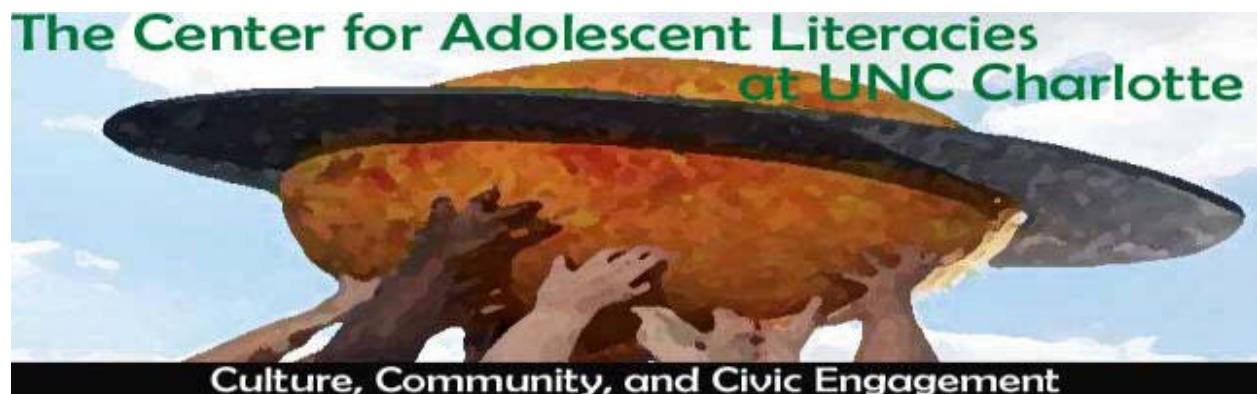# Freedom School Partners
Children's Defense Fund Freedom Schools® Program
Evaluation Report

Submitted
by

Written by

D. Bruce Taylor, Ph.D.

Sandraluz Lara-Cinisomo, Ph.D.

November, 2017

Table of Contents

Tables

Figures

Freedom School Partners Children's Defense Fund Freedom Schools® Program
Summer 2017 Evaluation Report

OVERVIEW

This is the eighth report in a series of evaluation studies conducted by the Center for Adolescent Literacies at UNC Charlotte for the Freedom School Partners' Children's Defense Fund Freedom Schools® programs in Charlotte, N.C. The focus of this report is on reading outcomes for Level I, Level II and Level III Scholars in 10 of the 16 Charlotte Freedom School program sites during the summer 2017.

Freedom School Partners' CDF Freedom Schools Program

The Children's Defense Fund (CDF) is a private, nonprofit child advocacy organization that was founded in 1973 to champion the rights of all children, especially those living in poverty. Based in Washington, DC, CDF grew out of the Civil Rights Movement under the leadership of Marian Wright Edelman, who is president of CDF. The Children's Defense Fund Leave No Child Behind® mission states that it seeks "to ensure every child a Healthy Start, a Head Start, a Fair Start, a Safe Start and a Moral Start in life and successful passage to adulthood with the help of caring families and communities."[1] CDF describes Freedom Schools as a program that "seeks to build strong, literate, and empowered children prepared to make a difference in themselves, their families, communities, nation and world today." In short, Freedom School is a summer program with a mission of empowerment that includes a significant focus on literacy.

Created by the Children's Defense Fund, the Freedom Schools program engages children in grades K-12[2] in a six week summer program designed to prevent the "learning loss" that students (known as Scholars in the program) typically experience over the months when school is not in session. Freedom Schools program also aim to have a positive impact on children's character development, leadership, and community involvement. The CDF Freedom Schools program provides enrichment with the stated goals of "helping children fall in love with reading, increase[ing] their self-esteem, and generate[ing] more positive attitudes toward learning." CDF reports that more than 137,000 children have participated in Freedom School programs since its inception in 1995. In the summer, 2017, there were 12,225 Scholars in Freedom School programs in 89 cities and 27 states including Washington D.C. The Scholars are

---

[1] Information about the Children's Defense Fund and its programs is available at the CDF website: http://www.childrensdefense.org/.
[2] Grade and age ranges vary by Freedom School site. Some programs serve children across the K-12 span while others focus on K-5 or K-8.

grouped by grade levels with Level I Scholars having just completed Kindergarten, first or second grade. Level II Scholars come from grades three through five and Level III Scholars from grades six through eight. There is a Level IV high school program but that group was not included in this study.

The Freedom Schools programs provide a literature based reading program, the Integrated Reading Curriculum or IRC. About 80 books are on the IRC's booklist and these books feature the work of many well-known authors. CDF has developed six weeks of lesson plans for approximately half of the books to help staff and Scholars reflect on the themes *I Can Make a Difference in: My Self, My Family, My Community, My Country, and My World with Hope, Education and Action*. The remaining titles are used to create on-site libraries of books for use during silent sustained reading and read-alouds as well as for research on history and community service projects. Servant Leader Interns are recruited and provided with training that includes how to implement the Integrated Reading Curriculum. The majority of Interns are college-age students.

In Charlotte, *CDF* Freedom Schools are hosted by Freedom School Partners, a 501(c)(3) organization founded in 1999 that is dedicated to serving children and families living in poverty. FSP's mission is to "promotes the long-term success of children by preventing summer learning loss through igniting a passion for reading and inspiring a love of learning." Freedom School Partners began hosting Freedom Schools programs in 2004 at one location serving 100 scholars. In 2017, Freedom Schools served 16 sites and more than 1,200 Scholars. FSP partners with community groups, faith-based organizations, colleges and universities, and corporations, which provide volunteer and financial support.

Freedom School sites in Charlotte range in size from 50 to 100 scholars and operate five days a week, from 8:00 to 3:00 p.m. Transportation is provided. Scholars are served breakfast, lunch and a healthy snack. Freedom School programs are offered at no charge to participating families beyond a $40 per family activity fee, and parents are asked to attend parent meetings and volunteer in the program.

A typical day at a Freedom School follows a pattern. After breakfast, the Scholars and site staff, including the Interns, come together for Harambee, a Kiswahili word for "let's pull together." Harambee is a time of celebration and affirmation akin to a daily pep-rally that includes songs, chants and read aloud of a short book by a community member. Integrated Reading Curriculum, or IRC, follows Harambee. During IRC students go to their classrooms with their Intern for a 2-1/2 hour period of literacy activities built around the reading of culturally-diverse books. The program has a 1:10 Intern to Scholar ratio, and Scholars and Interns read, discuss and engage in activities drawn from the books. Following IRC, Scholars eat lunch and then engage in afternoon enrichment activities. The enrichment activities vary by site but include a

mix of traditional summer activities like swimming and sports but also yoga, field trips to museums and other local sites, cooking and hands-on co-curricular activities that include an academic focus that connect to science, engineering, art and technology.

As has been noted, this research builds on a pilot evaluation study conducted at two Freedom School sites during the summer 2009. The evaluation was extended to additional sites in the Summer 2010, 2011, 2012, 2013, 2015 and 2016. Findings across all evaluation years have remained relatively consistent. Although there has been some variation across these years, between 80% and 90% of Freedom School Scholars grew or maintained in their ability to read as measured by the BRI. Furthermore, important data were gathered in 2010 regarding students' attitudes towards the reading component of Freedom School with the overwhelming majority demonstrating positive attitudes towards the program (as determined in an analysis of the Scholar interviews). The Scholars comments pointed to the engaging nature of the books and activities that are part of the IRC and the role of the college-age interns as positive aspects of the program.

Evaluation History

Since 2009, the Center for Adolescent Literacies at UNC Charlotte has conducted seven evaluation studies of Freedom School Partners' Children's Defense Fund Freedom Schools® programs in Charlotte, N.C. This work has grown from a two-site pilot study in 2009 followed by multi-site studies in each year from 2010 through 2013 and from 2015 to 2016.

