analysis of the data are presented followed by the results of the inferential statistical analysis and ANCOVAs.

Findings

Descriptive Statistics

For Grade 6, posttest scores indicated that after 7 weeks, the test preparation site had a mean score of 33.79 (*SD* = 23.43). The pretest mean score for Grade 6 test preparation site was 26.41 (*SD* = 22.43). The mean had an increase of 7.38. Grade 6 from the no test preparation site had a mean score of 22.47 (*SD* = 18.06). The pretest score for Grade 6 of the no test preparation site was 23.59 (*SD* = 19.50). The mean had a decrease of -1.12.

For Grade 7, posttest scores indicated that after 7 weeks, the test preparation site had a posttest mean score of 32.25 (*SD* = 22.37). The pretest score for Grade 7 of the test

preparation site was 31.05 (*SD* = 25.35). The mean had an increase of 1.20. The no test preparation site had a posttest mean score of 32.61 (*SD* = 20.77). The pretest score for Grade 7 at the no test preparation site was 32.77 (*SD* = 23.97). The mean had a decrease of -0.16.

For Grade 8, posttest scores indicated that after 7 weeks, the test preparation site had a posttest mean score of 31.87 (*SD* = 24.27). The pretest score for Grade 8 of the test preparation site was 38.91 (*SD* = 25.69). The mean had a decrease of -8.04. The no test preparation site had a posttest mean score of 33.73 (*SD* = 24.50). The pretest score for Grade 8 no test preparation site was 34.73 (*SD* = 22.48). The mean has a decrease of -1.00. See Table 4.

Table 4

Pretest and Posttest Mean Scores with Standard Deviations for Treatment and Alternative Treatment Groups plus Results from ANCOVAs of Treatment Groups for Grades 6, 7, and 8

		Pretest (T0)		Posttest (T1)		df	*F*	*p*
	N	M	SD	M	SD			
Treatment								
Female Grade 6	64	27.40	19.91	32.09	22.21			
Male Grade 6	86	29.08	23.49	35.06	24.35			
Students with Disabilities	23	18.96	16.92	27.13	24.17			
Overall Pretest Grade 6	10			26.41	22.43			
Overall Posttest Grade 6	150			33.79	23.43	1, 217	5.687	.018
Female Grade 7	53	36.29	26.42	36.81	23.21			
Male Grade 7	79	27.35	23.93	29.20	21.39			

(*table continues*)

		Pretest (T0)		Posttest (T1)		df	*F*	*p*
	N	M	SD	M	SD			
Students with Disabilities	19	15.70	14.32	28.78	15.95			
Overall Pretest Grade 7	132			31.05	25.35			
Overall Posttest Grade 7	132			32.25	22.37	1, 209	3.529	.062
Female Grade 8	77	44.16	25.85	34.22	22.76			
Male Grade 8	71	32.95	24.29	29.32	25.72			
Student with Disabilities	23	20.33	13.65	10.69	10.42			
Overall Pretest Grade 8	148			39.91	25.69			
Overall Posttest Grade 8	148			31.87	24.27	1, 226	1.352	.246
Non-Treatment								
Female Grade 6	37	24.87	19.54	23.08	18.55			
Male Grade 6	39	21.85	19.57	21.90	17.80			
Students with Disabilities	14	13.18	8.99	15.00	12.50			
Overall Pretest Grade 6	76			23.59	19.50			
Overall Posttest Grade 6	76			22.47	18.06			
Female Grade 7	37	35.10	22.40	34.64	19.10			
Male Grade 7	49	31.77	25.65	31.08	22.02			
Students with Disabilities	10	18.18	18.29	10.80	8.67			
Overall Pretest Grade 7	86			32.77	23.97			
Overall Posttest Grade 7	86			32.61	20.77			
Female Grade 8	42	41.95	23.31	44.48	26.00			
Male Grade 8	46	28.28	19.50	24.15	18.63			
Students with Disabilities	15	19.37	12.89	15.13	14.53			

(*table continues*)

		Pretest (T0)		Posttest (T1)		df	*F*	*p*
	N	M	SD	M	SD			
Overall Pretest Grade 8	88			34.73	22.48			
Overall Posttest Grade 8	88			33.73	24.50			

Note. The pretest scores serve as a covariate, because it is theorized that the posttest treatment is, in part, dependent on students' pretest scores, and there are statically differential results between pre and posttests.

Group Comparisons

For Grade 6, I conducted a three-way ANCOVA, with treatment versus control, student gender, and disability status used as two-level factors. Pretest scores were included as a covariate to determine if there was a statistically significant difference between treatment and control schools on the outcome variable. Discovery B, was used as a control for the pretest score. The main effect of treatment was significant for the 6th grade students, F (1, 217) = 5.687, $p < .018$.

For Grade 7, I conducted a three-way ANCOVA, with treatment versus control, student gender, and disability status used as two-level factors. Pretest scores were included as a covariate to determine if there was a statistically significant difference between treatment and control schools on the outcome variable. Discovery B, was used as a control for the pretest score. The main effect of treatment was marginally significant for the Grade 7 students, F (1, 209) = 3.529, $p > .062$.

For Grade 8, I conducted a three-way ANCOVA, with treatment versus control, student gender, and disability status used as two-level factors. Pretest scores were

included as a covariate to determine if there was a statistically significant difference between treatment and control schools on the outcome variable. Discovery B, was used as a control for the pretest score. There was no main effect, and no significance of treatment for Grade 8 students, F (1, 226) = 1.352, $p > .246$. ANCOVA results are presented in Table 2.

Grade 6 showed that main effect of treatment was significant. Though the effect of treatment had less marginal to no significance in Grades 7 and 8. Although, significant effects can be found, after 7 weeks of treatment, in smaller subgroups of study participants. The overall pretest and posttest mean score improved for both 6th and 7th Grade participants. Outstandingly, students with disabilities mean score in Grades 6 and 7 showed benefit from participating in the auxiliary Powerful Words test preparation program. Based on mean scores from pre and posttests, I believe that if the program was extended past the 7 weeks, students in all grades would show significant main effects of treatment.

