A Quantitative Investigation of the Relationship Between Seventh Grade Reading Scores and High School Office Discipline Referrals

by

Dan Palombit

Dissertation Proposal Submitted to the Doctoral Program

of the American College of Education

in partial fulfillment of the requirements for the degree of

DOCTOR OF EDUCATION

December 2019

A Quantitative Investigation of the Relationship Between Seventh Grade Reading Scores and High School Office Discipline Referrals

by

Dan Palombit

Approved by:

Dissertation Chair: Katrina Schultz, Ed.D.

Committee Member: Patricia Krumnow, Ed.D.

Program Director: Elizabeth Johnson, Ed.D.

Assistant Provost: Jerry Ausburn, Ed.D.

A Quantitative Investigation of the Relationship Between Seventh Grade Reading Scores and High School Office Discipline Referrals

By

Dan Palombit

Approval

The candidate for graduation named above submitted the above applied dissertation by direction of the dissertation committee listed below. The signatures and dates below attest that it was approved and submitted to American College of Education in partial fulfillment of the requirements for the Doctor of Education degree.

Dr. Katrina Schultz — 11/7/2019

Katrina Schultz, EdD
Committee Chair — Date

Dr. Patricia Krumnow, Ed.D — 11/12/2019

Patricia Krumnow, Ed.D.
Committee Member — Date

December 4, 2019

Elizabeth Johnson, Ed.D.
Program Chair — Date

Jerry Ausburn — December 4, 2019

Jerry Ausburn, Ed.D.
Assistant Provost, Education Professions — Date

Abstract

Students' time in class and exposure to academics are essential to academic success. Disruptive student behavior can result in time out of class and missed academic instruction, which can lead to grade retention. Literature exists on reading ability, behavior, and grade retention all independent of each other but little exists on relationships between reading scores, office discipline referrals, and grade retention occurrences. With a focus on Gesell's theory of neuromaturation, the study investigated data to support brain development theories and effects on executive functioning. The quantitative study was designed to determine whether a positive correlation existed between reading scores on standardized tests, office discipline referrals and grade retention occurrence. The sample consisted of a random, anonymous 100 student data covering a three-school year period: 2012-2013, 2014-2015, and 2015-2016. Data analyses were performed using IMB SPSS and G*Power and included Spearman correlation and Kruskal-Wallis *H* test. Spearman correlation analysis found statistical correlational significance between reading scores and office discipline referrals; as reading scores improved, office discipline referrals declined. Results from Kruskal-Wallis *H* test showed statistical significance ($H(2) = 10.42$, $p = .005$) among office discipline referrals as a function of reading scores meaning students' high school office discipline referrals are correlated with seventh-grade reading scores and category. Since statistical analysis was not performed between reading scores and grade retention due to lack of variability on grade retention data ($n = 4$), a descriptive profile was generated for all grade retained students. The study's findings showed relationships between higher standardized test reading scores and lower office discipline referrals and grade retention which can be used by educators when reviewing data, implementing interventions, and developing school- or system-wide improvement plans.

Dedication

This dissertation is dedicated to my magnificent family who provided encouragement and support throughout this process. This journey required an incalculable amount of time and I thank you for being at my side, offering seemingly endless support and encouragement. I can never reimburse to you the time I was working on this dissertation, however, I can promise and declare the rest of my life to you.

To my beautiful, brilliant wife, Megan; you are my rock. My inspiration. My source of happiness. My teammate on this breathtaking journey of life. Quitting or slowing down was never an option thanks to you. Our relationship has grown to one of true love and I know we will continue to grow even closer. I can't thank you enough for being you.

To my kind-spirited stepson, Caiden; you are my source of relief during high-stress times. Because of your kind and free-spirted personality and our relationship, I truly know what it means to be connected to someone. You remind me each and every day of the reason I chose a career in education. I learn so much from you and am a better person because of you. You are truly a remarkable individual and have had a tremendous impact on my life.

To babies June; I cannot wait to meet you. This amazing life gets better each day.

To Mom, Dad, and Katie; thank you for being a fundamental part of this remarkable life. Your never-ending encouragement, faith in me, and love means the world. Somehow, somewhere, and some way I will show you how truly thankful I am to each of you. Your continual words of encouragement and genuine love have kept me going from day one.

To Mr. Trigger and Mr. Diesel; you taught me what loyalty is and how to relax. I undeniably loved and love our time together.

I did this dissertation with each and every one of you. For you. For us.

Acknowledgements

I would like to take this time to offer my thanks to the many individuals who supported me throughout this remarkable journey. I would like to thank my family, friends, and colleagues who continually offered support through this process. Your words of encouragement and humor served as fuel along my journey.

I am thankful to my committee for your patience, supportive feedback, and timely responses. I learned a tremendous amount through our conversations, emails, and edits. Through these communications and interactions, I have refined many skills, namely resilience. I owe special acknowledgement to my chairperson, Dr. Katrina Schultz. Dr. Schultz started as my third dissertation chairperson at a crucial point in my journey. You provided structure, encouragement, and helped me develop a vision and roadmap for my dissertation.

Table of Contents

List of Tables

Table

List of Figures

Figure

Chapter 1: Introduction

A goal of every school is to enable academic success resulting in students' graduation and preparation for post-secondary opportunities by providing excellent educational opportunities to students (Mazyck, 2013). Perhaps the most urgent and important requirement for achieving this goal is positively related to students' presence in class (Özkanal & Arıkan, 2011). Most studies prove consistent attendance is positively related to a student's academic performance (Herman, 2011). At the same time, studies show teachers' frequent requests for support, typically resulting from a student concern, encompasses disruptive student behavior, often resulting in student removal from class for a period of time (Oliver, Wehby, & Reschly, 2011). Positive and non-disruptive behavior affects school and classroom climates as well as allows instructional staff opportunities to focus on curriculum providing students a variety of educational opportunities (Jimerson, Woehr, & Kaufman, 2007). McIntosh, Horner, Chard, Boland, and Good (2006) indicated both academic and behavioral challenges among students are closely linked and suggested educators need skills and tools to effectively reach and connect with students who enter a classroom and display either or both challenges.

Chapter 1 presents an overall introduction to this quantitative correlational research study including background, problem statement, purpose, and significance. Research questions, hypotheses, and theoretical framework are detailed in Chapter 1. Definitions of key terms, assumptions, scope and delimitation, as well as limitations are discussed. Chapter 1 concludes with a summary of important details and transition to Chapter 2.

Background of the Study

Student attendance and student discipline are two of many factors affecting overall student success in school (Mundy et al., 2017). Many existing studies address student reading

levels and abilities, student discipline, student impulse control, and effects of grade retention for students who do not achieve a passing score on a standardized assessment. Student grade retention is used throughout schools systems as a common result for lack of progression or failing grades for particular classes, as in the case in high schools, or grade levels, as is the case in elementary schools (Kenneady, 2004). Grade retention generally reveals unfavorable academic and social emotional results for individuals subject to repeating a course or grade level (Kenneady, 2004; Martorell & Mariano, 2018). A gap in literature as related to methods used to estimate or predict a likelihood of a student being at an identified behavioral risk level based on the student's reading ability is identified in Chapter 1.

Supported by Henderikus' (2010) three theory-related requirements, theory of neuromaturation was identified for this study. According to the theory of neuromaturation, frontal lobes of the brain develop for at least the three first decades of life (Gesell, 1925). The main function of the frontal lobes is to provide executive control, such as planning, impulse control, memory, and overall coordination of these and other cognitive characteristics into a rational self (Meltzer et al., 2017). Executive control characteristics are of concern specifically with an early adolescent individual experiencing independence and unstructured environments commonly seen in middle grades (Friedman & Miyake, 2017; Kalpakci, Ha, & Sharp, 2018; Nitschke, Köstering, Finkel, Weiller, & Kaller, 2017). Because this theory of function includes the idea of affordance, or what fits between the student and his or her environment, students without proper environmental affordance may suffer consequences of missed learning opportunities. Missed learning opportunities can be recognized as time out of academic classroom settings and can lead to immediate and long-term academic challenges (Meltzer et al., 2017).

Gesell's theory of neuromaturation is predominantly applicable to this study. Since the theory of neuromaturation addresses brain development (Gesell, 1925), researchers can draw similarities between academic and behavioral implications as a result of the timeliness of brain development. Further examination of existing literature shows a comparison between reading performance and behavioral performance in like peers (Follmer, 2018). Literature indicates those who struggle in either or both reading and behavioral performance may not be as neuromature as like peers (Berninger et al., 2017; Follmer, 2018; Zhang, Bingham, & Quinn, 2017).

Özek examined grade retained students, standardized reading scores, and discipline occurrences and events. Findings showed each point below the minimum score, the cutoff rate, school administrators observed a heightened disciplinary rate of five incidents per 100. School leaders' viewpoints and subsequent interpretation of said viewpoints were the focus of the study rather than viewpoints from students' perspectives. The current study examined real, existing data attributed to individual students.

Grade retention is not favorably looked upon as a solution for struggling students (Jimerson et al., 2007). Grade retention can have long-lasting academic and developmental outcomes for grade retained students (Kenneady, 2004; Martorell & Mariano, 2018). As well as inhibiting a grade-retained student's ability to continue with his or her social-emotional development, little evidence of long-term academic gains is evident by examination of standardized test scores and classroom evaluation data (Jimerson et al., 2007; Klapproth et al., 2016). Research suggests a likelihood of correlation between a student's reading ability and his or her likelihood of an office discipline referral applicable to non-grade retained students (Martorell & Mariano, 2018; Mundy et al., 2017; Özek, 2015; Pimperton & Nation, 2014).