Table 1. 2010-2016 Independent Level Evaluation Results from the Basic Reading Inventory

Year	Number of Scholars Pre/Post	Performance on the BRI Independent Reading Measure over time (percent of participating Scholars for the year shown).		
		Declined	Maintained	Improved
2010	132	10.6%	36.6%	50.8%
2011	182	9.3%	37.4%	53.3%
2012	196	8.2%	29.2%	62.6%
2013	167	16.5%	25.9%	57.6%
2015	225	12.0%	32.4%	55.6%
2016	295	14.63	39.8	45.58

This report includes data collected during the summer 2017. Between 2009 and 2016, the research design included pre- and post-testing of children participating in the program (known as Scholars) using the *Basic Reading Inventory* (BRI), 10th Ed. (Johns, 2010) Forms A and B.

Results from these evaluations are reported as two types of scores: Independent reading level (what a child can read and comprehend on their own or independently) and Frustration reading levels (that point at which a child can no longer read and comprehend even with support). Results have varied somewhat between but followed a pattern with the largest percentage of Scholars showing gains in reading as measured by the BRI, another significant percent maintaining in their ability to read and smaller percent showing some sort of decline in reading. Table 1 provides an overview of those results of the BRI Independent reading level scores from 2010 through 2016. Results show that between 45 and 63% of Scholars show gains, 25 and 40% maintain, and between about 8 and 16% show a decline as measures on the BRI at the Independent reading level. Table 2 reports the scores from 2010 through 2016 for the BRI Frustration reading levels.

Table 2. 2010-2016 Frustration Reading Level Evaluation Results from the Basic Reading Inventory

Year	Number of Scholars Pre/Post	Performance on the BRI Frustration Reading Measure over time (percent of participating Scholars for the year shown).		
		Declined	Maintained	Improved
2010	132	9.8%	25.0%	65.2%
2011	182	5.4%	33.9%	60.6%
2012	196	4.6%	35.7%	59.7%
2013	167	9.6%	28.7%	61.7%
2015	225	9.8%	28.9%	61.3%
2016	295	15.3%	30.7%	54.0%

Table 2 shows that approximately 54% to 65% of Scholars showing gains in reading, 25% to 35% maintaining in their ability to read, and 5% to 15% declining in their ability to read as measured by the BRI at the Frustration level (a ceiling level in a child's ability to read).

In 2016, the Center for Adolescent Literacies recommended piloting the use of a different reading inventory, the *Ekwall/Shanker Reading Inventory* or ESRI (Shanker & Cockrum, 2013). The ESRI was used at four Freedom School Partner Freedom School sites in Charlotte and the BRI at 10 sites. The BRI and ESRI are similar in structure in that both use Graded Word Lists, passages and questions and allow us to report Independent and Frustration reading level scores. Our goal in introducing the ESRI was two-fold: 1) to determine if the ESRI would be a more efficient but useful measure of reading to be used (i.e., take less time per administration),

and 2) to provide an alternate assessment tool to compare BRI results against. Table 3, below, shows the ESRI evaluation Independent and Frustration level results from 2016 report.

Table 3. 2016 Evaluation Results using the Ekwall/Shanker Reading Inventory at four sites

Year	Number of Scholars Pre/Post	Performance on ESRI Reading Measures over time (percent of participating Scholars for the year shown).		
		Declined	Maintained	Improved
ESRI IND. Scores	91	11.4%	35.4%	53.2%
ESRI FRU Scores	87	10.3%	28.7%	60.9%

As we wrote in the 2016 report, the results on the BRI varied somewhat from previous years with a somewhat smaller group of Scholars showing increases, and the ESRI results were consistent with previous year's results. In both cases, the results (of BRI and ESRI) were fairly consistent with the pattern of results seen between 2010 and 2015. Additionally, we learned that the ESRI took, on average, 8 to 12 minutes per administration less time than the BRI. The BRI pretest took an average of 26.44 minutes to administer and the posttest took an average of 25.16 minutes. In contrast, the ESRI pretest took an average of 14.51 minutes and the posttest took 17.04, on average. These results addressed our two questions about the ESRI: it is an assessment that seems to provide valid and reliable results, consistent with the BRI, but in less time per administration.

This year's evaluation continues with a pre/post format at 10 Freedom School sites in Charlotte using the ESRI to understand the effects on reading performance of students served by Freedom School Partners during the summer of 2017.

RELATED RESEARCH

Note: This review of related research is updated with each evaluation cycle. We review the research literature and add to this section but retain much of what has been reported in earlier reports.

Freedom Schools programs are six-week, literacy-based summer learning programs designed for children at risk of school failure. The risk factors that children in poverty face include lower academic achievement as measured by grades and on standardized tests, lower graduation rates, and difficulties with reading and literacy. Literacy is a key aspect of school completion. Results from the 2015 National Assessment of Educational Progress (NAEP) indicate that 27% of fourth-grade and 28% of eighth-grade public school students in North Carolina scored below the Basic level in reading. Only 38% of fourth-grade and 30% of eighth-grade students scored at

or above the Proficient level. While these scores are not significantly different from 2013, they continue to raise concerns about the reading ability of school-age children in North Carolina.

Youth from low-income households tend to have lower reading achievement scores than children from middle- and high-income households. Each school year, the reading achievement gap grows and much of the distance accrues during the summer when children are not as inclined to read. A recent study by Hughes-Hassell and Rodge (2007) examined the leisure reading habits of 584 urban adolescents (grades 5 – 8). One of their findings indicated that summer reading was not a "popular" activity for either male or female urban youth. However, it is known that for at-risk children, summer reading is essential to bridge the reading achievement gap (Allington & McGill-Frazen, 2003; Kim, 2004). Schacter (2003) studied the summer reading achievement of 61 first graders in Los Angeles. His study found that an 8-week summer reading "camp" experience had bearing on vocabulary, comprehension, phonics, and oral reading. Thus, for at-risk urban children, a summer program that focuses on reading has the potential to positively influence reading achievement.

Research on the CDF Freedom Schools programs has focused on the historical context of the program (Watson, 2014), ideological contexts (Smith, 2010), leadership aspects of the Freedom School program, the impact on college-age Servant Leader Interns (Jackson, 2009a) and implications for teacher education (Coffey, 2009; Davis, 2010; Jackson, 2009b; Jackson, 2011). An overview of the Freedom School program was published in Teaching Tolerance (Williamson, 2013). Bethea (2012) published results of a study that indicate that involvement in the Freedom School program in Oakland, California had a positive influence on Scholars' racial identity and views toward African/African American culture. Pre- and post-test results also showed an increase on social skills strategies and a future commitment to social action; however, the study showed no statistically significant increase in attitudes toward reading. Howard (2015) examines Freedom Schools as a model for reimagining education for Black children that re-centers learning, literacy and culture in an atmosphere that is free from police presence in schools.

Summer Learning Loss

The 9-month school schedule currently in widespread use has its roots in 19th and 20th Century society in which 85% of Americans were involved in agriculture. It made sense at the time to standardize school schedules and to have children at home during the summer months to help with farming. Today fewer than 3% of Americans are involved in agriculture and research shows that students' learning is impacted negatively by this block of time away from school.

There is a growing body of research about summer learning loss including the publication in the last year of a comprehensive book on the subject, *The Summer Slide: What We Know and Can*

Do about Summer Learning Loss (Alexander, Pitcock, & Boulay, 2016). A review of research (meta-analysis) by Kim and Quinn (2013) on summer reading interventions conducted in the United States and Canada from 1998 to 2011 showed that summer reading interventions that employed teacher-directed literacy lessons had a positive effect on K-8 participants' reading comprehension. The effect of these summer interviews was stronger for children from low-income backgrounds than from a mix of income backgrounds. A recent study by Gershenson and Hayes (2017) on the summer activities of exceptional students, which they define as English language learners and students with an individualized educational plan (IEP), shows that these students are less likely to participate in organized summer programs but show greater gains in reading than other groups of students. Meta-analyses conducted by Cooper et al. (2000 and 1996) integrating studies examining the effects of summer vacation on standardized achievement test scores showed that summer learning loss equaled at least one month of instruction as measured by grade level equivalents on standardized test scores, on average. An analysis of the research of Hayes and Grether (1983) with high- and low-poverty students in 600 New York City schools showed that rich and poor students had seven-month difference in scores at the beginning of second grade but this widened to a difference of two years and seven months by the end of grade six. What made this particularly striking was the research showing little or no difference in these students' achievement when school was in session: They learned at the same pace. As Hayes and Grether noted: "The differential progress made during the four summers between 2nd and 6th grade accounts for upwards of 80 percent of the achievement difference between economically advantaged ... and ... ghetto schools."