Section 5: Discussion, Conclusions, and Recommendations

Introduction

Ninety-six percent of the students in this study were Black, with 15% of students categorized by the (Tennessee State DOE, n.d.) as Black students with disabilities. I used an intact-group comparison, matching the test preparation site with another similar middle school. Next, I added the 12 Powerful Words vocabulary program that familiarizes students with key terms related to test items to improve reading test performance of underachieving Black students in middle school. This research study corresponds to the U.S. national problem in that Black males attending urban city schools are chronicled to be non-proficient, and failure continues as the student progresses to secondary grade levels of education (Yaffe, 2012). After I collected data, I used ANCOVA (Lodico et al., 2010; Mertler & Vannatta, 2005) to compare test preparation and no test preparation site scores on the December 2012 administration of the DEA reading test. I used initial status in reading as a covariate.

Interpretation of Findings

The results of school A, test preparation site showed that posttest scores were significant for Grade 6, marginally significant for Grade 7, and not significant for Grade 8. These findings in this specific study suggest that the Powerful Words (Bell, 2005) were more effective with younger students than with older students within the study participant sample. However, the DEA measuring tool is vertically scaled to show growth in student scores after each benchmark and is based on the Rasch measurement model (Discovery Education, n.d.). Discovery Education, n.d. states that vertically scaled tests use common

items embedded into each testing tool, linking the assessments. However, student retention of acquired learning is less likely with linear, step by step and sequenced teaching (Montgomery & Groat, 1998). Rather, students learn applicably, with synthesis of knowledge in connection to other skills, and retain learned material when taught holistically (Montgomery & Groat, 1998).

According to Discovery Education (n.d.), students attain larger achievement gains in elementary intermediate Grades 3-5. Yet, students gradually decline in performance on achievement tests when reaching middle school secondary level, Grades 6-8. If highly-qualified educators know that achievement rates decline at the start of secondary instruction, it would behoove those that serve at-risk students to equalize potential rates, mentoring to extend learning, start test-wiseness at primary levels, elementary Grades K-2, restructuring test preparation to sequence into curriculum throughout student learning over many years of time (Biggs, Musewe, & Harvey, 2014; McCormick, O'Connor, Cappella, McClowry & Sandee, 2013; Rowley & Wright, 2011).

The Implications for Social Change

Kontovourki and Campis (2010) proposed that test preparation strategies that focus specifically on reading must be planned, explicit, and consistent. Once motivated, educators have purpose or intent for this test taking strategy. It is documented by researchers that test preparation can improve test proficiency rates among student participants (Lam, 2013; Matsumura et al., 2010) Therefore, test-wiseness certainly can improve performance of underachieving Black students on standardized reading tests. So, the first implication of social change in this research study is that the 12 Powerful Words

are purposely effective (Bell, 2005), and consequently should be embedded into core learning programs. The second implication for social change is that test-wiseness should begin within curricula before tested Grade 3, scoped and sequenced to Grade 8 to streamline academic vertical growth scales of test preparation.

In the United States, school test accountability for student achievement under the mandates of NCLB (2002) are attained by a student's proficiency score on standardized tests in core subject areas (DiGaetano, 2015; Fagioli, 2014). If test accountability is not met, according to NCLB guidelines (2002), there must be a school restructuring or closure of Grades K-12 learning institution/s. For Tennessee school districts such as the one that I studied, use of this test preparation strategy could result in the reduction or elimination of schools being taken over and closed by Departments of Education, due to low achievement test scores (see De Witte & Moccia, 2011; Engberg, Gill, Zamarro & Zimmer, 2012; Irwin & Seasons, 2012), and school restructuring and closures are widespread in this state (Tennessee State DOE, n.d).

When the target index performance objective is not met on achievement tests, schools have to provide students with the opportunity to attend a school that met the target index performance for proficiency as well as supplemental intervention services (NCLB, 2002). As an added variable, closures have become a common strategy to expand charter school density in this area of the U.S., but student achievement performance does not necessarily change with the different learning provider (see also, Blackwell, 2013; Brownstein, 2012; Chapman & Donnor, 2015; Farrell, Wohlstetter & Smith, 2012; Gronberg, Jansen & Taylor, 2012; Heise, 2012; Karp, 2014; Prusinski,

Ruddy, Plucker & Cierniak, 2015). Although charter schools, specifically, offer an alternative learning program, these entities have the option to withdraw operations when desired achievement test performance is not met. Consequently, local blight becomes greater. Parents and school personnel get angry, and school communities have to circumnavigate their way (DeWitte & Van Klaveren, 2014; Karp, 2014). Closing schools may perhaps be counterproductive (Grant, Floch Arcello, Konrad, & Swenson, 2014; Lytton, 2011). Hence, supporting the stability of district schools with collaborative efforts of parents and educational teams withstands neighborhood schools for generations (Brevetti, 2014). As of 2015-2016, state representatives have petitioned to the Tennessee Comptroller's department to revoke many charter schools' licensure to operate (Tennessee State DOE, n.d).

With improved test performance of students due to the streamlining of test-wiseness, curricula embedded 12 Powerful Words vocabulary program, and motivated teachers who explicitly train students in this skill-set from the commencement of kindergarten (Ortiz et al., 2012) to Grade 8-12, students and communities will have the ability to thrive in knowing that they are college or career ready, beginning at the early stages of school-age education (Urban Child Institute, 2009). Change occurs in communities when citizens are equipped with learning and have job embedded or higher level learning skills to work, subsequently produce (ASCD, 2016). Students' preparedness is imperative to participate fully in a global society, extending range of the state of Tennessee (ASCD, 2016).