Martorell and Mariano's (2018) research did not discuss a connection between reading ability and behavior outcomes, allowing the contribution of new literature to the field of educational leadership and school improvement.

Statement of the Problem

The problem is students' disruptive classroom behavior and removal from class impacts on reading instruction, which has an impact on students' performance and assessment scores (McIntosh et al., 2006). Students who are behind or struggling with one or more concepts find academic challenges multiply quickly and some struggling students may engage in disruptive behaviors (Bruhn & Watt, 2012). Data indicate any student who is two or more years behind in reading ability is at risk of truancy, time out of class, and dropping out of school (Kane County Regional Office of Education, 2016). Ford, Olmi, Edwards, and Tingstrom (2001) suggested students with chronic challenges, such as staying focused and on task while in class, are more likely to receive office discipline referrals and time out of class resulting in missing valuable reading instruction. Teachers often use behavior modification techniques to address problematic behaviors, but fail to investigate reading-specific academic deficits manifesting as poor performance on standardized tests or lower grades (Oliver et al., 2011).

A review of the literature revealed a gap exists when examining students' disciplinary data as a function of reading performance as measured by standardized academic test scores. Existing empirical research provided insight on and relationships among student grade retention, behavioral-related disruptions, and material hardships related to student reading achievement (Kenneady, 2004; Nunes, Balcão Reis, & Seabra, 2018; U.S. Department of Health and Human Services, 2004). There exists evidence in current literature showing retention as an ineffective intervention to address reading ability (Kenneady, 2004). Relationships between reading ability

and behavioral outcomes lack clarity and needs to be explored in further detail. Little evidence in research literature leads to an understanding of relationships and likelihood of a particular student being at a level of risk for behavior or disciplinary incident based on said student's reading performance ability.

Student reading achievement and behavior are indicators monitored regularly by schools and state agencies and are causes of grade retention throughout an educational system (Marchbanks et al., 2013; Mathys, Véronneau, & Lecocq, 2017). Jimerson et al., (2007) discovered trends between students with behavioral and disciplinary actions are at high risk for dropping out of school or facing grade retention. Data from Jimerson et al.'s study did not indicate nor categorize behavior as major or minor or provide any other characterization for said behavior events. Data from Jimerson et al.'s study were inclusive of all behavioral events regularly collected and recorded. One gap in Jimerson et al.'s study was investigating whether grade retention or repetition was effected by behavioral concerns and dropout potential. Özkanal and Arıkan (2011) suggested a connection between student attendance at school and comparable attendance in workforce environments drawing a comparison to how behaviors learned in school carry through to employment. Existence of such a relationship compliments Marchbanks et al.'s findings relating economic effect to grade retention or high school dropout status. Ryan's (2011) research supports economic claims of Özkanal and Arıkan and Marchbanks et al. citing struggles of individuals who do not have a high school diploma to earn a living wage or effectively and happily engage within a community.

Research shows post-secondary options become limited for students who have been retained or dropped out of school, those with behavioral concerns, and students who struggle with reading ability all of which have an effect on the individual's life and on society as a whole

(Marchbanks et al., 2013; Xia & Kirby, 2009). Retention or dropping out of school has an economic impact on society. Marchbanks et al. (2013) suggested grade retention and high school dropout amount to over $752 million in local, state, and national economic disruption. Not only is the problem current, relevant, and important, but academic and behavioral challenges are ongoing challenges for school systems, society, and students who struggle with reading ability, behavioral challenges, and students facing grade retention as well as for fellow peers who have or do not share those same challenges.

Given the importance of the current study's focus, conducting further research to gather information on reading ability and behavioral events is imperative. The current study holds potential to contribute to the gap in literature through data collection and analysis taking a closer look at reading ability and behavioral outcomes. School leaders, educators, educational researchers, and parents can benefit from the current study which contributed to current literature and practices among educational leaders.

Purpose of the Study

The purpose of this quantitative investigation was to investigate the relationship between seventh-grade reading scores on standardized tests and office discipline referrals and indicators of grade retention among seventh grade students in Northwest Suburban Chicagoland school district. The current quantitative correlational research study includes an analysis of reading scores of seventh grade students in Northwest Suburban Chicagoland school district and identifies trends in and relationships among reading scores as compared to office discipline referrals. The study includes an analysis of reading scores of the same cohort of students and identifies trends in reading scores as compared to grade retention occurrence. Obtained office discipline referral data and grade retention occurrence data were the study's dependent variables

with scores from Illinois Standards Achievement Test (ISAT) from 2014-2015 and 2015-2016 serving as independent variables. Data were formally requested from school district officials. The study collected standardized reading scores data of 100 students from 2012-2013 school year when students from Northwest Suburban Chicagoland school district were in seventh grade, and office behavior referral data from 2014-2015 and 2015-2016 school years when students were in ninth and tenth grade respectively. Grade retention status was collected in terms of occurrence frequency from 2014-2016 school years while participants are enrolled in high school. Student data were referred to as cohorts for purposes of identifying specific groups.

Data, discoveries, and results from the study provides educational leaders with a method of identification schools can use to make data-based decisions to identify students scoring lower on reading portions of a standardized test or those students incurring office discipline referrals and implement academic reading interventions at an earlier age. The study was designed to provide a means of early identification, which may lead to reduction of future occurrences of office discipline referrals. School leaders can evaluate and determine needed support relating to middle to high school transition among middle and high school feeder patterns.

Significance of the Study

This study was designed with the intention of providing educational practitioners with real-time, data driven strategies employable when examining student academic and behavioral data. School leaders, educational instructors, and support staff are charged with data driven decision making including lesson planning and assessment creation–differentiation–on an ongoing, regular basis (Marshall, 2016). School and educational leaders can use data driven methods to identify students in need of academic and behavioral interventions, possibly before either manifests into a major challenge for a student.

The significance of this current study was to provide indication of possible student discipline issues based on a reading score or score range. This study contributed new literature to educational leadership theories and practices, educational improvement processes, student supports, and data driven decision-making. Providing further literature on educational leadership and educational improvement movements, results were narrowly focused on relationships between reading scores and behavioral data.

Research Questions

Research questions were designed to guide this quantitative correlational research study. Data were requested from school district officials and included an anonymous group of 100 students' reading scores, behavioral data, and retention status. Studying reading scores as a predictor of student behavioral incidents provided new insight to existing literature and implementable strategies for school leaders. This current study was designed to address three research questions.

Research question 1. What degree of correlation, if any, exists between seventh-grade standardized reading assessment scores from ISAT and future occurrence of office discipline referrals during students freshman and sophomore years?

Research question 2. What degree of correlation, if any, exists between student office discipline referrals and average to above average standardized reading assessment scores?

Research question 3. What degree of correlation, if any, exists between seventh-grade standardized reading assessment scores from ISAT and grade retention throughout students' freshman and sophomore year in high school?

Hypotheses

Students with below-average scores on seventh grade ISAT standardized reading assessment do account for more office discipline referrals than students with average to above-average scores on the same assessment. An overall hypothesis was scores in below average ranges, indicate a predictor of future student office discipline referral events and time out of class. Office discipline referral data and grade retention occurrence frequency serve as the study's dependent variables with scores from ISAT assessment from 2014-2015 and 2015-2016 serving as independent variables.

$H_0$1: There is no significant correlation between seventh-grade standardized reading assessment scores from ISAT and future occurrence of office discipline referrals during students' freshman and sophomore years.

H_a1: There is significant correlation between seventh-grade standardized reading assessment scores from ISAT and future occurrence of office discipline referrals during students freshman and sophomore years.

$H_0$2: There is no significant correlation between student office discipline referrals and average to above average standardized reading assessment scores.

H_b2: There is significant correlation between student disruptive behaviors and average to above average standardized reading assessment scores.

$H_0$3: There is no significant correlation between seventh-grade standardized reading assessment scores from ISAT and future occurrence of office discipline referrals during students' freshman and sophomore years.

H_c3: There is significant correlation between seventh-grade standardized reading assessment scores from ISAT and future occurrence of office discipline referrals during students' freshman and sophomore years.

Theoretical Framework

This current study was centered on Gesell's theory of neuromaturation. Gesell's (1925) neuromaturation theory explains ongoing frontal lobe brain development through an individual's first three decades of life and how experiences effect said development. The brain's frontal lobes provide executive functioning within each person, including memory and cognitive characteristics and abilities (Gesell, 1925; Kalpakci et al., 2018). Cognitive development within an individual's brain is reliant on appropriate academic and behavioral learning opportunities and social interactions (Follmer, 2018). Time in class, physically being present, is important to acquisition of academic and behavioral skills but consistent time in class can be interrupted due to behavior incidents (Meltzer et al., 2017). Meltzer et al., describe missed learning opportunities as being out of the classroom setting due in part to behavior, grade retention, or small group or individual interventions.