Research from the past decade shows that the impact of summer learning loss may be greater than found in earlier studies (Allington & McGill-Franzen, 2003). This deficit is so pronounced that Allington and McGill-Franzen dub summer reading loss as the "smoking gun." Their research has reported that the cumulative effects of summer reading loss can mean that struggling readers entering middle school may lag two years behind peers in their ability to read. Additional research (Alexander, Entwisle, & Olson, 2007) traces the achievement gap between high–socioeconomic and low–socioeconomic 9th grade students to the loss in reading proficiency that occurs over the summer months throughout the elementary grades. Summer learning loss across the elementary school years accounted for more than half the difference in the achievement gap between students from high–socioeconomic and low–socioeconomic families. A study by Kim (2004) published by The Center for Evaluation of the American Academy of Arts and Sciences highlights that low-income and minority students experience greater summer reading loss but suggest that summer reading mitigates this negative impact. A 2014 study by Menard and Wilson suggests that the effect on students with reading disabilities (RD) is greater than on non-RD students while another study (Sandburg Patton & Reschly, 2013) suggests greater impact on younger students.

The issue of summer learning loss is not only debated in scholarly journals. In 2010, *Time Magazine* published a cover story entitled "The Case against Summer" (Von Drehle, 2010) in which it reported:

> The problem of summer vacation, first documented in 1906, compounds year after year. What starts as a hiccup in a 6-year-old's education can be a crisis by the time that child reaches high school. After collecting a century's worth of academic studies, summer-learning expert Harris Cooper, … concluded that, on average, all students lose about a month of progress in math skills each summer, while low-income students slip as many as three months in reading comprehension, compared with middle-income students.

Calls to reorganize school calendars and extend the school year have been suggested as a way to deal with the issue of summer learning loss (Aronson, Zimmerman & Carols, 1998; Dechenes & Malone, 2011; Dessoff, 2011; Jimerson, Woehr, Kaufman & Anderson, 2003; Silva, 2007; WestEd, 2001; Woelfel, 2005). Additional research focuses on policy and funding towards mitigating summer learning loss as a way to address gaps in academic achievement (Leefatt, 2015) while other research suggests parent tutoring during the summer as a means for helping many struggling readers (Mitchell & Begeny, 2014). More recent research indicates that summer programs with a math and literacy component can help students realize gains in their math and reading abilities during the summer months (Graham, McNamara & Van Lankveld, 2011; Smith, 2011-2012). Recent scholarship has included more on the role of summer programs to mitigate summer learning loss (McCombs, et al., 2012) and even "do-at-home" activities (Nikirk, 2012). Research on summer learning loss has recently extended to the post-secondary level with research on summer and between-semester knowledge decay (Dills, Hernandez-Julian, & Rotthoff, 2016).

OBJECTIVES AND RESEARCH QUESTIONS

History

Given the challenges of summer learning loss and literacy attainment and their potential impact on such issues as graduation rates, there is a need for more research on summer programs and their potential to address these issues. A 2005 evaluation of the Kansas City Freedom School Initiative demonstrated a significant improvement in reading abilities for Freedom School Scholars. The pilot evaluation conducted in 2009 by UNC Charlotte was the first effort to evaluate outcomes for participating Scholars in Charlotte. In early 2009, Freedom School Partners approached the University of North Carolina at Charlotte's Institute for Social Capital, Inc. (ISC) to develop an outcomes evaluation for the program. A pilot program evaluation was

conducted at two Freedom School sites for summer 2009. Results from the pilot evaluation were promising. This pilot study showed that of the 51 participants in grades two through five, 57% showed an increase in their reading levels as assessed in the *Basic Reading Inventory*, 9th Ed (Johns, 2005). Twenty-nine percent maintained their reading performance and just under 14% showed some decline. A recommendation that stemmed from the pilot evaluation was the continuation of programmatic evaluation.

In 2010, Freedom School Partners contracted with the Center for Adolescent Literacies at UNC Charlotte to implement an outcome evaluation project to examine the effect of Freedom Schools on children participating at all ten FSP Freedom School sites. The program evaluation sought to assess the extent to which the CDF Freedom Schools program met the following objectives for the K-8 students (Scholars) enrolled:

- To increase children's reading performances
- To maintain or to increase children's reading levels from the end of the school year until the beginning of the proceeding school year
- To increase children's "love" of reading

Present Study – Summer 2017

Following the 2016 evaluation, our team met with the leadership of Freedom School Partners and it was agreed that for 2017, we would use the Ekwall/Shanker Reading Inventory because it yielded results that were consistent with the BRI but required less time. The research questions that guided the evaluation were adjusted accordingly. This evaluation was guided by the following questions:

- *Did Freedom School Scholars show any change in their Independent and Frustration reading levels from pre- to posttest as measured by the Ekwall/Shanker Inventory (ESRI)?*
  - Specifically, what proportion of Freedom School Scholars maintained or improved reading performance over time?
  - Was there a significant difference in number of Freedom School Scholars who maintained or improved compared to those whose reading performance declined over time?
  - Did differences in performance over time differ by Scholar demographic characteristics? If so, how?

Rationale for this Study and the Use of IRIs

The ESRI, like the BRI, is a reading evaluation measure that takes into account different aspects of reading including word knowledge, fluency and comprehension with greatest emphasis on comprehension, or meaning making. The ESRI is a good fit to both the goals of the Freedom School reading program but also to the contexts of that program. The IRC component of Freedom Schools, described earlier, engages students in the reading of culturally diverse books. Scholars and Interns read, discuss and engage in activities related to the books. The focus of this culturally diverse literature-based experience is on comprehension rather than a subskills approach to reading.

Reading inventories like the ESRI and BRI are well suited to reading programs like Freedom Schools. They have compatible forms for pre- and post-test administration to measure change over the relatively short duration of the program. The span the K-8 grade levels, the grade and age range of Scholars in the Level I, II and III classrooms and are practical in terms of cost, time and resources to administer. They allow for fidelity in administration so that multiple evaluators could be trained to assess Scholars using common guidelines for administration and scoring and have a solid research base.

Reading assessments have their roots in the early 20th Century but came of age in the 1940s with the study of skills that comprise comprehension (Davis, 1944; Davis, 1968). Today, reading comprehension assessments are the most common type of published reading test that is available, and the most common reading comprehension assessments involve reading of passages followed by questions about the passage (usually literal recall) and then repeat this process with additional "disconnected" passages (p. 6, ETS, 2012). These traditional approaches to measuring comprehension focus on creating items that consist of lists of content and skills rather than an approach that focuses on what students know and should be able to do (ETS, 2012). Variations on this include asking inferential questions in addition to recall questions. Most reading assessments include what is thought of as the basic skill components of comprehension which include word identification, inferences, strategies, vocabulary and lexical knowledge (Sabatini, O'Reilly & Albro, 2012a and 2012b).

Reading assessments fall into two broad categories: formal and informal. Formal assessments are commonly known as standardized tests or measures and have data which often support conclusions about how a student's reading can be compared to other students his or her age. Formal measures are used to assess overall achievement and to compare a student to others at their age or grade. Scores are often given in percentiles or stanines and many are helpful as diagnostic tools or for measuring change over longer periods of time (year to year in schools,

for example). Informal reading measures are content and performance driven and are often used to inform instructional practices or progress monitoring over short intervals for individual students. Leslie and Caldwell (2009) define informal measures as assessments that "do not interpret scores using comparative or normative data or employ standardized procedures for administration and scoring" (p. 410). Informal measures are often used by classroom teachers and others to gain insight into student performance and to inform instruction. Examples of informal assessments with a focus on comprehension include: questions, retellings, informal reading inventories (i.e., the ESRI), think-alouds, and most assessments that fall under the heading of performance or authentic assessments.