Recommendation for Action

A directed test preparation strategy to improve reading achievement performance of Black students may have implications for educational practice, particularly within the supplementary services of curriculum at Title I schools. Therefore, as a researcher, I will be able to use this study as a communicative tool for stakeholders, which include U.S. national organizations such as the Black Alliance for Education Options (BAEO), educational state and local representatives/ groups as School Seed, Stand, Memphis Child Advocacy Center, and Leadership Memphis, school boards, administrators, teachers, parents, and business owners in the Southeastern United States.

As an educator who has done presentations at the university, national teacher conference, and school levels, I plan to disseminate this research by developing partnerships within professional learning communities with stakeholders. I also plan to provide consultation to educational teams that would like to offer this auxiliary test-preparation strategy that familiarizes Black students with key vocabulary terms related to reading test items and improves reading test performance.

Recommendations for Educational Practice for Action

Findings from this study may also provide some insight regarding the degree to which standardized reading tests measure pure reading achievement or some combination of reading achievement and test-wiseness. Improved test scores not only affect a student's academic and career choices but also have ramifications for teacher placements, school accountability, and school closures (NCLB, 2002). Reducing below proficient scores of Black students could eventually promote a more culture- or race-blind society

in which racial stereotypes about intelligence, ability, and achievement motivation are reduced or eliminated, although such optimal changes may require a generation or more.

References

Allington, R. L. (2013). What really matters when working with struggling readers. *Reading Teachers, 66*(7), 520-530. doi:10.1002/TRTR.1154

Alon, S. (2010). Racial differences in test preparation strategies: A commentary on shadow education, American style: Test preparation, the SAT and college enrollment. *Social Forces, 89*(2), 463-474. doi:10.1353/sof.2010.0053

ASCD. (2016.). *College, career, and citizenship education.* Retrieved from http://www.ascd.org/public-policy/well-rounded-education.aspx

Ashby, C., Burns, J., & Royle, J. (2014). All kids can be readers: The marriage of reading first and inclusive education. *Theory Into Practice, 53*(2), 98-105. doi:10.1080/00405841.2014.885809

Barone, D. (2013). Important revelations about school reform: Looking at and beyond reading first. *School Effectiveness & School Improvement, 24*(4), 392-416. doi:10.1080/09243453.2012.693104

Bayer-Hummel, T. (2009). The effects of jeopardy as a test preparation strategy for nursing students. *Teaching and Learning in Nursing, 5*(1), 12-15. doi:10.1016/j.teln.2009.05.002

Bean, R. M., Dole, J. A., Nelson, K. L., Belcastro, E. G., & Zigmond, N. (2015). The sustainability of a national reading reform initiative in two states. *Reading & Writing Quarterly, 31*(1), 30-55. doi:10.1080/10573569.2013.857947

Bell, L. I. (2005). *12 powerful words that increase test scores and help close the achievement gap. A resource for educations and parents*. Retrieved from https://www.larry-bell.com/product/12-powerful-words/

Bicak, B. (2013). Scale for test preparation and test-taking strategies. *Educational Sciences: Theory & Practice, 13*(1), 279-289. Retrieved from ERIC Database. (ED EJ1016653)

Bicer, A., Capraro, R. M., & Capraro, M. M. (2013). Integrating writing into mathematics classroom to increase students' problem solving skills. *International Online Journal of Educational Sciences, 5*(2), 361-396. Retrieved from http://aggiestem.tamu.edu/sites/aggiestem.tamu.edu/files/drmarymargaretcapraro/Writing%20in%20Mathematics%20Classroom.pdf

Biggs, S. A., Musewe, L. O., & Harvey, J. P. (2014). Mentoring and academic performance of Black and under-resourced urban middle grade students. *Negro Educational Review, 65*(1-4), 64-86. Retrieved from EBSCO Database. (ED 100601726)

Blackwell, W. H. (2013). An era of charter school expansion: an examination of special education in Massachusetts' charter schools. *Journal of Disability Policy Studies, 24*(2), 75-87. doi:10.1177/1044207312455061

Brevetti, M. (2014). Reevaluating narrow accountability in American schools: The need for collaborative effort in improving teaching performances. *Delta Kappa Gamma Bulletin, 81*(1), 32-35. Retrieved from

http://ezp.waldenulibrary.org/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=98474484&site=ehost-live&scope=site

Brownstein, A. (2012). Aid to low-performing schools also a budget patch. *Education Week, 31*(28), 18-19. Retrieved from http://www.edweek.org/ew/articles/2012/04/18/28sig-recession.h31.html

Buckendahl, C., Davis, S., Plake, B., Sireci, S., Hambleton, R., Zenisky, A., & Wells, C. (2009). Evaluation of the national assessment of educational progress. U.S. *Department of Education*. Jessup, MD. Retrieved from https://www2.ed.gov/rschstat/eval/other/naep/naep-complete.pdf

Burchinal, M., McCartney, K., Steinberg, L., Crosnoe, R., Friedman, S. L., McLoyd, V., & Pianta, R. (2011). Examining the Black-White achievement gap among low-income children using the NICHD study of early child care and youth development. *Child Development, 82*(5), 1404-1420. doi:10.1111/j.1467-8624.2011.01620.x

Burdick-Will, J., Keels, M., & Schuble, T. (2013). Closing and opening schools: The association between neighborhood characteristics and the location of new educational opportunities in a large urban district. *Journal of Urban Affairs, 35*(1), 59-80. doi:10.1111/juaf.12004

Cadeau, K. M. (2011). Bridging the gap between middle school and university through partnership: A reflection. *Journal of Multidisciplinary Research, 3*(3), 147-151. Retrieved from Retrieved from EBSCO Database. (ED 71701547)