This current study provided school leaders and educational researchers with data and literature on reading scores as possible predictors of student behavior events. School leaders can address systematic change through transformational leadership methods. Bass and Riggio (2006) describe transformational leadership as a process where leaders address subordinates' needs through leaders' actions, which lead to improvement with subordinates' motivation leading to overall improvement and collective organizational success. Transformational leaders recognize subordinates' and colleagues' wishes to be challenged and empowered and can accomplish said wish by showing trusting relationships as well as allowing colleagues to demonstrate leadership

abilities at all levels (Bass, 1999; Bass & Riggio, 2006). Research shows school leaders alone cannot effect change but can influence colleagues by modeling, empowering, and posing intellectually challenging and rewarding tasks such as instructional changes within a school.

Educational leaders, building administration, and staff can use outcomes from this current study to make data-based informed decisions, school-wide goals and values which may, through a shared common understanding, provide academic improvement, increased student success, and staff engagement (Andersen, Bjørnholt, Bro, & Holm-Petersen, 2018; Lee & Kuo, 2019). Educational-related transformational leadership focuses on motivating others to be actively engaged in ongoing improvement and change which may lead to improved experiences for all students (Maier et al., 2016). Data-driven educational decisions, whether at the school level or classroom level, include input from many professionals in order to fulfill a larger mission or vision shared by all (Lee & Kuo, 2019).

This quantitative correlational study's research questions and hypotheses were developed based on transformational leadership and data-driven decision making opportunities for school leaders and staff. Research questions and hypotheses identify student office discipline referrals and grade retention occurrence as dependent variables and ISAT student scores as independent variables. Chapter 2 further provides in depth evidence addressing standardized reading scores and relationships with student discipline. Chapter 2 addresses school leaders' and teams' use of discovered data as a means of addressing school improvement in the areas of academics and student behavior.

Definitions of Terms

The study examined existing data, some of which needing clarification, from standardized assessments and school district specific student academic and behavior information.

Definitions are provided for the study's independent, or predictor, variable and two dependent, or criterion, variables. Terms referenced in this study having multiple or universal interpretations or applications are included within definition of terms. Defined terms are included to provide context to specific use within the study.

Absenteeism. Not coming to school or not following the courses of being absent without having a valid excuse (Özkanal & Arıkan, 2011).

Cohort. Group of students based on the number of students who enter Grade 9 for the first time, identified by school year, such as 2013-2014 (Illinois State Board of Education, 2017).

Freshman. A student in the first year of high school (Merriam-Webster, n.d.-a).

Grade retention. [dependent, criterion variable] Retained in a grade; repeating a grade (Rhodes, Thomas, & Liles, 2018).

Illinois Standards Achievement Test (ISAT). [independent, predictor variable] Assessment measuring student achievement based on Illinois learning standards which reports reading results as Lexile reader measures (Illinois State Board of Education, 2013a).

Lexile reader measure. A person's reading ability, as measured by Lexile Framework, based on reading ability and text measurement (MetaMetrics, n.d.).

Neuromaturation. Process of development in which children and adolescents progress through both physical and mental growth sequences or milestones at his or her own pace (Gesell, 1925).

Office Discipline Referral. [dependent, criterion variable] A formal submission by a school staff member, typically to a principal or designee, documenting a student's behavior concern or policy violation (Allman & Slate, 2011).

Reading Comprehension. The process of simultaneously extracting and constructing meaning through interaction and involvement with written language (Ness, 2016).

Response to Intervention (RtI). A multi-tier approach to the early identification and support of students with learning and behavior needs (National Center for Learning Disabilities, n.d.).

Sophomore. A student in the second year of high school (Merriam-Webster, n.d.-b).

Limitations

Price and Murnan (2004) describe limitations as methodological or other design characteristics having an effect or influence on the execution of this study or findings. This study was limited to 100 students and 4 schools within a school district. One hundred students is a narrow group of participants compared to actual enrollment in the seventh grade; the sample size was appropriate based on results obtained from a priori power analysis. The computed minimum sample size of 84 was increased to 100 to account for missing data pieces and errors in data. The sample size was determined using a 95% significance level and 80% power of test for a correlational statistical analysis (Creswell, 2013; Sullivan & Feinn, 2012).

Illinois Student Achievement Test (ISAT) reading data for seventh-grade students during 2012-2013 school year is the only standardized testing data available. Data for this study were requested from school district officials specific to assessment results of ISAT seventh-grade reading assessment; examining each individual student holistically was not feasible. Attendance habits of randomly chosen participants was not be examined. Attendance habits may have provided insight on days attended, days missed, and may be used to hypothesize what if any, effect non-disciplinary related time out of class has on reading scores. Although attendance may be a factor in student academic achievement, attendance is an external factor schools cannot

necessarily control as attendance encompasses illness, truancy, court appearance, family issues, and vacation (Kane County Regional Office of Education, 2016).

Data were collected and analyzed without having any specific information about student participant data. Data were random and did not specifically include or exclude a student who (a) had repeated any grade(s), (b) was on track with his or her high school cohort, or (c) had an eligibility for special education or Section 504 services. In an effort to control internal and external influential factors, raw data were requested from the school district's student management information system, in the form of an Excel spreadsheet. Standardized testing data were reported as a scaled score based on ISAT's Interpretive Guide 2013. Behavioral data were reported as frequency of event and event name. To avoid bias in this current study, behavioral data were reported for analysis without making interpretations of or grouping like behavioral events.

Limitations to internal validity may have included events occurring with participants outside of the control of the school district, such as trauma, medical issues, mobility within the school district, or change in family dynamics. Any such event may not necessarily have had a direct effect on reading ability or behavior events but may have had an overall influence within a participant. Any such event would not be identifiable from requested data. A limitation to external validity may include an inability to identify existing discrepancies in standardized test administration among the participants. Scores are reported as a standalone data set devoid of explanation of exceptionalities or discrepancies with test administration.

Scope and Delimitations

This current study focused on four schools within a large school district in Illinois. To determine a required minimum sample size, a priori power analysis was conducted. Most

quantitative studies rely on 95% significance level and an 80% power of test to provide an appropriate sample size (Creswell, 2013; Sullivan & Feinn, 2012). Using G*Power 3.1.9.2, identifying a 95% significance level, 80% power of test, and statistical test of correlational analysis, the minimum sample size is 84 (Faul, Erdfelder, Buchner, & Lang, 2009; see Appendix A). To account for missing data or any other data related error, 100 student records were requested.

The sample consisted of 100 students' ISAT standardized test scores from students' seventh-grade year and behavioral data from students' freshman and sophomore years. The scope can likely allow for generalizability of findings to other educational institutions outside this current study's schools and district since data on both topics are readily available for most, if not all, schools and districts within Illinois. Should an educational institution choose an assessment different from ISAT, reading levels or Lexile reader measure can be applied to any valid and reliable standardized assessment. Data requested provided raw assessment scores and behavior events for each participant and did not provide external factors such as teacher characteristics: teacher effectiveness, staff intervention, benefits or shortfalls in curriculum design or delivery. Said external factors may or may not have had an effect on participants' assessment scores or behavior events.

The study covered a four-year school year period from 2012 to 2016. The timeframe was appropriate in order to allow data collection with fidelity from three different, sequential school years. School years 2012-2016 were chosen to comply with school district officials' request to use data for a cohort who has graduated and exited the system. Data were requested from school district officials and were requested to include pseudonyms rather than specific student-identifying data pieces such as names and student identification numbers. Data were requested

via random collection from existing data sources within the student information management system with permission from appropriate school district officials.

No students who were at the time receiving or had received special education services or accommodations from a Section 504 or healthcare plan were deliberately included or excluded. No distinction was made among aforementioned populations of students identified as receiving or have received special education services or accommodations from a Section 504 or healthcare plan since all students take standardized assessments unless a student-specific education plan states, resulting from a multidisciplinary team determination, an alternative assessment is appropriate for said student. All ISAT assessments administered were administered to appropriate populations.

Assumptions

This study assumed equal assessment administration conditions and setting existing among participating school sites. All students had equal opportunities and equal educational settings while sitting for seventh-grade ISAT assessment administration. This assumption was necessary since no indicator may exist to address opportunities and settings. Equal access to curriculum, instruction, and other academic-relevant school offerings providing academic preparation to students are assumed to exist among participating school sites within Northwest Suburban Chicagoland school district. This assumption was necessary since there was no existing data to show specific curriculum, instructional, or other academic-relevant offering other than a district curriculum scope and sequence.

Continual enrollment at Northwest Suburban Chicagoland school district schools starting with participants' seventh-grade year through the completion of participants' sophomore year was assumed. This assumption was necessary to ensure consistency among students' school

settings and experiences. Students classified as non-native English speakers were assumed to be evaluated by appropriate educational professionals and, based on results, administered an appropriate assessment for seventh-grade. This assumption was necessary to ensure students were administered the ISAT assessment in students' native language and a potential language barrier did not influence scores. Behavior data were assumed accurate and reported incidents aligned to Northwest Suburban Chicagoland school district's student handbook.

Chapter Summary

The purpose of Chapter 1 was to provide an introduction and overview of this study. The purpose of this quantitative correlational study was to investigate the relationship between seventh-grade reading scores on standardized tests and office discipline referrals and indicators of grade retention among seventh grade students. Data were collected on standardized reading scores data from 2012-2013 school year when students from the Northwest Suburban Chicagoland school district were in seventh grade. Student office discipline behavioral referral data and grade retention occurrence frequency from 2014-2015 and 2015-2016 school years, were requested. These data correspond to student participants when said participants were in ninth and tenth grade respectively. The problem addressed is how students' disruptive classroom behavior and removal from class impacts on reading instruction, which has an impact on students' performance and assessment scores (McIntosh et al., 2006).