Formal and informal assessments measure comprehension but informal measures of reading are better suited for this research because they measure change over a short duration and typically require less time and fewer resources. Formal assessments are usually more expensive to purchase and may require computer administration and/or scoring. There are numerous reading assessments but many of these focus on a narrow range of grades and ages. For example, there are several early literacy assessments such as the Developmental Reading Assessment (DRA) and the Dynamic Indicators of Basis Early Literacy Skills (DIBELS) that span grades K-2 or 3. Informal Reading Inventories (IRI), however, span a larger range of grades from Kindergarten or first grade through grades eight or nine. Costs for IRIs are relatively low and most come in paper/pencil formats although they do require adult one-on-one administration.

The IRI has a long history as a tool for measuring comprehension and reading ability. Most IRIs include measures of word recognition using leveled lists of words and leveled passages read aloud or silently. Nilsson (2008) states in his review of eight IRIs that these assessments provide information about students' strengths and needs as well as charting reading progress over time. Leslie and Caldwell (2009), authors of the *Qualitative Reading Inventory* (Leslie & Caldwell, 2016), raise the following issues about IRIs: 1) readability formulas used to determine passage levels may not accurately measure difference in difficulty of one passage from another, 2) passage equivalency across forms may vary, and 3) questions used to measure comprehension may work differently with different text types and topics. Research conducted by Applegate, Quinn and Applegate (2002) further suggest that IRIs focus more on text-based recall rather than inferential questions. These issues are not unique to IRIs but also reflect concerns with other forms of reading assessments including standardized assessments. Research by Spector (2005) suggests that IRIs are best suited for low-stakes decisions such as assessing reading levels (which aids in book selection and evaluation) but should not be used for diagnosing reading difficulties. Also, IRIs typically do not offer a fine-grained analysis of growth but, rather, measure reading difference in grade-level increments.

Both formal and informal reading assessments are used for program assessment. However, IRIs were used more frequently for short-term pre/post-test administration while standardized measures were more likely to be administered when repeated measures were not used for purposes of evaluation. The STAR Reading by Renaissance Learning is another test used by some programs for evaluation purposes, but was found unsuitable by our team for several reasons. According to the publisher STAR Reading is "designed for students who can read independently" (Renaissance Learning, Inc., 2010), and some Freedom School Scholars are emergent readers, not yet reading at an independent level. Moreover, STAR is a timed test providing each student a fixed amount of time for reading a passage and then between 45 and 60 seconds to answer questions after which moves to the next question. We felt this also would present problems for emergent and struggling readers. Finally, STAR is a computer-based test, and some Freedom School sites lack access to computers or sufficient numbers of computers to properly administer this type of assessment.

Our analysis of reading assessments, outlined above, led us to the conclusion that IRIs were best suited to the Freedom School evaluation project, and in 2008 we determined that the BRI would be our IRI of choice and based on results from 2016, we adopted the ESRI for the 2017 evaluation.

METHODS

Study Design and Measures

At the beginning of the summer 2017, there were 903 Scholars at the 10 sites where data collection was conducted. The sample was stratified by level, gender, and ethnicity (see Table 2). The evaluation included a pretest-posttest design using only an intervention group (i.e., children who were exposed to the Freedom School Program). This design allows investigators to measure change in reading performance from the start of the program to the end. The results presented in this report are based on children for whom we obtained complete pre- and posttest data.

Table 4. Criteria for Stratification

Criteria			
Level	I (K-2)	II (3-5)	III (6-8)
Gender	Male	Male	Male
	Female	Female	Female
Ethnicity	African-American	African-American	African-American
	Hispanic	Hispanic	Hispanic
	Other	Other	Other

Recruitment Procedures

Participants were recruited for the study through the enrollment process for the Freedom School Program. Parents were informed about the research project and were invited to participate. Parental consent was obtained by the Freedom School staff. Each Scholar was randomly selected for the study based on the stratification criteria described above and was administered a child assent/permission prior to assessing his/her reading performance.

Assessment Instrument

The *Ekwall/Shanker Reading Inventory* (ESRI; Shanker & Cockrum, 2013) is an individually administered reading inventory with multiple measures used to assess reading. For this evaluation, the research team used Form A and Form C for the pretest and posttest, respectively. These are equivalent measures used to assess students' oral reading. Form A and C include a Graded Word List (GWL), Graded Reading Passages, and Oral Reading Comprehension Questions that accompany each passage. The ESRI has a single GWL, the San Diego Quick Assessment (SDQA), that has lists of 10 words each. The single set of ESRI word lists are used for the pre and post administration. The ESRI instructs assessors to start all students on the pre-primer (PP) lists of words and to have the student continue reading until he or she makes three or more errors on any one list. Once a student makes three errors the GWL administration is stopped. The lowest list with three or more errors (where the administration was stopped) is the Frustration level. The Instructional level is the list with two errors, and the high level (list) with one error or less is scored Independent.

The Graded Reading Passages on the ESRI consist of short, leveled passages of text that are read aloud by the Scholar while the assessor documents reading accuracy by noting miscues. The passages on the ERSI go through the 9th grade level, one grade level beyond the BRI. Miscues include words skipped, words inserted, and words said incorrectly. The ESRI has assessors say any unknown words that a student cannot read after a five second pause. Scores

are reported at the Independent, Instructional, and Frustration levels based on scales provided for each passage. Passages are a mix of expository and narrative form with accompanying comprehension questions about details from the text. Scores for the ESRI for each passage are computed using a matrix that includes a dimension for the number of comprehension questions missed and number of word recognition errors. More weight is given to comprehension than word errors. Scores are reported at the Independent, Instructional, and Frustration levels (Shanker & Cockrum, 2013). The ESRI computes the Independent and Frustration levels using the same percentages as the BRI (Table 5).

Table 5. Levels of Reading Assessed with the *Ekwall/Shanker Reading Inventory*

Level	Characteristics
Independent (easy)	Comprehension (90%+) Word Recognition (99%+)
Frustration (too hard)	Comprehension (50%+) Word Recognition (90%+)

Scores on the ESRI are computed for each outcome range from pre-primer to ninth grade. For analysis purposes, Scholars who perform at pre-primer are assigned a score of zero. Scholars who reach a ceiling score of ninth grade at the Independent, Instructional or Frustration level at pre- and posttest are assigned a score of 10 to capture their upper limit. While those Scholars may be able to read beyond 9th grade level, assigning a 10 allows us to capture the Scholar's minimum upper limit.

Data Collection Procedures

Prior to enrollment in this evaluation, demographic information about the child (e.g., date of birth, age, grade, race or ethnicity, and prior Freedom School Program participation) was collected by Freedom School Partners which shared the data with the research team for the purpose of this study. Each year of this evaluation, the assessment team selects a purposeful sample of Scholars from among those for whom consent has been obtained. The sample reflects the demographics of the Freedom School program in Charlotte (race, gender, age, grade level). Our goal was to identify approximately 300 Scholars for the pre-test across the 10 sites to create the sample using the ESRI knowing that this would yield a sufficient number of Scholars with pre and post-test scores to provide a high rate of confidence for our findings.

As in previous years, Scholars were selected to participate in the pretest during the first two weeks of the program (June 19-26, 2017) and Scholars who participated in the pretest and who

were present at the time of the assessment participated in the posttest during the last week of the program (the final day of the program was July 27, 2017). Participants were assigned an identification number for data tracking purposes, to de-identify them to protect their identity, and for data analysis purposes. As described above, the ESRI was used to determine a pre- and posttest Independent (floor) and a Frustration (ceiling) reading score equivalent to a grade level based on the exam's scoring procedures.

Sample

For this evaluation, there were 903 eligible Scholars at the 10 participating sites and 393 were randomly selected, exceeding our goal of 300 Scholars for pre-testing. Figure 1 shows the Scholar test flow. Of those, 89 Scholars were unable to reach the Independent level in the pre-test and one was unable to complete the Frustration level on the pretest. Of those who were not able to achieve independence on the pretest (i.e., unable to receive the lowest level of pre-primer or primer), 68 were in Level I, 19 were in Level II and two were in Level 3. Due to absences, program drop out and scheduling a conflict, 93 Scholars did not complete the independence posttest and 46 were unable to achieve the lowest score (pre-primer). Ninety-four Scholars did not complete the frustration posttest. A total of 225 Scholars had complete data. The report is based on all available data.