Chapman, T. K., & Donnor, J. K. (2015). Critical race theory and the proliferation of U.S. charter schools. *Equity & Excellence in Education, 48*(1), 137-157. doi:10.1080/10665684.2015.991670

Chen, T.-L., & Lan, Y.-L. (2013). Using a personal response system as an in-class assessment tool in the teaching of basic college chemistry. *Australasian Journal of Educational Technology, 29*(1), 32-40. doi:10.14742/ajet.95

Chung-Herrera, B. Ehrhart, K.H., Ehrhart, M., Solamon, J. & Kilian, B. (2009). Do test preparation and strategies reduce the black-white performance gap? *Journal of Management, 35*(5), 1207-1227. doi:10.1177/0149206308328506

Coelho, V., Marchante, M., & Sousa, V. (2015). "Positive attitude": A multilevel model analysis of the effectiveness of a social and emotional learning program for Portuguese middle school students. *Journal of Adolescence, 43,* 29-38. doi: 10.1016/j.adolescence.2015.05.009

Common Core State Standards Initiative. (n.d.). *Key shifts in English language arts.* Retrieved from http://www.corestandards.org/other-resources/key-shifts-in-english-language-arts/

Condron, D. J., Tope, D., Steidl, C. R., & Freeman, K. J. (2013). Racial segregation and the Black/White achievement gap, 1992 to 2009. *Sociological Quarterly, 54*(1), 130-157. doi:10.1111/tsq.12010

Creswell, J. W. (2012). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research.* Boston, MA: Pearson Education.

De Witte, K., & Van Klaveren, C. (2014). The influence of closing poor performing primary schools on the educational attainment of students. *Educational Research & Evaluation, 20*(4), 290-307. doi:10.1080/13803611.2014.940979

De Witte, P. M., & Moccia, J. (2011). Surviving a school closing. *Educational Leadership, 68*(8), 54–57. Retrieved from Retrieved from EBSCO Database. (ED 60719870)

Dietz-Uhler, B., & Lanter, J. (2009). Using the four-question technique to enhance learning. *Teaching of Psychology, 36*(1), 38-41. doi:10.1080/ 00986280802529327

DiGaetano, A. (2015). Accountability school reform in comparative perspective. *Urban Affairs Review, 51*(3), 315-357. doi:10.1177/1078087414537606

Discovery Education (n.d.). *Assessment*. Retrieved from http://static.discoveryeducation.com/de/docs/assessment/14111_Tennessee%20Research.pdf

Discovery Education (2008). *Discovery assessment brochure*. Retrieved from http://www.discoveryeducation.com/pdf/DEAssessment_Brochure.pdf

Doe, C., & Fox, J. (2011). Exploring the testing process: Three test takers' observed and reported strategy use over time and testing contexts. *Canadian Modern Language Review, 67*(1), 29-54. Retrieved from Retrieved from EBSCO Database. (ED 58531875)

Ebanks, M., Toldson, I., Richards, S., & Lemmons, B. (2011). Project 2011 and the preparation of Black and Latino students for admission to specialized high

schools in New York City. *Journal of Negro Education, 81*(3), 241-251. Retrieved from Retrieved from EBSCO Database. (ED 83853082)

Engberg, J., Gill, B., Zamarro, G., & Zimmer, R. (2012). Closing schools in a shrinking district: Do student outcomes depend on which schools are closed? *Journal of Urban Economics, 71,* 189-203. doi:10.1016/j.jue.2011.10.001

Enriquez, G. (2013). "But they won't let you read!": A case study of an urban middle school male's response to school reading. *Journal of Education, 193*(1), 35-46. Retrieved from http://www.academia.edu/3651182/_But_They_Wont_Let_You_Read_A_Case_Study_of_an_Urban_Middle_School_Males_Response_to_School_Reading

Erickson, M. (1972). Test sophistication: An important consideration. *Journal of Reading, 16*(1), 140-144. Retrieved from http://www.jstor.org/stable/40010257?seq=1#page_scan_tab_contents

Espey, L., & Brindle, S. (2010). Click, click, wow! Engaging students with student response systems. In D. Gibson & B. Dodge (Eds.), *Proceedings of Society for Information Technology & Teacher Education international conference* (pp. 2708-2714). Chesapeake, VA: Association for the Advancement of Computing in Education. Retrieved from http://www.editlib.org/p/33779

Fagioli, L. P. (2014). A comparison between value-added school estimates and currently used metrics of school accountability in California. *Educational Assessment, Evaluation and Accountability, 26*(2), 203-222. doi:10.1007/s11092-013-9183-9

Farnsworth, T. (2013). Effects of targeted test preparation on scores of two tests of oral English as a second language. *TESOL Quarterly, 47*(1), 148-155. doi:10.1002/tesq.75

Farrell, C., Wohlstetter, P., & Smith, J. (2012). Charter management organizations: An emerging approach to scaling up what works. *Educational Policy, 26*(4), 499-532. doi:10.1177/0895904811417587

Fede, J. L., Pierce, M. E., Matthews, W. J., & Wells, C. S. (2013). The effects of a computer-assisted, schema-based instruction intervention on word problem-solving skills of low-performing fifth grade students. *Journal of Special Education Technology, 28*(1), 9-21. doi:10.1177/016264341302800102

Fike, D., Fike, R., & Lucio, K. (2012). Does clicker technology improve student learning? *Journal of Technology and Teacher Education, 20*(2), 113-126. Retrieved from http://www.learntechlib.org/p/38442

Finkel, E. (2010). *Black children still left behind*. Retrieved from http://www.districtadministration.com/article/black-children-still-left-behind

Gadbury-Amyot, C. C., Austin, K., & Overman, P. R. (2013). Development and implementation of online national board dental examination review courses. *Journal of Dental Education, 77*(12), 1556-1565. Retrieved from https://www.researchgate.net/publication/259249568_Development_and_Implementation_of_Online_National_Board_Dental_Examination_Review_Courses