Chapter 1 included background of the problem, study's significance, research questions, hypotheses, theoretical framework, assumptions, scope and delimitations, and limitations. Chapter 1 provided an explanation of how this study contributed to existing literature and provided data-driven methods immediately implementable by educational professionals. Chapter 1 provided an abbreviated overview of relevant literature. Chapter 2 will present an extensive

review of literature related to relevant framework, relationships among reading level and behavioral referrals, summary of existing research, and identification of gaps in existing literature.

Chapter 2: Literature Review

The degree to which seventh-grade reading scores influence office discipline referrals is a concerning topic among educators and educational scholars. By some estimation, academic costs of behavioral discipline parallel issues of limited learning opportunity and lack of achievement in an academic arena (Marchbanks et al., 2013). In combating effects of this suggested correlation, relationships between behavioral intervention and academic success is bridged to benefit students. The purpose of this quantitative study was to examine possible relationship between the independent variable of performance on reading tests and the dependent variable of disciplinary referrals.

The purposes of literature review include (a) discussing theories relevant to posited relationship between reading level and disciplinary referrals, (b) discussing and synthesizing empirical findings relevant to relationship between reading level and disciplinary referrals, (c) discussing academic initiated grade retention effects on student academic and behavioral achievement, and (d) identifying gaps in existing literature base. Each of these purposes of literature review is addressed in a distinct section within Chapter 2. Identified empirical studies are discussed in substantial detail, regarding analysis of individual studies, interrelation of key studies to each other, and association of studies with theoretical framework of neuromaturation. The main identified gap in literature was inability of existing studies and approaches to estimate a likelihood of a particular student being at a level of risk for a disciplinary infraction based on reading performance.

Literature Search Strategy

The literature strategy was based on searching peer-reviewed, scholarly articles. Literature was retrieved from academic databases such as: ERIC, Google Scholar, Academic

Search Ultimate, Education Abstracts, Education Full Text, Education Research Complete, JSTOR, ProQuest Education Database, and Web of Science. State and National Departments of Education resources were utilized for literature review. Reviewing searched articles' reference list provided additional relevant resources for exploration and review. Table 1 contains a comprehensive list of search terms utilized.

Table 1

Search Terms for Scholarly Article Search

Key Term	Search Term	Two Search Terms	Three Search Terms
Reading performance	disciplinary referrals	disciplinary referrals and standardized test	disciplinary referrals and standardized test and grade retention
	behavioral problems	behavioral problems and standardized test	behavioral outcomes and standardized test and grade retention
	behavioral outcomes	behavioral outcomes and standardized test	
	grade retention		
Reading level	disciplinary referrals	disciplinary referrals and standardized test	disciplinary referrals and standardized test and grade retention
	behavioral problems	behavioral problems and standardized test	behavioral outcomes and standardized test and grade retention
	behavioral outcomes	behavioral outcomes and standardized test	
	grade retention		
Reading comprehension	disciplinary referrals	disciplinary referrals and standardized test	disciplinary referrals and standardized test and grade retention
	behavioral problems	behavioral problems and standardized test	behavioral outcomes and standardized test and grade retention
	behavioral outcomes	behavioral outcomes and standardized test	
	grade retention		

The combination of search strings and databases listed in Table 1 resulted in identifying several hundred possible empirical articles for inclusion. The following inclusion criteria were applied: (a) publication in a peer-reviewed journal with named editors and an explicitly identified peer-review process, (b) publication in a scholarly journal rather than an industry journal, (c) publication in the English language, (d) quantitative methodological orientation, (e) treatment of reading, in some form, as an independent variable or as a correlate of behavioral or disciplinary outcomes, (f) accessibility and availability of full-text articles, and (g) publication from 2014 onwards.

The search procedure involved application of the seven criteria above to all results for all search strings applied to each database consulted. Additional searches involved all search strings with a less restrictive application of search criteria resulting in a small number of non-peer reviewed sources. Despite the breadth of databases and search terms, relatively few empirical studies met identified inclusion criteria. Because of a relatively low number of relevant studies identified through the search procedures, one may (a) offer a relatively detailed discussion of each relevant study, (b) highlight conceptual and methodological links between studies, (c) provide a more detailed discussion of gaps, and (d) tie each study closely to identified theoretical framework of the study.

Theoretical Framework

Henderikus (2010) described a theory as typically containing three elements. First, researchers use a theory to describe a phenomenon. Second, researchers use a theory to explain a phenomenon or relationships between phenomena. Third, researchers use a theory to make predictions. Based on Henderikus' framework, a theory relevant to topics of reading

performance and disciplinary referrals should contain each of these three elements. One matching theory to each of Henderikus' functions is neuromaturation.

According to theory of neuromaturation, frontal lobes of the brain develop for at least the three first decades of life (Gesell, 1925). The main function of the frontal lobes is to provide executive control such as planning, impulse control, memory, and coordination of these and other cognitive characteristics into a rational self (Friedman & Miyake, 2017; Kalpakci et al., 2018; Meltzer et al., 2017; Nitschke et al., 2017). Because Gessell's theory of function included an idea of affordance –what fits between a student and his or her environment – worth noting is students without proper environmental affordance may suffer consequences of missed learning opportunities. Missed learning opportunities are described as time out of class due to student behavior, receiving academic or behavioral intervention outside of classroom settings during instructional times, and grade retention (Meltzer et al., 2017). Grade retention is not favorably looked upon and research indicates grade retention can have a negative result on student's academic, behavioral, and social-emotional development (Jimerson et al., 2007). Research suggests providing cognitively appropriate reading interventions for students considered at-risk, can increase struggling students' reading skills and retaining a student, who would be developmentally and academically behind same age peers, can result in negative or no effect on student achievement (Jimerson et al., 2007).

The theory of neuromaturation is particularly applicable to current study. Studies, such as Johnson, Sudhinaraset, and Blum's (2009) examination of neuromaturation and adolescent risk taking, have indicated one can use neuromaturation to explain both vulnerabilities and potential of the adolescent brain. Johnson et al. (2009) determined neuromaturation research provided unique opportunities to investigate roles or effects of social environments of an

adolescent and what factors occurred when developmental processes were shaped during this time. Researchers can use the theory of neuromaturation to describe both reading skill and unruly or noncompliant behavior (Johnson et al., 2009). Researchers have associated both reading and appropriate behavior with developing frontal lobes to (a) plan complex actions, (b) inscribe learning events into memory, and (c) control impulses (Savioja, Helminen, Fröjd, Marttunen, & Kaltiala-Heino, 2017; Schwartz, Connolly, & Brauer, 2017; Yang, Joshi, Jahanshad, Thompson, & Baker, 2017).

In a behavioral context, neuromature individuals can (a) plan responses to disciplinary requirements, physical settings, and interpersonal dealings in a manner which reduces conflict and increases cooperation; (b) remember and be influenced by consequences of both positively and negatively reinforced behavior; and (c) control impulses based on violent verbal or physical responses (Savioja et al., 2017; Schwartz et al., 2017; Wang et al., 2016; Yang et al., 2017). In context of reading, neuromature individuals can (a) plan to read, (b) remember what has been read and integrate into an interpretation of a text or a real-life response, and (c) control impulses to abandon or not initiate reading (Berninger, Abbott, Cook, & Nagy, 2017; Follmer, 2018; Zhang et al., 2017). Researchers have suggested individuals who have struggled with reading performance, as well as behavior, may not be as neuromature as like peers (Berninger et al., 2017; Follmer, 2018; Zhang et al., 2017).

In addition to providing phenomenological descriptions of reading and behavior, researchers can use theory of neuromaturation to explain and forecast frameworks for a link between two theories (Berninger et al., 2017; Follmer, 2018; Zhang et al., 2017). If reading performance and behavior are both linked to neuromaturation, there may be a significant negative correlation which may exist. Meaning, individuals with higher levels of reading

performance may have lower levels of disciplinary problems, and individuals with higher levels of disciplinary problems may have lower levels of reading performance.

Henderikus (2010) noted theories required empirical testing. Theory-testing through empirical analysis is a complex process because one can define concepts in theories as quantified and explored (King & Finn, 2017). In the case of the theory of neuromaturation, empirical analysis may reveal an absence of a correlation or even a positive correlation between reading performance and disruptive behavior. Conversely, empirical analysis can indicate a negative correlation between these two variables. The findings of empirical analysis can depend on ways one can define variables of reading performance and student behavior. Regardless of what empirical analysis may reveal, theory of neuromaturation appears highly suited to the current study because neuromaturation theory's explanation of a possible link between reading performance and student behavior as likely meriting disciplinary referrals (Henderikus, 2010).

Research Literature Review

The purpose of this section of the literature review is to discuss, analyze, and synthesize empirical studies relevant to topics of study and is subdivided into four sections. In the first section, studies based on grade retention and low levels of reading performance are discussed. In the second section, a longitudinal study on grade retention is discussed. The third section contains a discussion of articles on reading skills and behavioral outcomes and effects on grade-retained students. The fourth section consists of identifying gaps in literature.