Figure 1. 2017 Scholar Test Flow
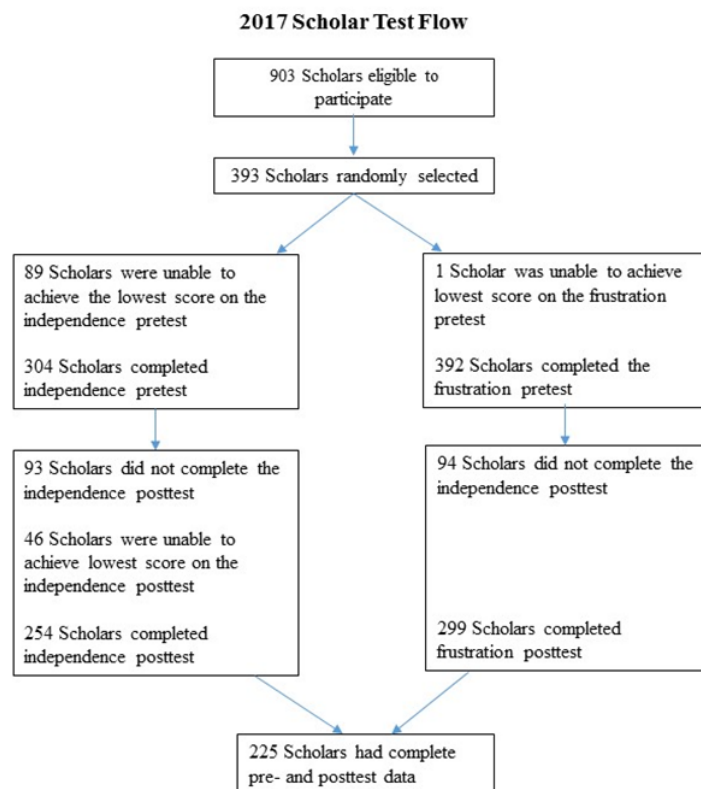
Analysis Plan

Scores on the ESRI are computed for each outcome range from pre-primer to ninth grade. For analysis purposes, Scholars who perform at pre-primer or primer were assigned a score of zero. In previous years, we used -1 to capture these pre-reading levels. However, we found that using zero was just as reliable as -1 in an analyses of the multi-year data used for a peer-reviewed manuscript that is currently under review. Thirty-seven Scholars received a pre-primer score on the independence pretest. Among those, 19 were in Level I, 14 were in Level II and 4 in Level III. Only one Level II Scholar received a primer score on the independence pretest. Twenty-eight Scholars received a pre-primer score on the posttest independence test: 16 were in Level I, 9 in Level II, and 3 in Level III. One Level I Scholar received a primer score on the independence posttest. Fifty-four received a pre-primer score on the pretest frustration measure. Of these 54 with pre-primer score at pre-test, 45 were in Level I and nine were in Level II; none in Level III. Twenty-five Scholars received a pre-primer score on the posttest frustration score: 23 were in Level I and two in Level II; none in Level III. Scholars who reached a ceiling score of ninth grade at the Independent, Instructional or Frustration level at pre- and posttest were assigned a score of 10 to capture their upper limit. While those Scholars might be able to read beyond 10th grade level, assigning a 10 allows us to capture the minimum of their upper limit. Fifteen Scholars received a 10 as the frustration pretest score and 33 received the max score in the posttest. With regard to pretest frustration, three Scholars achieved a score of 10 on the instructional measure, meaning the exceeded the maximum level; only one Scholar was unable to perform well enough to receive a frustration score.

To answer our research question, we computed change scores from pre- to posttest for Independent and Frustration reading performance based on the composite score in both assessments, which captures performance on the Graded Reading Passages miscues and Oral Reading. The results begin with demographic data of the entire sample and those who had complete pre- and posttest data for the Independent and Frustration measure. To determine whether there were differences between Scholars who had Independent pre- and posttest data and those who did not, we conducted bivariate analysis using chi-square test and Fisher's exact test as well as ANOVAs for continuous demographic variables. We also performed these tests to determine if there were differences in the sample who had pre- and posttest Frustration data and those who did not. The subsequent sections provides three sets of results for each outcome (Independent and Frustration). The first set of results shows means and standard deviations for the pre- and posttest by Level. The next set of results provide a distribution that shows the proportion of children whose reading performance declined, was maintained, or improved over time for each test. To determine whether there is a statistically significant difference from pre- to posttest (or within subjects also referred to as a within subject test), we conducted the Wilcoxon Signed-Ranks Test, a non-parametric hypothesis test designed to test

differences in a sample that is not normally distributed and who are assessed using repeated measures as is the case in this study. The Wilcoxon Signed Ranks Test allows investigators to determine whether there is a statistically significant difference in means or groups (declined and improved) among Scholars from pre- to posttest. Also, non-parametric methods allow us to work with data that is ranked such as the use of grades. The third set of results, which we address within the results and the Discussion are based on additional analyses not previously conducted to help us assess associations between demographic characteristics and our outcomes of interest. To explore associations between our outcomes of interest and key Scholar demographic characteristics, we conducted Pearson correlations for continuous variables and analyses of variance (ANOVAs) for categorical predictors (e.g., Scholar level). It is important to note that these additional analyses were conducted to help advance the work and provide insight into possible future directions. An alpha level of .05 was used to determine significance for all tests.

Results

As Table 6, column three shows, more than half the sample was African American/Black (55.5%). Hispanic or Latino children were the second largest group at close to 32%. A substantially smaller proportion of Asian, White and mixed heritage children were also represented in the sample. Females made up 48.5% of the sample. The vast majority of Scholars participated in the free lunch program (94.4%) and more than half had previously attended a FSP summer reading program. The mean age was 10.15 (SD = 2.40) and the largest proportion of Scholars were in Level II. As described above, we explored whether there were differences in the sample who had pre- and posttest data and those who did not. As Table 6 shows, there were significant differences in the Scholar population by race/ethnicity, grade repeated, age and Level between those who had complete data for either the Independent or Frustration measure—differences we explored in our analyses. There were also significant differences by gender and prior FSP experience in the Independent samples.

Table 6. Demographic information for entire sample (N = 393) and those who had complete Independent (n = 225) and Frustration data (n = 299).

|	Complete Independent Data (n = 225)	Complete Frustration Data (n = 299)	All Scholars (N = 393)
	Percent/M(SD)		
Race/Ethnicity			
African American/Black	49.8	50.8	55.5
Hispanic/Latino	36.9	35.1	31.6
Asian/Asian American	4.4	4.3	4.5
European American/White, non-Latino	4.4	4.3	4.1
Mixed Heritage	4.4	4.7	4.1
Other	0	0.7	0.5
χ2	p = .054	0.036	
% Female	53.6	49.0	48.5
Fisher's Exact Test	0.025	0.724	
% Participation in Free/Reduced Lunch	95.1	94.9	94.4
Fisher's Exact Test	1.000	1.000	
% Grade repeated	4.9	6.0	7.6
Fisher's Exact Test	.021	.041	
% Prior FSP experience/enrollment	63.1	55.5	55
Fisher's Exact Test	.000	.722	
Age	10.84	10.13	10.15 (2.40)
ANOVA	F = 11.47, p = .000	F = 2.06, p = .039	
Grade	4.28	3.65	3.55 2.25)
ANOVA	F = 1.10, p = .409	F = 1.24, p = .263	
Level			
I	21.8	34.4	35.6
II	50.2	43.8	43.8
III	28	21.7	20.6
χ2	.000	.533	

Independent Measure Findings

Table 7 shows that the mean grade level among Level I Scholars was close to second grade at 1.65 (SD = 1.32). However, results from the posttest show that the mean grade declined to 1.48 (SD = 1.43). The mean among Level II Scholars was 2.55 (SD = 1.63) for the pretest score and 3.29 (SD = 2.25) on the posttest, an increase of almost a full grade. Level III Scholars showed a substantial improvement from pre- to posttest with more than a full grade improvement from pre-to posttest. Further analyses revealed that there was a significant difference by Scholar Level in mean pre- and posttest scores (F (3, 301) = 23.93, p < .001 and F (2, 251) = 42.42, p < .001), with those in higher Levels earning higher scores.