Gamse, B. C., Jacob, R. T., Horst, M., Boulay, B., & Unlu, F. (2009). *Reading first impact study* [Final report]. Jessup, MD: National Center for Educational

Evaluation and Regional Assistance. Retrieved from http://ies.ed.gov/ncee/pdf/20094038.pdf

Garo, L. A. (2013). Children's exposure to neighborhood poverty and violence: Implications for Black student middle school literacy in Charlotte, North Carolina. *Vulnerable Children & Youth Studies, 8*(1), 60-96. doi:10.1080/17450128.2012.668231

Grant, C. A., Floch Arcello, A., Konrad, A. M., & Swenson, M. C. (2014). Fighting for the 'right to the city': Examining spatial injustice in Chicago public school closings. *British Journal of Sociology of Education, 35*(5), 670-687. doi:10.1080/01425692.2014.919844

Grier-Reed, T., Ehlert, J., & Dade, S. (2011). Profiling the African American student network. *The Learning Assistance Review, 16*(1), 21-30. Retrieved from ERIC Database. (ED EJ919576)

Gronberg, T. J., Jansen, D. W., & Taylor, L. L. (2012). The relative efficiency of charter schools: A cost frontier approach. *Economics of Education Review, 31*, 302-317. doi:10.1016/j.econedurev.2011.07.001

Heise, M. (2012). Law and policy entrepreneurs: Empirical evidence on the expansion of school choice policy. *Notre Dame Law Review, 87(*5), 1917-1940. Retrieved from http://scholarship.law.nd.edu/cgi/viewcontent.cgi?article=1041&context=ndlr

Hoffman, L., & Nottis, K. (2008). Middle schools students' perceptions of effective motivation and preparation factors for high-stakes tests. *National Association of Secondary School Principals, 92(*3), 209-223. doi:10.1177/0192636508322662

Hornof, M. (2008). Reading tests as a genre study. *The Reading Teacher, 62*(1), 69-73. doi:10.1598/RT.62.1.8

Hughes, C. A., & Schumaker, J. B. (1991). Test-taking strategy instruction for adolescents with learning disabilities. *Exceptionality, 2*, 205-221. doi:10.1080/09362839109524784

Hughes, C., Schumaker, J. B., Deshler, D. D., & Mercer, C. (1988). *Learning strategies curriculum: The test-taking strategy.* Lawrence, KS: Edge. doi:10.1177/106342669300100307

Irizarry, Y. (2015). Selling students short: Racial differences in teachers' evaluations of high, average, and low performing students. *Social Science Research, 52*, 522-538. doi:10.1016/j.ssresearch.2015.04.002

Irwin, B., & Seasons, M. (2012). School closure decision-making processes: Problems and prospects. *Canadian Journal of Urban Research, 21*(2), 45-67. Retrieved from http://search.proquest.com/openview/09882020bc32c534170befa7127f9bc0/1?pq-origsite=gscholar

Jiang, L., & Elen, J. (2011). Instructional effectiveness of higher-order questions: The devil is in the detail of students' use of questions. *Learning Environments Research, 14*(3), 279-298. doi:10.1007/s10984-011-9095-x

Johnson, L. (2014). “Can I come in here?” Winston’s discovery of edge-of-school spaces and meaningful literacy engagement. *Changing English: Studies in Culture & Education, 21*(3), 201-214. doi:10.1080/1358684X.2014.929281

Kantor, H. (1991). Education, social reform, and the state: ESEA and federal education policy in the 1960s. *American Journal of Education, 100*(1), 47-83. doi:10.1086/444004

Karp, S. (2014). Tough lessons for Rahm: Education promises to be a central issue in the 2015 mayoral election, especially in black communities that have borne the brunt of school closings, teacher layoffs and charter expansion. *The Chicago Reporter, 43*(3), 7-13. Retrieved from http://chicagoreporter.com/tough-lessons-rahm/

Kawasaki, S. & Tominaga, H. (2010). Execution test series and partial scoring in support for introductory c programming exercise. In J. Herrington & C. Montgomerie (Eds.), *Proceedings of EdMedia: World Conference on Educational Media and Technology* (pp. 3189-3196). Toronto, CA: Association for the Advancement of Computing in Education (AACE). Retrieved from https://www.learntechlib.org/p/35096/

Keels, M. (2013). Getting them enrolled is only half the battle: College success as a function of race or ethnicity, gender, and class. *American Journal of Orthopsychiatry, 83*(2-3), 310-322. doi:10.1111/ajop.12033

Kontovourki, S., & Campis, C. (2010). Meaningful practice: Test prep in a third-grade public school classroom. *Reading Teacher, 64*(4), 236-245. doi:10.1598/RT.64.4.2

Kucan, L., & Palinscar, A. S. (2011). Locating struggling readers in a reconfigured landscape: A conceptual review. *Handbook of Reading Research, 4*, 341-358.

Kulik, C.-L. C., Kulik, J. A., & Bangert-Drowns, R. L. (1990). Effectiveness of mastery learning programs: A meta-analysis. *Review of Educational Research, 60*, 265-269. doi:10.3102/00346543060002265

Lam, R. (2013). Formative use of summative tests: Using test preparation to promote performance and self-regulation. *Asia-Pacific Education Researcher, 22*(1), 69-78. doi:10.1007/s40299-012-0026-0

Lancaster, P. E., Schumaker, J. B., Lancaster, S. C., & Deshler, D. D. (2009). Effects of a computerized program on use of the test-taking strategy by secondary students with disabilities. *Learning Disability Quarterly, 32*(3), 165–179. doi:10.2307/27740366

Lauen, D. L., & Gaddis, S. M. (2012). Shining a light or fumbling in the dark? The effects of NCLB's subgroup-specific accountability on student achievement. *Educational Evaluation and Policy Analysis, 34*(2), 185-208. doi:10.3102/0162373711429989

Lenarduzzi, G. P. (2015). Critical incident effects on principals: Using school closure as the context. *Educational Management Administration & Leadership, 43*(2), 253-268. doi:10.1177/1741143213513190

Lodico, M. G., Spaulding, D. T., & Voegtle, K. H. (2010). *Methods in educational research: Theory to practice.* San Francisco, CA: Jossey-Bass.