Academic Outcomes of Grade-Retained Students

Retention as an educational tool to address under-achievement has been a controversial issue, especially since results in literature are not consistent. Most studies used propensity scores to compare two groups of students about impacts of grade retention: those students grade

promoted to the next grade and students retained in the same grade. Academically at risk is used to describe one-third of all students within the United States (Schargel & Smink, 2001). Students considered at risk, especially when identified at early grade levels such as elementary school, are commonly associated with behavioral challenges and discipline, dropping out of high school, and grade retention (Alexander, Entwisle, & Kabbani, 2001; Hickman, Bartholomew, Mathwig, & Heinrich, 2008). Students who struggle academically reported feeling alienated, not supported, or disconnected from peers and teachers and suggest said feeling lingers with a student even when a student realized some success (Bailey & Stegelin, 2003).

Klapproth et al. (2016) examined how grade retention in secondary school would affect academic achievement and psychosocial adjustment of students. Klapproth et al. (2016) investigated short-term and medium-term effects of grade retention on student outcomes. Klapproth et al. (2016) used propensity scores to select a control group of students who were promoted who were similar to characteristics of students who were retained, and as a result, found three major findings.

The first major finding is grade retention only has short-term benefits for retained students. The second major finding is results of standardized achievement tests revealed no significant differences in results between students who were retained and students who were promoted. The last major finding was the two groups of students only differed in self-concept, with non-retained students having a more favorable sense of self-concept. Self-concept was described as a student's view of him or herself compared to peers and included comments on a student's perception of academic and physical abilities and emotional state (Klapproth et al., 2016).

Similar to Klapproth et al. (2016), Nunes et al. (2018) measured impact of grade retention of students to subsequent academic performance since there was no way to control for initial level of ability of students during time students were retained. Nunes et al. (2018) used propensity scores to match two groups of students regarding socioeconomic status and scores in national exams. The results indicated a situation where grade retention had a positive impact to future academic achievement of students for a short, immediate time. Nunes et al.'s study did not measure or otherwise evaluate or observe effects on grade retained students' social or emotional development. In cases where there were significant impacts to academic achievement of retained students, results were small and short-lived. Jacob and Lefgren's (as cited in Nunes et al., 2018) 2004 study of retained 3rd and 6th grade students in Chicago showed positive results for retained 3rd grade students for year one and either null or negative results for subsequent years; there was no effect for 6th grade retained students during the same time period.

Students who received discipline at school are at risk for being retained or being a school dropout (Jimerson et al., 2007). Marchbanks et al. (2013) examined impacts of school discipline on students' risk for grade retention and school dropout. Marchbanks et al. (2013) track a statewide sample of 7th grade through 12th grade students. Results revealed school discipline is linked with at least 4,700 grade retention cases annually in Texas. Marchbank et al.'s Texas study showed trends of high discipline rates among minority groups and data show higher frequency among behavior-related interruptions. One disadvantage of grade retention is the additional costs of instruction amounting to almost $41 million dollars (Marchbanks et al., 2013). Economic effects of dropouts amount to $711 million annually (Marchbanks et al., 2013). High school dropouts not only cost educational institutions millions of dollars in remediation but also prove detrimental to the high school dropout's lifelong earning potential. Educational agencies

should review current academic and behavioral intervention as means of disrupting behavior- and disciplinary- related student retention resulting in undesirable academic outcomes and somber economic liabilities for students, school systems, and society as a whole (Marchbanks et al., 2013).

Mathys et al., (2017) tested whether grade retention during students' transition to secondary school has an impact to psychosocial adjustment of students. Mathys et al. (2017) conducted a quasi-experimental design and used propensity score matching with 181 students in a Belgium secondary school. The results revealed grade-retained students experienced decreased self-esteem and intrinsic and extrinsic motivation. Grade-retained students failed to show a decrease in delinquent and aggressive behavior compared to match promoted students. Grade retention appeared to have negative impacts on psychosocial adjustment of students (Mathys et al., 2017).

Like Mathys et al., Council, Cartledge, Green, Barber, and Gardner (2016) conducted a descriptive study to examine whether computer-based, repeated reading intervention is correlated with improved reading performance and social behavior for three urban African American girls presenting with both academic and behavioral risk. The culturally relevant reading passages were delivered through computer software with an attempt to increase reading fluency of young students. Data surrounding reading performance and student behavior were collected before and after intervention program time. After implementation of reading intervention over a period of time, reading and behavioral outcomes of participants improved significantly. Results indicated through AIMSweb reading intervention, each girl increased reading fluency scores by a minimum of 7.0 points with one participant increasing 30.7 points (Council et al., 2016).

Post-secondary educational options are not favorable for individuals who drop out of school (Xia & Kirby, 2009). Allensworth (as cited in Xia & Kirby, 2009) studied students who were retained one time during an academic career and found students who were retained one time were up to 50% more likely to drop out of school than someone not retained. Allensworth (as cited in Xia & Kirby, 2009) studied students who were retained two times during an academic career and found those individuals were up to 90% more likely to drop out of school than someone who was not retained. Pursuing post-secondary education, including vocational training or military enlistment services, is challenging for students who have not graduated high school.

Fine and Davis (2003) studied post-secondary enrollment trends among struggling and grade retained students. Fine and Davis (2003), much like Klapproth et al. (2016), measured grade retention effects specifically regarding students' academic performance and not academic potential or ability. Fine and Davis' (2003) study results indicated students who struggled through school but graduated, regardless of any student-specific academic indicators such as grade point average and ACT or SAT score, were more than 50% as likely to pursue post-secondary educational opportunities as students who dropped out of school or did not obtain a high school diploma or equivalent.

Longitudinal Outcomes of Grade-Retained Students

Given recent increasing rates of grade retention and associated negative effects, better understanding is needed to identify intervention practices to mediate risk factors of grade retention (Kenneady, 2004). Several researchers conducted longitudinal studies about effects of grade retention to students. Hughes, West, Kim, and Bauer (2017) conducted a 14-year prospective study investigating effects of retention of students in Grades one to five to high

school completion data of student participants. The study started with 734 Grade 1 students who were academically at-risk. After 14 years, 477 students already earned a diploma and 21 students earned a GED, both indications of high school graduation. After the same 14 years, 110 students dropped out of school and 126 students have missing school completion data indicating no official record of high school graduation. Hughes et al. (2017) found grade retention led to a significant likelihood of dropping out of high school. Hughes et al.'s research did not indicate the likelihood of a grade retained student, who was also a high school dropout, obtaining a GED or high school equivalent at any time in the student's life.

In another prospective, longitudinal study, Yang, Chen, Rhodes, and Orooji (2018) examined relationships between school engagement, material hardship, and grade retention among 4,329 elementary school students. Yang et al. (2018) used multilevel logistic regression to address nesting effects of students within schools. Results indicated 42.34% of students in the sample repeated at least one grade (Yang et al., 2018). Material hardship was associated with a greater chance of grade retention as a result of student's lack of basic needs (Yang et al., 2018). Furthermore, this relationship was mediated by school engagement or lack of partially due to a student's family dynamics or external influences (Yang et al., 2018). The results of the study suggest the need for interventions to address material hardship and school engagement of students. Few studies addressed material hardship effects on family situations and what interventions, if any, were provided outside of the school setting.

Material hardship cannot be and is not defined in one way (U.S. Department of Health and Human Services, 2004). Researchers suggest material hardship varies based on how people experience hardship (U.S. Department of Health and Human Services, 2004). Researchers agree material hardship evaluations should exist to determine what extent a person or people are able

to meet basic needs, exclusive of income-based criteria used to determine poverty (U.S. Department of Health and Human Services, 2004). The U.S. Department of Health and Human Services (2004) suggests assessing basic needs for human survival such as clothing and shelter, medical and dental care, and food as a primary influencing factor when determining material hardship.

Material hardship was associated with a greater chance of grade retention. This relationship was mediated by school engagement. Yang et al. (2018) study's results suggest a need for interventions addressing material hardship and school engagement among students. Research indicates families and students experiencing one or multiple material hardships most likely have multiple challenges while in school and through development and neumaturation compared to higher income peers (U.S. Department of Health and Human Services, 2004).

Competitive employment opportunities for students who have dropped out of school are limited (Fine & Davis, 2003). Although post-secondary education is not necessary for competitive employment, individuals with post-secondary education or vocational training are more likely to employed earning higher wages than peers without education or vocational training (Fine & Davis, 2003). Eide and Showalter (2001) discovered students who were grade retained were compensated lower per hour wages and received lower benefits than like peers who completed high school.

In context of grade-retention policies in Florida and New York City, behavioral outcomes by be measured as a function of low reading performance (Martorell & Mariano, 2018; Özek, 2015). Low reading performance is, in both of these academic environments, a factor in grade retention. Measuring behavioral outcomes of reading performance-based grade-retained students, particularly in Florida, is a means of estimating effects of low reading performance on

behavioral outcomes. Neither study discussed specifics of reading passages used for assessments and whether said reading passages were provided to students in English only or made available in a student's native, fluent language.

Studies regarding grade retention due to low reading and mathematics performance were limited. Of specific concern was the lack of data addressing separately computed data analysis among reading performance and mathematics performance; many studies combined both reading and mathematics data for computational purposes. For this reason, identification and discussion of other studies is a necessity, where researchers have measured and modeled reading performance as a predictor of future behavioral outcomes. For example, King, Gonzales, and Reinke (2018) conducted a study longitudinal in nature; meaning, relationships were measured between low reading achievement measured at a certain point in time, often in lower grades, and several subsequent years of behavioral outcomes. King et al. (2018) used the technique of linear discriminant analysis and discovered three subgroups among students with disciplinary problems: (a) a subgroup who did not experience any academic risk including a subgroup with acceptable levels of both math and reading performance, (b) a subgroup who was below average in both reading and math, and (c) a subgroup who was substantially below average in both reading and math.