Table 7. Mean Independent scores by Scholar Level

	Pretest		*Posttest*	
	N	*M (SD)*	*N*	*M (SD)*
Level				
I	72	1.65 (1.32)	65	1.48 (1.43)
II	153	2.55 (1.63)	124	3.29 (2.25)
III	79	3.55 (2.04)	65	5.00 (2.63)
Total	304	2.60 (1.81)	254	3.26 (2.51)

A comparison of performance over time shows that more than half of the sampled population improved. Nearly a quarter were able to maintain their pretest performance. However, 21.8% declined over time. Additional analysis revealed that among those who declined (n = 49 of 225), 30.6% were Level I Scholars, 46.9% were in Level II, and 22.4% were Level III Scholars.

Figure 2. Distribution of Independent Reading Performance Over Time (N= 225)
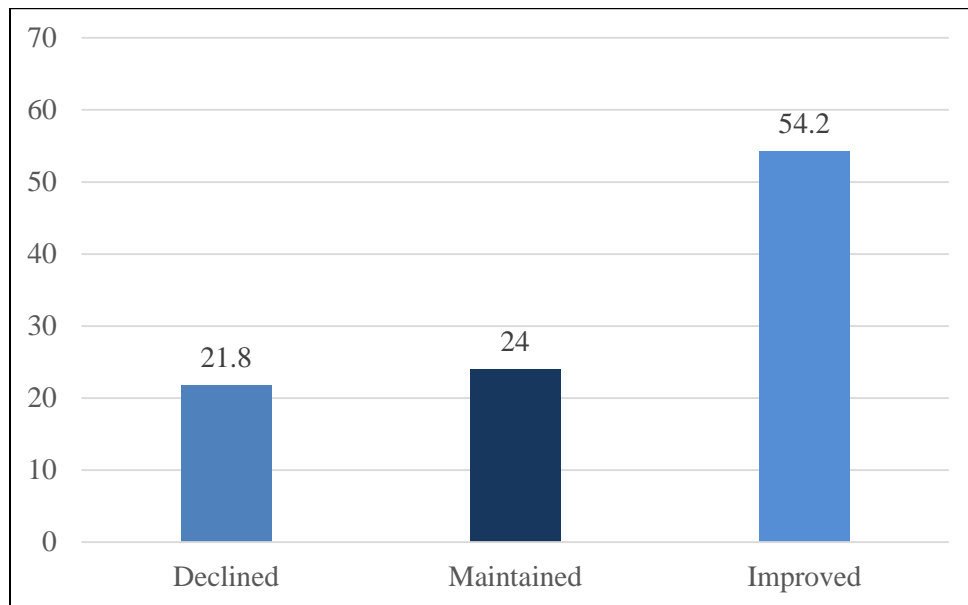
To better understand whether the change from pre- to posttest was statistically significant, we conducted a Wilcoxon Ranks Signed Test. Results revealed that there the differences were statistically significant (Z = -6.70, *p* = .000).

Frustration Measure Findings

Unlike the pattern observed in the previous section, Level I Scholars showed an improvement from pretest (*M* = 2.46, SD = 2.20) to posttest (*M* = 2.63, SD = 2.11) (see Table X below). Scholars at Level II had a mean pretest score of 4.89 (SD = 2.09) and more than a full grade mean difference at posttest (*M* = 6.06, SD = 2.43). The largest change was observed among Level III Scholars whose mean score difference was more than 1.5 grade improvement from pretest (*M* = 6.36, SD = 2.11) to posttest (*M* = 7.88, SD = 2.07). As with the Independence test, results from the ANOVA revealed that there were significant differences in mean pre- and posttest scores by Scholar Level (F (2, 389) = 96.63, *p* < .001 and F (2, 296) = 122.73, *p* < .001, respectively). In other words, Level I Scholar mean scores were significantly different than Level II and III Scholars. Similarly, Level II Scholars' mean pre- and posttest scores were significantly different than Level I and II. We also found a significant difference in pre- and posttest scores by prior FSP experience (F (1, 390) = 23.84, *p* < .001 and F (1, 297) = 23.81, *p* < .001, respectively). Scholars who had previously attended the program scored significantly higher, on average, on the Frustration pre- and posttest (*M* = 4.89, SD = 4.32 and *M* = 6.01, SD = 2.71, respectively) than those who had not (*M* = 3.64, SD = 3.18 and *M* = 4.35, SD = 3.17, respectively). Results

from the correlations revealed a significant and positive association between number of years of prior attendance and mean pre and post Frustration scores (r (213) = .232, p = .001 and r (164) = .313, $p < .001$, respectively).

Table 8. Mean Frustration scores by Scholar Level

	Pretest		Posttest	
	N	M (SD)	N	M (SD)
Level				
1	139	2.46 (2.20)	103	2.63 (2.11)
II	172	4.89 (2.09)	131	6.06 (2.43)
III	81	6.36 (2.11)	65	7.88 (2.07)
Total	392	4.33 (2.20)	299	5.27 (3.03)

An analysis of the proportion of Scholars who declined, maintained or improved in performance shows that more than half the sample improved (52.8%) over time. Another 35.8% were able to maintain their previous performance, and only 11.4% declined on this measure from pre- to posttest. Here, the largest proportion of Scholars who declined on this measure were those in Level II, with 58.8% of the 34 in this category. Level I Scholars made up more than a quarter of the decline group 26.5% (n = 9); Level III was 14.7% (n = 5) of this group.

Figure 3. Distribution of Frustration Reading Performance Over Time (N= 299)
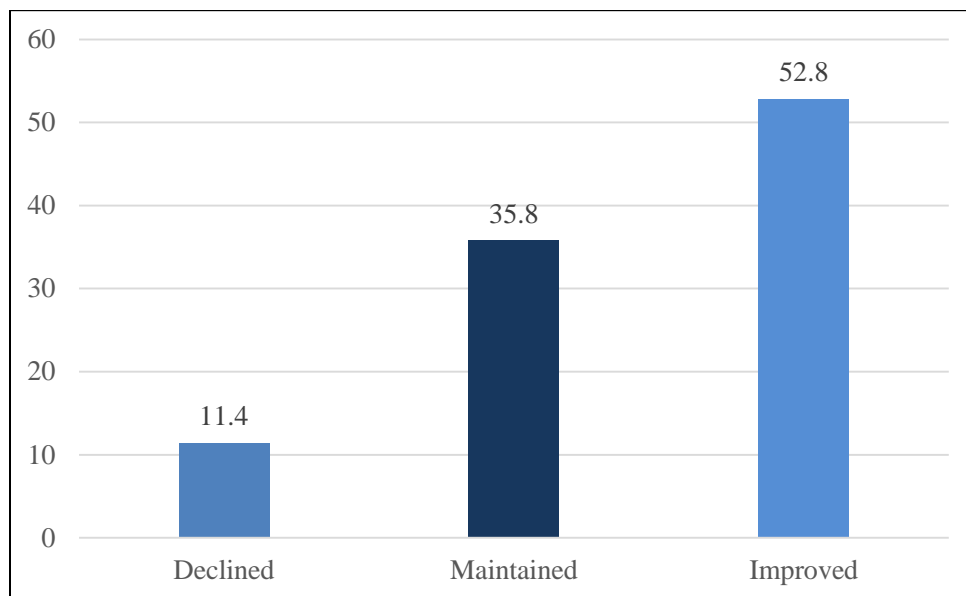
Results from the Wilcoxon Ranks Signed Test indicated that there was a significant difference from pre- to posttest on this measure (Z = -9.09, *p* = .000).