Lundy-Wagner, V., & Gasman, M. (2011). When gender issues are not just about women: Reconsidering male students at historically Black colleges and universities. *Teachers College Record, 113*(5), 934-968. Retrieved from http://www.academia.edu/16135607/When_Gender_Issues_Are_Not_Just_About_Women_Reconsidering_Male_Students_at_Historically_Black_Colleges_and_Universities

Lytton, M. (2011). Have all the costs of closing a school been considered? *CELE Exchange. Centre For Effective Learning Environments, 2011*(5-8), 1-4. doi:10.1787/5kg5c8bwqs31-en

Mangold, K. (2007). Educating a new generation: Teaching baby boomer faculty about millennial students. *Nurse Educator, 32*(1), 21-23. Retrieved from www.ncbi.nlm.nih.gov/.../17220763

Matsumura, L.C., Garnier, H.E., & Resnick, L.B. (2010). Implementing literacy coaching: The role of school social resources. *Education Evaluation and Policy Analysis, 32*(2), 249-272. doi:10.3102/0162373710363743

Mayr, S., Erdfelder, E., Buchner, A., & Faul, F. (2007). A short tutorial of GPower. Tutorials. *Quantitative Methods for Psychology, 3,* 51-59. Retrieved from http://www.tqmp.org/RegularArticles/vol03-2/p051/p051.pdf

McCormick, M. P., O'Connor, E. E., Cappella, E., & McClowry, S. G. (2013). Teacher-child relationships and academic achievement: A multilevel propensity score model approach. *Journal of School Psychology, 51,* 611-624. doi:10.1016/j.jsp.2013.05.001

Mertler, C. A., & Vannatta, R. A. (2005). *Advanced and multivariate statistical methods* (3rd ed.). Glendale, CA: Pyrczak.

Millman J., Bishop, H., & Ebel, R. (1965). An analysis of test-wiseness. *Educational and Psychological Measurement, 25*, 707-726. doi:10.1177/001316446502500304

Milner, H. R. (2012, May). But what is urban education? *Urban Education, 47*(3), 556-561. doi:10.1177/0042085912447516.

Misco, T. (2010). Remediation for another high-stakes test. *Kappa Delta Pi Record, 46*(3), 121-126. Retrieved from ERIC Database. (ED EJ898411)

Mizumoto, A., & Takeuchi, O. (2009). Examining the effectiveness of explicit instruction of vocabulary learning strategies with Japanese EFL university students. *Language Teaching Research, 13,* 425-449. Retrieved from http://kuir.jm.kansai-u.ac.jp/dspace/bitstream/10112/4495/1/KU-1100-20091000-01.pdf

Montgomery, S., & Groat, L. (1998). Student learning styles and their implications for teaching. *The CRLT occasional paper series, 10,* 1-8. Retrieved from http://web.uri.edu/teach/files/Student-Learning-Styles-and-Their-Implications-for-Teaching.pdf

Mosley, Wetzel, M., & Rogers, R. (2015). Constructing racial literacy through critical language awareness: A case study of a beginning literacy teacher. *Linguistics and Education*. Advance online publication. doi:10.1016/j.linged.2015.03.014

Nation's Report Card (2013). *Grade 8 national results.* Retrieved from http://nationsreportcard.ed.gov/reading_math_2013/#/executive-summary

National Assessment of Educational Progress (n.d.). *NAEP*. Retrieved from http://nces.ed.gov/nationsreportcard/

National Center for Educational Statistics (2009). *Fast Facts*. Retrieved from https://nces.ed.gov/

National Middle School Association. (2010). *This we believe—Essential attributes and characteristics of successful schools*. Retrieved from http://www.amle.org/AboutAMLE/ThisWeBelieve/The16Characteristics/tabid/1274/default.aspx

No Child Left Behind Act of 2001, Pub. L. No. 107-110, § 115, Stat. 1425 (2002).

Oakland, L., & Weilart, E. (1972). The effects of test-wiseness materials on standardized test performance of preschool disadvantaged children. *Journal of School Psychology, 10,* 355-360. Retrieved from ERIC Database. (ED047050)

Ortiz, M., Folsom, J. S., Al Otaiba, S., Greulich, L., Thomas-Tate, S., & Connor, C. M. (2012). The componential model of reading: predicting first grade reading performance of culturally diverse students from ecological, psychological, and cognitive factors assessed at kindergarten entry. *Journal of Learning Disabilities, 45*(5), 406-417. doi:10.1177/0022219411431242

Pan, Y., & Newfields, T. (2011). Teacher and student washback on test preparation evidenced from Taiwan's English certification exit requirements. *International Journal of Pedagogies & Learning, 6*(3), 260-272. doi:10.5172/ijpl.2011.6.3.260

Pearlman, K. (1984, August). Validity generalization: Methodological and substantive implications for meta-analytic research. In H. Wing (Chair), *Meta-analysis:*

Procedures, practices, and pitfalls. Symposium conducted at the meeting of the American Psychological Association of Toronto, Canada. Retrieved from http://psycnet.apa.org/psycinfo/1987-36191-001

Phillip, A. (2011). Community College of Philadelphia programs provide a support system for black male students. *Diverse Issues in Higher Education, 28*(23), 7. Retrieved from https://www.questia.com/magazine/1G1-276517618/community-college-of-philadelphia-programs-provide