King et al. (2018) suggested a low level of reading performance might not be a unique predictor of disciplinary problems or behavioral challenges among students. If reading performance was indeed a unique predictor, then linear discriminant analysis or other, related techniques, such as *k*-cluster analysis, might have identified a subgroup of behaviorally disturbed students uniquely classifiable through low reading scores. One could consider King et al.'s (2018) findings alongside those of Özek (2015), who found low levels of reading performance

were predictors of behavioral problems among students in Florida. The state of Florida retains Grade 3 students who have unacceptably low reading levels. Özek's (2015) findings were unclear about whether gathered Floridian data indicated a unique effect of reading or whether students who were grade-retained in Florida happened to have low levels of math performance, as well as low levels of reading performance. The underlying conceptual and methodological issue, which was better addressed by a cluster analysis conducted by King et al. (2018) than in other scholar's designs, is whether low reading performance—more than, for example, math performance or a combination of math and reading performance—was a distinct predictor of subsequently disciplinary outcomes.

The theory of neuromaturation can be applied when addressing King et al.'s (2018) empirical finding of low levels of reading performance and math performance are part of the same cluster of predicting disciplinary referrals. The consensus in neuromaturation theory is complex cognitive acts, such as reading and mathematics performance, are similarly reliant on frontal lobes and reliance on association with planning, memory, impulse control, and unification of these cognitive skills as part of a brain's executive function (Friedman & Miyake, 2017; Gesell, 1925). Researchers have supported this consensus in a substantial cross-section of empirical work relying on neuromaturation theory (Berninger et al., 2017; Follmer, 2018; Friedman & Miyake, 2017; Kalpakci et al., 2018; Meltzer et al., 2017; Nitschke et al., 2017; Savioja et al., 2017; Schwartz et al., 2017).

Behavioral Outcomes of Grade-Retained Students

Noncompliant or disruptive student behavior in schools is an increasing and prevalent issue and a problem, when conjoined with poor performance in core subjects such as reading, often crosses paths with unarmed teachers (Kenneady, 2004). Recent surveys taken in United

States over past years have shown a significant increase in school staff who has dealt with what are considered extreme behaviors (Martorell & Mariano, 2018). Along with extreme behavior is the inevitable result of substandard learning opportunities, as teachers assume roles of mediator and disciplinarian, rather than educator (Martorell & Mariano, 2018). Researchers have established a trend between misbehavior and poor grade performance. These studies were examined in an effort to highlight current issues surrounding topical problems educators are facing.

In addressing issues surrounding influence of seventh-grade reading scores on adolescent behavior and disciplinary referrals within school systems, several studies were consulted. Two recent quasi-experimental researchers have examined relationships between reading performance and discipline-related behavior through phenomenon of grade retention (Martorell & Mariano, 2018; Özek, 2015). For example, in Florida and New York City, students who do not meet a reading performance cutoff score are retained a grade. In such cases, grade retention is a function of reading level, providing an obvious opportunity to examine relationships empirically between reading level and discipline-related outcomes (Martorell & Mariano, 2018; Özek, 2015).

Researchers of relationships between reading scores and poor behavior in school systems have contributed to overarching issues of environment-related causes of poor academic performance. Özek (2015) conducted an empirical study on relationships between reading performance and discipline-meriting behavior in context of the state of Florida. Özek explained Florida has a rule requiring students with below-grade reading skills to be retained in Grade 3. Özek took advantage of Florida's measurement of reading performance being measured as a continuous variable. Considering student discipline, operationally defined as number of

disciplinary referrals, could be measured as a continuous variable, data from Florida were used to fit linear models to relationships between reading performance and discipline (Özek, 2015).

Contextually Özek's (2015) study's stated purpose was to measure an effect on low reading performance initiated grade retention on student discipline. Özek (2015) defined student discipline through four variables: (a) incident rate, (b) students involved in disciplinary incidents requiring in-school suspension, (c) students involved in disciplinary incidents requiring out-of-school suspension, and (d) number of days suspended. In synthesizing information, Özek (2015) provided a clearer view of summative grades and discipline-worthy behavior. Özek (2015) discovered an existence of a statistically significant (at $p < .05$) relationship between independent variable and all dependent variables.

Özek's (2015) findings indicated lower levels of reading performance correlated with higher rates of disciplinary incidents. Özek's (2015) study's main limitation was the unit of analysis was school principals, rather than individual students. From the perspective of school's principal's office, scholars have noted the importance of forecasting negative outcomes for specific students. If principals and other educational personnel can better understand specific risks and environmental challenges faced by specific students, there is a greater likelihood students' risks can be preemptively lessened through timely application of support resources. In this context, a provision of school-level data, such as those of Özek (2015), are not necessarily of practical use, although such studies remain useful for establishing a general link between reduced reading performance and increased discipline-meriting behavior.

A more specific study covering grading policy and long-term behavioral impacts indicated a pattern of poor performance regarding behavior. Martorell and Mariano (2018) studied relationships between grade retention and behavioral outcomes in New York City, in

which there was a policy of grade retention based on unacceptable levels of academic performance. In New York City, grade retention policy covers Grades 3 to 8; students in any of these grades who do not demonstrate adequate academic preparation can be grade-retained. Chicago Pubic Schools had strict guidelines and policies on retention and promotion which were ended after a seven year study when the Board of Education realized said policies were not benefitting students nor were the policies resulting in raised or increased test scores (Herszenhorn, 2004). Some credit-based grade retention still occurs at high schools but academic and social promotion policies have been abandoned in the Chicago Public School system (Herszenhorn, 2004).

Özek (2015) noted Florida's grade retention policy in Grade 3 was based on low reading performance. Martorell and Mariano (2018) noted of New York City, grade retention could be based on low performance in either reading or mathematics. While New York City leaders keep data records related to both math and reading performance, Martorell and Mariano developed a statistical model which did not differentiate between individual effects of (a) low math performance and (b) low reading performance on disciplinary outcomes. For this reason, Martorell and Mariano's results were not as likely as Özek's (2015) results to measure effects of low reading performance on behavioral outcomes reliably and validly. Given the possibility students with acceptable reading scores were grade-retained because of math scores, Martorell and Mariano's (2018) results did not necessarily measure effects of reading performance on disciplinary outcomes. Martorell and Mariano (2018) discovered no statistically significant effects of grade retention on behavior, but this finding was not necessarily relevant to questions regarding whether low reading performance predicted increased behavioral problems among students.

In another study, Klapproth et al. (2016) examined effects of grade retention on academic well-being and personal well-being of students. Klapproth et al. (2016) concluded effects of grade retention appears to be negative to both academic and personal well-being. Moreover, Klapproth et al. (2016) concluded an effect is more salient for academic outcomes compared to personal well-being outcomes. Klapproth et al. (2016) acknowledged other factors might affect a negative influence of grade retention to students' well-being. Klapproth et al. (2016) recommended schools do not rely on grade retention as an immediate or default response to academic and behavior difficulties of students. Other than grade retention, there are other effective educational and policy responses to poor student academic performance.

Most states retain students especially when students lack basic reading proficiency (Martorell & Mariano, 2018; Özek, 2015). Schwerdt, West, and Winters (2017) examined retention probabilities under test-based promotion policy of Florida and effects on student outcomes through high school. Schwerdt et al. (2017) found large positive effects on achievements fade out when comparing retained students and non-retained students. Students who are retained have a high probability of increasing grade point averages and may take fewer remedial course in high school. Grade retention does not influence graduation rates of retained students.

In another study, Mahjoub (2017) estimated treatment effect of grade repetitions among junior high school students in France. In contrast to Schwerdt et al. (2017), Mahjoub (2017) found grade repetition in junior high school increases probability students may graduate from junior high school. These studies brought together sourced information. Researchers have established importance of this subject and variables to use when determining the relationships

between poor reading scores and poor or disruptive behavior in adolescents (Martorell & Mariano, 2018; Özek, 2015).

Martorell and Mariano (2018) and Özek (2015) utilized disciplinary outcomes, such as suspension rates, as dependent variables. Contrary to Martorell and Mariano and Özek, Chung (2015), Mundy et al. (2017), and Pimperton and Nation (2014) treated behavior by focusing on behavior itself. For example, these researchers have expressed interest in behavioral characteristics, such as impulsivity and aggression, as differentially distributed among poor readers and good readers. These studies are worth including in review of empirical literature because of a likelihood of a direct connection between behavioral outcomes and disciplinary outcomes.

In an efficient school environment, principals, teachers, and other stakeholders can successfully identify and refer unacceptable behavior (Gage, Sugai, Lunde, & DeLoreto, 2013; Hoffman, 2014; Skiba, 2014). If done consistently, the dependent variable of behavioral dysregulation—however operationally defined—may be an appropriate causation for disciplinary referrals. Consequently, researchers who use behavioral dysregulation as dependent variables should be included while reviewing of empirical literature (Chung, 2015; Mundy et al., 2017; Pimperton & Nation, 2014).