DISCUSSION

This first year of full implementation of the ESRI provided insightful results with regard to Scholar performance and differences by key demographic characteristics. Like prior years, we found that a significant percentage of Scholars either maintained or improved reading performance on both measures over time. However, this year we found that a larger proportion of Scholars declined from pretest to posttest on the Independent measure (21.8% in 2017 versus 11.4% in the 2016, when a subset of sites were administered this test). In contrast, a similar proportion of Scholars declined on the Frustration test (11.4% in 2017 and 10.3% in 2016). Related, we found that a higher number of Scholars did not reach the baseline or minimum reading level necessary to be included in the analysis (see Figure 1). This meant that our final sample size was smaller than expected although it did allow us to test for significant differences. We also examined the distribution by Scholar Level. We found that Level II Scholars made up the largest proportion at 46.9%, followed by Level I (30.6%) and Level III (22.4%) Scholars. Among those whose performance declined on the Frustration test, we found that a similar rate of Level I and Level II Scholars performed less well on the posttest than the pretest (25% and 26.5%, respectively); Level III Scholars made up 14.7% of the group.

In our judgment, these differences are most likely due to differences between the BRI and the ESRI. Other factors could help us understand these differences. For example, changes in the Scholar population or to the curriculum; however, these seem unlikely given the consistency of the program and the population it serves, so we examined the BRI and ESRI more closely. In 2016, the ESRI was piloted in four sites. To determine whether there was a potential site effect, using 2017 data we examined distribution scores by site type (piloted in 2016 versus sites administer the ESRI for the first time this year). The results revealed that there were no meaningful differences. To help elucidate the results and help us explain why a larger number of Scholars did not reach the minimum level on the Independent measure to be included in the final analysis, we revisited the BRI and ESRI scoring procedures. It appears that the ESRI gives greater weight to word recognition errors than the BRI, which has a direct impact on Scholars' final score. Word recognition errors, or miscues, contribute to struggles with reading performance, and are captured to a larger extent in the ESRI. In comparison, the BRI score is based primarily on comprehension. We believe this difference in scoring procedures might help explain the high proportion of Scholars whose scores declined from pre- to posttest. In recent years, Freedom School Partners has integrated the teaching of sight words into its Level I

curriculum. It will be important for FSP to continue to examine reading strategies taught and reinforced during the course of the summer to ensure participants are learning important word recognition skills as well as comprehending the subject matter.

There are other seemingly minor differences between the two assessments that may also contribute to the differences in outcomes. While the ESRI is similar in structure to the BRI— both use graded word lists, leveled passages and comprehension questions—there some notable differences in the number and level of passages included. The BRI has an additional level of passage for emergent readers. The BRI has both a Pre-primer and Primer passage and questions that fall before the grade-level passages (PP, P and then 1, 2, 3…) for emergent readers while the ESRI has only a Pre-primer passage and questions (PP, 1, 2, 3…). At the other end of the reading level spectrum, the BRI stops at grade 8 for passages and questions while the ESRI goes to grade 9 with passages and questions. We find it likely that the BRI offers greater sensitivity than the ESRI for emergent readers and for emergent readers to show gains from Pre-primer to Primer levels. On the other hand, the ESRI provides more room to measure growth at the upper levels because it has a grade 9 passage and questions.

In addition to exploring test characteristics that might explain differences in Scholar performance this year, we conducted additional analysis (e.g., correlations and ANOVAs), which revealed that there were significant differences on the Independence and Frustration measures by Level. We also found that dosage or program exposure was also a significant factor on Frustration scores. Specifically, we found that there was a positive and significant association between years of prior FSP experience and Frustration scores. We also found that there was a significant difference by any prior exposure to the program, with those who attended receiving higher mean Frustration scores. Findings from previous evaluations have shown that as Scholars move up levels from Level I to Level II and III, they show greater growth in reading as measured in these evaluations. We believe that because the Freedom School program focuses on reading and discussion of high-interest books, that there is likely a greater impact on reading comprehension over reading sub-skills (decoding, for example). Our own observations of Freedom School classrooms confirm that great emphasis is placed on meaning making (comprehension). Many of the Level I Scholars are emergent or less experienced readers who are still working out reading skills (decoding) and learning sight words and whose ability to comprehend and show gains in comprehend are less developed than older readers.

Also new this year is an analysis of program experience. Our findings suggest that Scholars with multiple exposures to the program show greater gains than first-year students in the program. The benefits noted in this evaluation to Scholars with prior Freedom School experience merit further study but are a promising outcome. The Freedom School program (treatment, in

evaluation terms), like most summer programs, has a limited timeframe or dosage, so it is reasonable that multiple "doses" of the program obtained by returning over multiple summers could have a cumulative impact. However, it may be that multiple factors influence this dosage effect. The increased exposure to reading over multiple and, often, consecutive summers could help to offset summer learning loss and accelerate Scholars reading gains, but it may also be the case that affective influences such as motivation to read contribute.

Finally, we began the report with a summary of the demographic characteristics of the sample. As Table 6 showed, the sample was very diverse. Given the high number of Scholars who did not have complete pre- and posttest data due to poor performance on a measure or absences, we conducted additional analysis to determine whether there were significant differences by demographic characteristics. It will be important to increase retention in future evaluations to ensure the final sample represents the randomly selected sample as a whole.

We believe that this year's Freedom School evaluation furthers our understanding of the impact the program has on the reading of Scholars who participate. The additional analyses provide us with a deeper understand of the program's impact on particular groups of youth served in the program. We continue to learn about the benefits and limitations of the reading evaluation tools used to measure program impact.

THE CENTER FOR ADOLESCENT LITERACIES AT UNC CHARLOTTE

The Center for Adolescent Literacies at UNC Charlotte is an instructional center focused on developing instruction to make literacy and learning relevant and effective for adolescents and those who work with them. The Center also will conduct and support research and service in support of its primary mission.

The mission of the Center for Adolescent Literacies (CAL) at UNC Charlotte is to advance the literacy achievement of adolescents in urban school settings and to develop pedagogies for adolescents and those who work with them to prepare them to be productive and empowered 21st century citizens. Specifically, the objectives of our center are as follows:

- To provide community outreach
- To build cultural understanding and awareness
- To promote community engagements
- To encourage civic engagement through service learning
- To equip teachers, parents and pre-service teachers with knowledge, skills, and dispositions for supporting and scaffolding adolescent literacy and service learning
- To develop and provide collaborative professional development to promote adolescent literacy
- To encourage collaborative involvement among all stakeholders (including teachers, students, parents/guardians and university faculty).

Evaluation Leadership Team

Dr. Bruce Taylor is the Director of the Center for Adolescent Literacies at UNC Charlotte and is a Professor in the Department of Reading & Elementary Education. Over the past 14 years, Dr. Taylor has provided leadership in developing the ReadWriteServe (RWS) community-based literacy initiatives at UNC Charlotte. These programs include America Reads, the Urban Youth in Schools Internship, and RWS Tutor Training. He is the author and co-author of numerous peer-reviewed articles, book chapters, and technical reports and co-author of three books. His research examines the social and cultural aspects of literacy and learning of adolescents and, in particular, ways to meet the academic learning needs of diverse and marginalized students. He has led several reading program evaluation projects. Dr. Taylor teaches undergraduate, master's level, and doctoral courses that focus on content-area and adolescent literacy, digital literacies in education, and sociocultural aspects of language and literacy.

Dr. Sandraluz Lara-Cinisomo is an Assistant Professor at the University of Illinois at Urbana-Champaign (UIUC) in the Department of Kinesiology and Community Health. Dr. Lara-Cinisomo's research explores the association between biomarkers, psychosocial factors and

perinatal depression. Dr. Lara-Cinisomo's research focuses on English and Spanish speaking Latina mothers. Prior to joining UIUC, Dr. Lara-Cinisomo was an assistant professor an NIH-funded fellow at the University of North Carolina at Chapel Hill, an assistant professor at University of North Carolina at Charlotte and a behavioral scientist at the RAND Corporation. Her research includes qualitative and quantitative methods.

References

Alexander, K., Entwisle, D., & Olson, L. (2007). Lasting consequences of the summer learning gap. *American Sociological Review, 72* (2), 167 – 180.

Alexander, K., Pitcock, S., & Boulay, M.C. (2016). *The summer slide: What we know and can do about summer learning loss*. New York: Teachers College Press.

Allington, R. L., & McGill-Franzen, A. (2003). The impact of summer set-back on the reading achievement. *Phi Delta Kappan, 85*(1), 68-75.