Piazza, S. V., & Duncan, L. E. (2012). After-school literacy engagements with struggling readers. *Reading & Writing Quarterly, 28*(3), 229-254. doi:10.1080/ 10573569.2012.676363

Preston, H. (1964). *Test-wiseness: Is it good or bad?* Upper Saddle River, NJ: Prentice Hall Inc. Retrieved from https://msu.edu/~dwong/StudentWorkArchive/ CEP900F00-RIP/Haniza-TestWiseness.htm

Prusinski, E. L., Ruddy, A., Plucker, J. A., & Cierniak, K. A. (2015). Chartering new waters: The Indianapolis mayoral transition and the charter school community. *Educational Policy, 29*(4), 617-644. doi:10.1177/0895904813510776

Purtell, K., & McLoyd, V. (2013). Parents' participation in a work-based anti-poverty program can enhance their children's future orientation: Understanding pathways of influence. *Journal of Youth & Adolescence, 42*(6), 777-791. doi:10.1007/s10964-012-9802-7

Reardon, S. F., Valentino, R. A., & Shores, K. A. (2012). Patterns of literacy among U.S. students. *The Future of Children, 22*(2), 17-38. Retrieved from http://cepa.stanford.edu/content/patterns-literacy-among-us-students

Rice, S., Geels, K., Trafimow, D., & Hackett, H. (2011). *Our students suffer from both lack of knowledge and consistency: A ppt (potential performance theory) analysis of test-taking*. Retrieved from ERIC Database. (ED529376)

Rowley, R. L., & Wright, D. W. (2011). No "White" child left behind: The academic achievement gap between Black and White students. *The Journal of Negro Education, 80*(2), 93-107. Retrieved from https://www.questia.com/library/journal/1P3-2530927031/no-white-child-left-behind-the-academic-achievement

Ruberg, L. F. (2001, June). Evaluation of program impact based on teacher implementation and student performance. In *Proceedings from World Conference on Educational Multimedia, Hypermedia, and Telecommunications.* Norfolk, VA: Association for the Advancement of Computing in Education. Retrieved from https://www.learntechlib.org/p/8550/

Rudo, Z. (2012). *Corrective action in low-performing schools and school districts.* Retrieved from http://www.sedl.org/pubs/policy91/policy91.pdf

Scruggs, T. E., & Lifson, S. A. (1985). Current conceptions of test-wiseness: Myths and realities. *School Psychology Review, 14*(3), 339-350. Retrieved from http://psycnet.apa.org/psycinfo/1986-05013-001

Stough, L.M. (1993). *Research on multiple-choice questions: Implications for strategy instruction*. Paper presented at the Annual Convention of the Council for Exceptional Children. San Antonio, TX. Retrieved from ERIC Database. (ED366135)

Tennessee Comptroller Department of Education (n.d.). *Tennessee Comprehensive Assessment Program Repository*. Retrieved from http://www.comptroller.tn.gov/repository/re/tcap.pdf

Tennessee State Department of Education (n.d.). *Resources for students and families*. Retrieved from https://www.tn.gov/education/section/student-and-family-resources

Tennessee State Department of Education (n.d.). *State report card data*. Retrieved from www.tn.gov/education/topic/report-card

Tennessee State Department of Education (n.d.). *Tennessee value added system*. Retrieved from http://www.tn.gov/education/topic/tvaas

Tennessee School District (n.d.). *Legacy Memphis city schools*. Retrieved from http://www.mcsk12.net/

Thorndike. E. L. (1951). Reliability. In E. F. Lindquist (Ed.), *Educational measurement* (pp. 560-620). Washington, DC: American Council on Education.

Trafimow, D., & Rice, S. (2008). Potential performance theory: A general theory of task performance applied to morality. *Psychological Review, 115*(2), 447-462. doi.org/10.1037/0033-295X.115.2.447

Turner, S. (2009). Ethical and appropriate high-stakes test preparation in middle school: Five methods that matter. *Middle School Journal, 41*(1). 36-45. Retrieved from ERIC Database. (ED EJ854574)

United States Department of Labor (2012). *The African-American labor force in the recovery.* Retrieved from http://www.dol.gov/_sec/media/reports/BlackLaborForce/BlackLaborForce.pdf

Urban Child Institute (2009). *Updates on data, education and policy.* Retrieved from http://www.urbanchildinstitute.org/sites/all/files/2009-11-Cognitive_brief.pdf

Xie, Q., & Andrews, S. (2013). Do test design and uses influence test preparation? Testing a model of washback with structural equation modeling. *Language Testing, 30*(1), 49-70. doi:10.1177/0265532212442634

Welsh, M. E., Eastwood, M., & D'Agostino, J. V. (2014). Conceptualizing teaching to the test under standards-based reform. *Applied Measurement in Education, 27*(2), 98-114. doi:10.1080/08957347.2014.880439

West, C. (2012). Summer program aims to improve literacy skills of black male teens. *Education Digest, 77*(5), 43-45. Retrieved from ERIC Database. (ED EJ972610)

Winsler, A., Gupta Karkhanis, D., Kim, Y., & Levitt, J. (2013). Being Black, male, and gifted in Miami: Prevalence and predictors of placement in elementary school gifted education programs. *Urban Review, 45*(4), 416-447. doi:10.1007/s11256-013-0259-0

Witt, E. A. (1994). *Meta-analysis and the effects of coaching for aptitude tests.* Paper presented at the Annual Meeting of the American of the American Educational Research Association. Atlanta, GA. Retrieved from ERIC Database. (ED358120)

Wong, K. K., Shi, J., Gao, H., Zheteyeva, Y. A., Lane, K., Copeland, D., & Uzicanin, A. (2014). Why is school closed today? Unplanned K-12 school closures in the United States, 2011–2013. *PLOS ONE, 9*(12), 1-15. doi:10.1371/journal.pone.0113755

Wood, J. L., & Palmer, R. T. (2013). Understanding the personal goals of black male community college students: Facilitating academic and psychosocial development. *Journal of African American Studies, 17*(2), 222-241. doi:10.1007/s12111-013-9248-3

Wood, S., & Jocius, R. (2013). Combating "I hate this stupid book!": Black males and critical literacy. *Reading Teacher, 66*(8), 661-669. doi:10.1002/trtr.1177

Yaffe, D. (2012). Middle school matters: Improving the life course of black males. *Educational Testing Service*. Retrieved from https://www.ets.org/Media/Research/pdf/PICPNV20n4.pdf

Yang, P. (2000). *Effects of test-wiseness upon performance on the test of English as a foreign language* (Unpublished doctoral dissertation). University of Alberta, Edmonton, Canada.