In one study, Mundy et al. (2017) studied relationships between reading scores and misbehavior in school based on data collected from a sample of 1,239 Grade 3 students in Australia. Mundy et al. defined behavioral problems dichotomously into categories of either (a) normal or (b) borderline/abnormal. Behavioral categorization of students was performed by students' parents. Mundy et al. (2017) measured influences of both (a) reading scores and (b) mathematical performances on dependent variable of behavioral categorization.

Mundy et al. (2017) used a linear regression model, rather than a logistic regression model, and discovered students in abnormal/borderline behavioral category had reading scores 47.80 points lower compared to students in normal behavioral category. From this finding, Mundy et al. concluded lower reading performance scores were a significant predictor of behavioral problems likely to lead to disciplinary referrals in school environments. Mundy et al.'s (2017) findings did not indicate effects of reading performance on behavioral problems was unique. In Mundy et al.'s (2017) model, students in abnormal/borderline behavioral category had math scores 37.70 points lower compared to students in normal behavioral category. Among confidence intervals for inferiority in reading performance for abnormal/borderline behavioral group was between -32.80 and -62.80, with confidence interval for inferiority in math performance for abnormal/borderline behavioral group at -21.50 and -53.90.

This aspect of Mundy et al.'s (2017) findings indicated a comparison to King et al.'s (2018) findings, who likewise found effects of reading and math on behavioral outcomes was statistically similar. Chung (2015) examined relationships between self-regulatory behavior and reading performance among a sample of Chinese students. Chung (2015) found significant (at $p < .05$) and positive correlations between these pairs: (a) behavioral regulation and phonological awareness ($r = 0.41$), (b) behavioral regulation and vocabulary definition ($r = 0.26$), (c) sentence comprehension and behavioral regulation ($r = 0.54$), and (d) behavioral regulation and morphological construction ($r = 0.36$).

Chung (2015) examined relationships between self-regulatory behavior and reading performance among a sample of Chinese students. Chung sampled 78 Grade 1 students from Hong Kong, with 39 of the sampled individuals being poor readers and the remaining 39 being adequate readers. Chung purposefully stratified and matched the sample allowing for students in

each group to be similar to each other in age, nonverbal intelligence, and parental education levels. One strength of Chung's study was poor and adequate readers were compared to each other on several measures of reading skill, including (a) vocabulary, (b) phonological awareness, (c) rapid digit naming, (d) awareness of morphological construction, and (e) sentence construction.

Chung (2015) measured behavioral self-regulation as an index measurement using Head-Toes-Knees-Shoulders (HTKS) task. Chung's (2015) findings were important in establishing several types of reading deficits, specifically morphological construction, phonological awareness, vocabulary, and sentence construction, might predict behavioral problems among students. Chung's (2015) findings were closely connected to chosen to the current study's theoretical framework of neuromaturation. As Chung explained, HTKS was a measurement of planning, memory, and impulse control, which were three key functions of maturing frontal lobes (Berninger et al., 2017; Follmer, 2018; Friedman & Miyake, 2017; Gesell, 1925; Kalpakci et al., 2018; Savioja et al., 2017).

In another study, Bernard (2017) evaluated influences of a teacher professional development program called *Attitudes and Behaviors for Learning* (AB4L) on student achievement, especially for students with poor performance in reading. Two primary school teachers received three half-day training sessions on how to teach with positive attitudes and behaviors for learning and those teachers applied what was learned during literacy classes back in classroom settings. Ninety-eight students were taught AB4L program while 86 students were not (Bernard, 2017). Results indicated teacher-rated and student-rated learning behaviors were positively associated with objective reading performance indicating the AB4L program has a significant effect on students in the lower 50% of the class in terms of reading comprehension

(Bernard, 2017). Bernard's (2017) study suggested students who participated in AB4L program showed not only improvements in reading comprehension but improvements in both behavior and learning.

Pimperton and Nation (2014) conducted a study, using data obtained from 244 schoolchildren ages between 8 and 10 years of age, on relationships between poor reading comprehension, working memory, and teacher ratings of behavior. Like Chung (2015), Pimperton and Nation (2014) measured reading skill through numerous variables, specifically matrix reasoning, text reading accuracy, reading comprehension, and phonemic decoding. Like Chung (2015), Pimperton and Nation's (2014) study included working memory. Pimperton and Nation's study was directly connected to the theoretical framework of neuromaturation, involving improving working memory as one of the key characteristics of individuals who are more neuromature compared to a like peer group (Friedman & Miyake, 2017; Kalpakci et al., 2018; Meltzer et al., 2017; Nitschke et al., 2017).

Pimperton and Nation (2014) indicated absence of statistically significant differences between poor readers and adequate readers in all but two dependent variables. Pimperton and Nation (2014) found poor readers (M = 84.66, SD = 4.13) were significantly worse in reading comprehension than adequate readers (M = 104.91, SD = 7.99, $p < .001$). Pimperton and Nation (2014) found poor readers (M = 90.21, SD = 10.71) had worse verbal working memory compared to adequate readers (M = 102.50, SD = 11.51, $p < .001$). Pimperton and Nation's (2014) findings are of importance because of relevance to neuromaturation. These findings constituted empirical support for a claim stating poor readers and adequate readers have at least one form of brain-level difference in verbal working memory. Neuromaturation theory indicates brain-level differences in working memory are likely to distinguish cognitive mature from

cognitively immature individuals regarding appropriate levels of cognitive development in different age groups (Friedman & Miyake, 2017; Kalpakci et al., 2018; Meltzer et al., 2017; Micai, Kavussanu, & Ring, 2015; Nitschke et al., 2017).

Pimperton and Nation's (2014) empirical findings were relevant to questions of whether variations in reading skill could predict variations in behavioral outcomes. The behavioral aspect of findings was based on a teacher rating scale. Teachers were asked to rate both poor readers and adequate readers on variables of oppositional behavior, hyperactivity, attention deficit/hyperactivity, and cognitive problems/inattention. One important conceptual weakness of Pimperton and Nation's attempt to measure relationships between reading skill and behavioral problems was a nonexclusive nature of scale measures. In particular, hyperactivity was measured twice: first as hyperactivity and second as part of attention deficit/hyperactivity. Inattention was measured twice: first through attention deficit/hyperactivity and second through cognitive problems/inattention. Overlaps between variables lowered both reliability and validity of behavior-related findings.

Pimperton and Nation (2014) concluded, according to teacher ratings, poor readers and adequate readers were not statistically dissimilar to each other regarding oppositional behavior, hyperactivity, or attention deficit/hyperactivity. Significance of subscale outcomes should be considered to understand why relationships between behavioral issues and poor academic performance were linked. The only statistically significant difference identified by Pimperton and Nation was within cognitive problems/inattention domains, which was broken down into a subscale to measure specific problems of cognition and attention. Pimperton and Nation found poor readers were rated by teachers as being (a) less likely to remember things already been

learned and (b) more likely to lack interest in schoolwork. Findings conclude poor readers were more likely to experience cognitive difficulties, rather than behavioral problems.

Garwood, Varghese, and Vernon-Feagans (2017) conducted a study on relationships between reading skills, behavior, and gender. Garwood et al.'s (2017) study was based on a sample of 472 readers in either Grade 1 or kindergarten, 236 of whom were identified as struggling readers, and 236 of whom were identified as non-struggling readers. Reading achievement levels for the sample was measured through an overall standardized reading test, as well as a writing test administered by school staff; each of these two tests was administered twice a year. Garwood et al. (2017) were interested in three kinds of behavior, one of which, hyperactivity/inattention, was directly relevant to adverse behavioral and disciplinary outcomes. Garwood et al. found a statistically significant ($p < .001$) difference between hyperactivity/inattention levels of struggling readers ($M = 4.39$, $SD = 2.81$) and non-struggling readers ($M = 2.52$, $SD = 2.38$), with effect size of the difference, d, measured as 0.64, which Garwood et al. described as a large effect.

Garwood et al. (2017) found gender neither mediated nor moderated relationship between reading skill and hyperactivity/inattention. Garwood et al. concluded poor readers were likely to experience adverse disciplinary outcomes in early grades because of individuals' hyperactivity and inattention. An important limitation of Garwood et al.'s study was no forms of dysfunctional behavior, other than hyperactivity/inattention, were measured. In addition, without disciplinary referrals or another objective measure of discipline as an outcome variable, one might not reach the same conclusion as Garwood et al.: specific deficits in behavior might be reflected in higher levels or likelihoods of disciplinary referrals.

Gaps in Literature

The purpose of this section is to discuss existence of a gap or gaps most relevant to the study. An overall gap is the apparent absence of recent empirical studies modeling the outcome variable of student-level disciplinary referral solely as a function of reading performance. Of seven studies closely discussed in literature review, five did not operationally define an outcome variable as disciplinary referrals, but rather as measures of behavior (Chung, 2015; Garwood et al., 2017; King et al., 2018; Mundy et al., 2017; Pimperton & Nation, 2014). While measures of behavior conceptually and practically lead to outcomes, such as disciplinary referrals, measures of behavior are not operationally defined as disciplinary referrals (Chung, 2015; Garwood et al., 2017; King et al., 2018; Mundy et al., 2017; Pimperton & Nation, 2014). Only two studies included disciplinary referrals as dependent variables (Martorell & Mariano, 2018; Özek, 2015). Of these studies, Özek's (2015) was based on aggregated, school-level data.