Applegate, M. D., Quinn, K. B., & Applegate, A. J. (2002). Levels of thinking required by comprehension questions in informal reading inventories. *The Reading Teacher, (56)*2, 174-180.

Aronson, J., Simmerman, J. & Carols, L. (1998). Improving student achievement by extending school: Is it just a matter of time? WestEd. San Francisco: Far West Laboratory.

Bethea, S. L. (2012). The impact of Oakland Freedom School's summer youth program on the psychological development of African American youth. *Journal of Black Psychology 28*(4), 442-454.

Coffey, H. (2010). "'They' taught 'me'": The benefits of early community-based field experiences in teacher education. *Teaching and Teacher Education: An International Journal of Research and Studies, 26* (2), 335-342.

Cooper, H., Charlton, K., Valentine, J. C., & Muhlenbruck, L. (2000). Making the most of summer school: A meta-analytic and narrative review. *Monographs of the Society for Research in Child Development, 65*(1), 1-118. EJ 630 022.

Cooper, H., Nye, B., Charlton, K., Lindsay, J., & Greathouse, S. (1996). The effects of summer vacation on achievement test scores: A narrative and meta-analytic review. *Review of Educational Research*, *66*(3), 227-268. EJ 596 384.

Davis, F. B. (1944). Fundamental factors of comprehension in reading. *Psychometrika, 9,* 185-197.

Davis, F. B. (1968). Research in comprehension in reading. *Reading Research Quarterly, 3,* 499-545.

Davis, J. H. (2010). Guide our feet: Teacher education and servant-leadership in a Children's Defense Fund Freedom School (Doctoral Dissertation). *ProQuest LLC*.

Dechenes, S., & Malone, H. J., (2011). Year-round learning: Linking school, afterschool, and summer learning to support student success. *Harvard Family Research Project* (ED521123).

Dessoff, A. (2011). Is year-round schooling on track? *District Administration, 47*(7), 34-36.

Dills, A., Hernández-Julián, R., & Rotthoff, K. W. (2016). Knowledge decay between semesters. *Economics of Education Review*, 5063-74.

Educational Testing Service. (2012). ETS research spotlight 7: Innovation in reading assessment. ETS: Princeton, New Jersey.

Gershenson, S., & Hayes, M. S. (2017). The summer learning of exceptional students. *American Journal of Education*, *123*(3), 447-473.

Graham, A., McNamara, J. K., & Van Lankveld, J. (2011). Closing the summer learning gap for vulnerable learners: an exploratory study of a summer literacy programme for kindergarten children at-risk for reading difficulties. *Early Child Development and Care, 181*95), 575-585.

Hayes, D. P., & Grether, J. (1983). The school year and vacations: When do students learn? *Cornell Journal of Social Relations, 17*(1), 56-71.

Howard, T. C. (2015). Why Black Lives (and minds) matter: Race, Freedom Schools & the quest for educational equity. *Journal of Negro Education, 85*(2), 101-113.

Hughes-Hassell, S., & Rodge, P. (2007). The leisure reading habits of urban adolescents. *Journal of Adolescent & Adult Literacy, 51*(1), 22-33.

Jackson, T. O. (2011). Developing sociopolitical consciousness at Freedom Schools: Implications for culturally responsive teacher preparation. *Teaching Education, 22* (3), 277-290.

Jackson, T. O. (2009a). Toward collective work and responsibility: Sources of support within a Freedom School teacher community. *Teaching and Teacher Education, 25*, 1141-1149.

Jackson, T. O. (2009b). Making the readings come to life": Expanding notions of language arts at Freedom School. *The New Educator, 5*(3), 311-328.

Jimerson, S.R., Woehr, S.M., Kaufman, A.M. & Anderson, G.E. (2003). Grade retention and promotion: Information and strategies for educators. National Association of School Psychologists, S3-61 – S3-64.

Johns, J. L. (2005). *Basic Reading Inventory: Pre-Primer through Grade Twelve and Early Literacy Assessments, 9th Ed*. Dubuque, IA: Kendall Hunt.

Johns, J. L. (2010). *Basic Reading Inventory: Pre-Primer through Grade Twelve and Early Literacy Assessments, 10th Ed*. Dubuque, IA: Kendall Hunt.

Kim, J. (2004). Summer reading & the ethnic achievement gap. *Journal of Education for Students Placed at Risk (JESPAR), 9*(2), 169-188.

Kim, J. S., & Quinn, D. M. (2013). The Effects of summer reading on low-income children's literacy achievement from Kindergarten to grade 8: A meta-analysis of classroom and home interventions. *Review of Educational Research, 83*(3), 386-431.

Leefatt, S. (2015). The key to equality: Why we must prioritize summer learning to narrow the socioeconomic achievement gap. *Brigham Young University Education & Law Journal, (2)*, 549-584.

Leslie, L. & Caldwell, J. (2009). Formal and informal measures of reading comprehension. In S. E. Israel & G. G. Duffy, (Eds.), *Handbook of research on reading comprehension*. New York: Routledge, pp. 403-427.

Leslie, L. & Caldwell, J. (2016). *Qualitative reading inventory-6, 6th Ed.* New York: Pearson.

McCombs, J. S., Augustine, C., Schwartz, H., Bodilly, S., McInnis, B., Lichter, D., & Cross, A. B. (2012). Making summer count: How summer programs can boost children's learning. *Education Digest, 77*(6), 47-52.

Menard, J., & Wilson, A. M. (2014). Summer Learning Loss among Elementary School Children with Reading Disabilities. *Exceptionality Education International*, *23*(1), 72–85.

Mitchell, C., & Begeny, J. C. (2014). Improving Student Reading Through Parents' Implementation of a Structured Reading Program. *School Psychology Review, 43*(1), 41-58.

National Assessment of Education Progress. (2015). U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics.

Nilsson, N. L. (2008). A critical analysis of eight informal reading inventories. *The Reading Teacher, 61*(1), 526-536.

Nikirk, M. (2012). Stave off summer learning loss with do-at-home activities. *Tech Directions, 71*(10), 16-17.

Renaissance Learning, Inc. (2010). *Getting the most out of STAR ReadingTM: Using data to inform instruction and intervention.* Wisconsin Rapids, WI: Renaissance Learning.

Sabatini, J. P., O'Reilly, T. & Albro, E. R. (2012a). *Measuring up: Advances in how we assess reading ability* (Volume I). Lanham, MD: Rowman & Littlefield.

Sabatini, J. P., O'Reilly, T. & Albro, E. R. (2012b). *Reaching an understanding: Innovations in how we view reading assessment* (Volume II). Lanham, MD: Rowman & Littlefield.

Sandberg Patton, K. L., & Reschly, A. L. (2013). Using curriculum-based measurements to examine summer learning loss. *Psychology in Schools, 50*(7), 738-753.

Schacter, J. (2003). Preventing summer reading declines in children who are disadvantaged. *Journal of Early Intervention, 26(*1), 47-58.

Shanker, J. L., & Cockrum, W. (2013). *Ekwall/Shanker reading inventory, 6th Ed.* New York: Pearson.

Silva, E. (2007). On the clock: Rethinking the way schools use time. Washington, D.C.: Education Sector.

Smith, K. (2010). Fostering regimes of truth: Understanding and reflecting on the Freedom School way. *Pedagogy, Culture and Society, 19*(2), 191-209.

Smith, L. (2011-2012). Slowing the summer slide. *Educational Leadership, 69*(4), 60-63.

Spector, J. E. (2005). How reliable are informal reading inventories? *Psychology in the Schools, 42*(6), 593-603.

Von Drehle, D. (2010, July 22). The case against summer. *Time Magazine*. Available at http://www.time.com/time/nation/article/0,8599,2005654,00.html .

Watson, M. (2014). Freedom Schools then and now: A transformative approach to learning. *Journal for Critical Education Policy Studies, 14*(1), 170-190.

WestEd (2001). *Making time count*. San Francisco: Far West Laboratory.

Williamson, L. A. (2013). No school like Freedom School. *Teaching Tolerance 52*(43, 25-28.

Woelfel, K. (2005). Learning takes time for at-risk learners. *Principal*, 85 (Nov./Dec.), 18 -21.