Zhengdong, G. (2009). IELTS preparation course and student IELTS: A case study in Hong Kong. *RELC Journal, 40*(1), 23-40. doi:10.1177/0033688208101449

Appendix A: 12 Powerful Words & Larry Bell's Strategies

Words in This List: 12 Powerful Words (Bell, 2005).

1. **Explain-** to make plain or clear; render understandable or intelligible
2. **Summarize-** giving the short version
3. **Compare-** to consider or describe as similar; liken
4. **Trace-** to follow a course, trail, etc.; list in steps
5. **Formulate-** to devise or develop, as a method, system, etc.; create
6. **Predict-** to declare or tell in advance; prophesy; foretell; what will happen next
7. **Analyze-** to examine carefully and in detail so as to identify causes, key factors, possible results, etc.; break apart
8. **Infer-** to derive by reasoning; come to a conclusion based on evidence; read between the lines
9. **Support-** back up with details
10. **Contrast-** to show unlikeness or differences; note the opposite natures, purposes, etc.
11. **Describe-** to tell or depict in written or spoken words all about something; give an account of
12. **Evaluate-** to judge or determine the significance, worth, or quality of; assess

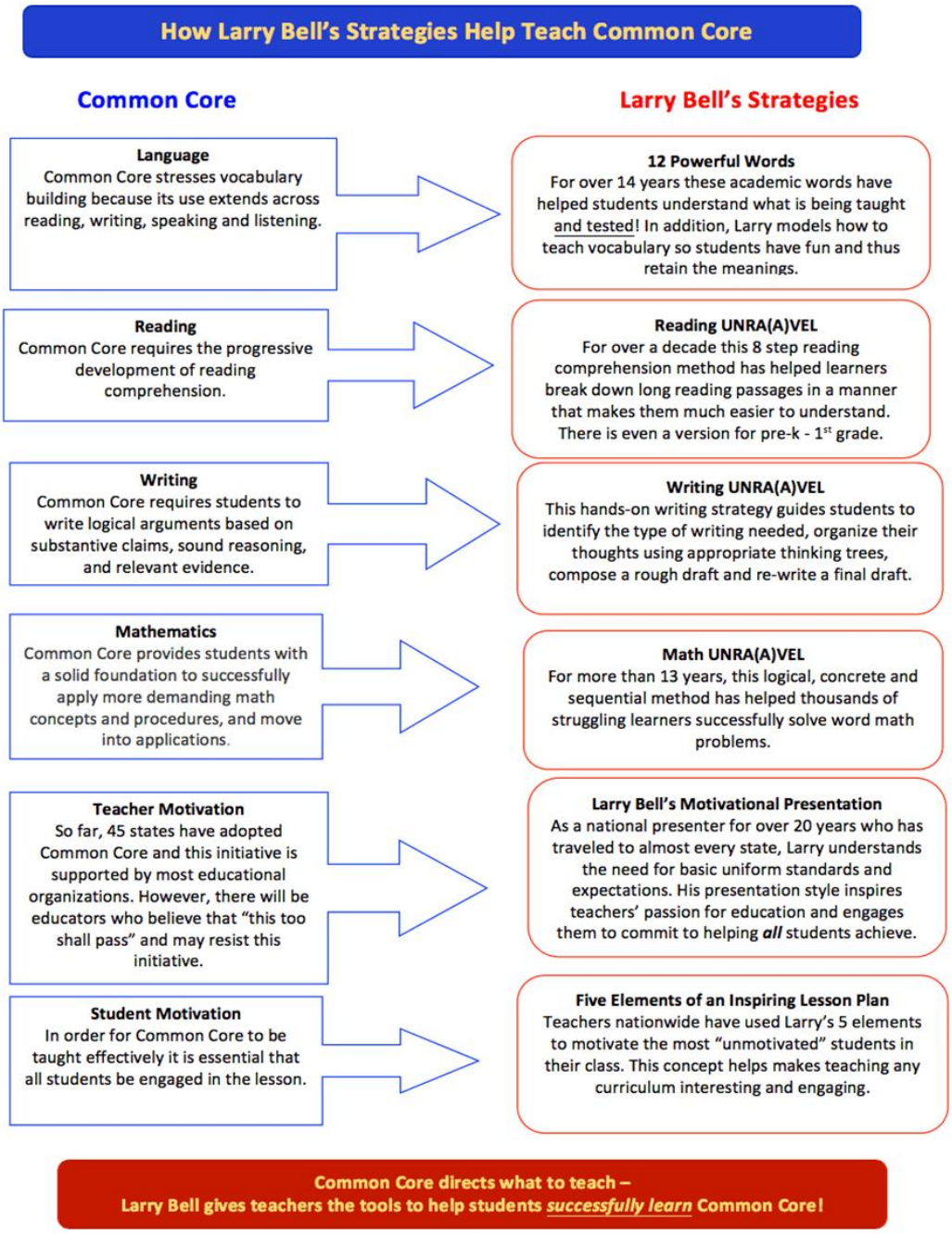

Figure A1. How Larry Bell's strategies help educators in teaching Common Core standards.