Özek's findings could not be applied in educational settings where principals, vice principals, teachers, social workers, and other stakeholders in student discipline would like to know likelihood of a particular academic outcome. Outcomes not specifically identified, yet thought to be beneficial, include a specific score on a reading assessment, relating to odds of a particular student, possibly identified as a male of color, undergoing a particular disciplinary outcome such as suspension (Özek, 2015). Although Martorell and Mariano (2018) treated disciplinary outcomes as dependent variables and drew on students for a level of analysis, Martorell and Mariano's use of linear models rather than odds ratios, as generated through logistic regression, resulted in a similar lack of practical applicability in actual education settings.

In a substantial cross-section of literature on educational leadership and school discipline in United States, there is an emerging consensus regarding addressing student discipline. Student disciplinary decision-makers are best served by trying to understand, respond to, and if possible, even predict and preempt, specific disciplinary circumstances applying to a student, group of students, or a school as a whole (Bottiani, Bradshaw, & Mendelson, 2017; DeMatthews, 2016; DeMatthews, Carey, Olivarez, & Moussavi Saeedi, 2017; Heilbrun, Cornell, & Lovegrove, 2015; Hoffman, 2014; Payne & Welch, 2015; Slate, Gray, & Jones, 2016). In this context, an ideal scenario is one in which school principals or other stakeholders in school discipline can assess a likelihood of a particular student being at a level of risk for a disciplinary infraction (Bottiani et al., 2017; DeMatthews, 2016; DeMatthews et al., 2017; Heilbrun et al., 2015; Hoffman, 2014; Payne & Welch, 2015; Slate et al., 2016). Such an ideal scenario is difficult, if even possible, to achieve, but there is evidence suggesting using data in guiding various aspects of disciplinary decision-making can be useful (Bottiani et al., 2017; DeMatthews, 2016; DeMatthews et al., 2017; Heilbrun et al., 2015; Hoffman, 2014; Payne & Welch, 2015; Slate et al., 2016). In this context, one overarching critique of empirical literature was none of the studies' models could generate an assessment of a likelihood of a particular student being at a particular level of risk for a disciplinary infraction based on reading performance, which was one of two universal performance measures available to American educational decision-makers at primary and secondary levels (Chung, 2015; Croft, Roberts, & Stenhouse, 2015; Garwood et al., 2017; Glover, Reddy, Kettler, Kurz, & Lekwa, 2016; King et al., 2018; Lewis & Hardy, 2015; Martorell & Mariano, 2018; Mundy et al., 2017; Özek, 2015; Pimperton & Nation, 2014; Plank & Condliffe, 2013; Singh, Märtsin, & Glasswell, 2015; Witte, Wolf, Cowen, Carlson, & Fleming, 2014).

Methodology

Examining methodology of existing literature proved beneficial. In this context, an important gap was scarcity of models based on students as a unit of analysis. Martorell and Mariano (2018) and Özek (2015) did not define students as a unit of analysis. When schools were units of analysis, principals and other educational personnel were not necessarily equipped with insight to apply to students fitting certain profiles based on demographics, reading score, and other variables. Gaps remain in environmental and preparedness senses. Preparedness was not a measure used in these studies, which reflected a gap in understanding results.

Gaps in studies existed when defining students as units of analysis (Chung, 2015; Garwood et al., 2017; King et al., 2018; Mundy et al., 2017; Pimperton & Nation, 2014). One such gap is based on possible insufficiency of linear regression and related statistical models. A linear model presumes 1-unit changes in a predictor variable or variables align with *n*-unit changes in a dependent variable, with scatterplot of these relationships taking a roughly linear form (García, García, López Martín, & Salmerón, 2015). Relationships between reading performance and discipline-related outcomes may not be linear. For example, there may be no effect of reading performance on discipline-related outcome until a critical threshold value for a low reading score. In such cases, linear regression models may be inferior to quantile regression models, hurdle regression models, or other nonlinear regression models (Aziz, Mahomed, & Mason, 2016; Indera, Yassin, Zabidi, & Rizman, 2017; Ma, Yan, Wei, & Wang, 2016; Taillardat, Mestre, Zamo, & Naveau, 2016).

Even if researchers applied regression models accommodating various types of relationships between reading performance and discipline-related outcomes, regression models might either be unsuited to or unable to drive actional information from data sources. With an

exception of logistic family of models, regression models typically require dependent variable in any analysis to be operationally defined as a continuous variable (Hong, Pradhan, Xu, & Bui, 2015; Li & Valliant, 2015). Regarding discipline-related outcomes, this methodological requirement of linear regression requires such outcomes to be defined as rates, raw counts, or continuous variables (Hong et al., 2015). Disciplinary distinctions between students might be better understood through dichotomizing the dependent variable of disciplinary referrals.

This particular gap in empirical literature with students as unit of analysis can be illustrated through zero-tolerance policies (Hoffman, 2014). In a zero-tolerance disciplinary environment, success is considered with prevention of any disciplinary events; in addition, a distinction between one and several disciplinary incidents loses meaning, as students may be expelled long before acquisition of several referrals (Gage et al., 2013; Hoffman, 2014; Skiba, 2014). In such an environment, principals and other school personnel may be more interested in the question of how to prevent any disciplinary referral from taking place, which requires a dependent variable of disciplinary referrals to be dichotomized (e.g., as *0 = no disciplinary referrals* or *1 = more disciplinary referrals*) referrals (Gage et al., 2013; Skiba, 2014). The dichotomization of disciplinary referral's variable will commit researchers to different statistical methods of analysis, such as using logistic regression with odds ratios referrals (Gage et al., 2013; Hoffman, 2014; Skiba, 2014). None of the student-based empirical studies appeared to take this approach (Chung, 2015; Garwood et al., 2017; King et al., 2018; Mundy et al., 2017; Pimperton & Nation, 2014). Chung (2015) reported an odds ratio (OR), but Chung's dependent variable was reading, categorized as poor reader versus adequate reader, not discipline or behavior.

Chapter Summary

The purposes of literature review were to (a) discuss theories relevant to posited relationship between reading level and disciplinary referrals, (b) discuss and synthesize empirical findings relevant to relationships between reading level and disciplinary referrals, and (c) identify gaps in existing literature base. The theoretical framework for the current study was identified as neuromaturation and was chosen because of the theory's explanation of existence and conceptual basis of a link between reading performance and discipline-meriting student behavior (Berninger et al., 2017; Chávez et al., 2016; Follmer, 2018; Friedman & Miyake, 2017; Gesell, 1925).

A synthesis of empirical research studies indicated a strong likelihood of a correlation between an extent of a student's reading performance and likelihood of subsequent discipline-meriting behaviors and outcomes (Chung, 2015; Garwood et al., 2017; King et al., 2018; Martorell & Mariano, 2018; Mundy et al., 2017; Özek, 2015; Pimperton & Nation, 2014). Only Martorell and Mariano (2018) failed to identify existence of a relationships between reading skill and behavioral or disciplinary outcomes. The existence of a high degree of consensus in literature strongly indicated poor readers were disproportionately likely to have behavioral and disciplinary problems. Regarding literature gaps related to statistical models in particular, relationships appear to have been understudied regarding (a) scarcity or absence of studies having applied quantile regression, hurdle regression, and other nonlinear regression models capable of accounting for a possibly nonlinear relationship between reading performance and discipline-related outcomes; and (b) scarcity or absence of logistic regression and other methods which could be useful to educational decision-makers who are interested in elimination of any disciplinary incidents which would require a dependent variable of disciplinary referrals to be

dichotomized. Chapter 3 will address the study's methodology and address statistical gaps in empirical literature.

The literature indicated a likelihood of a direct relationship between lower reading performance and higher levels of behavioral problems and disciplinary infractions (Chung, 2015; Garwood et al., 2017; King et al., 2018; Martorell & Mariano, 2018; Mundy et al., 2017; Özek, 2015; Pimperton & Nation, 2014). Few studies appeared to have tested disciplinary infractions as a function of reading performance. The only two recent researchers of this topic did not identify directly comparable empirical studies: Martorell and Mariano (2018), who provided an extensive literature review of both seminal and recent literature, suggested Özek (2015) was the first scholar to have applied statistical methods to quantify disciplinary infractions as a function of reading performance. There were few recent empirical studies on the topic.

Özek's (2015) findings were based on aggregated, school-level data. Martorell and Mariano's (2018) linear regression approach failed to treat discipline as a dichotomous event subject to OR generation. Neither of these two studies could provide what others have identified as a key need of disciplinary stakeholders: being able to generate an assessment of a likelihood of a particular student being at a particular level of risk for a particular disciplinary infraction based on reading performance or another predictor of interest in a given educational setting (Bottiani et al., 2017; DeMatthews, 2016; DeMatthews et al., 2017; Heilbrun et al., 2015; Hoffman, 2014; Payne & Welch, 2015; Slate et al., 2016). In Chapter 3, methodology is described and defined to indicate ways such a study can be designed and executed, thereby providing disciplinary stakeholders with a more practical model of disciplinary forecasting and management than exists in recent and relevant literature (Chung, 2015; Garwood et al., 2017;

King et al., 2018; Martorell & Mariano, 2018; Mundy et al., 2017; Özek, 2015; Pimperton & Nation, 2014).

Chapter 3: Methodology

Students who are behind or struggling in mastering reading abilities may face academic challenges and may engage in disruptive behaviors (Bruhn & Watt, 2012). Kane County Regional Office of Education (2016) reported any student who is two or more years behind in reading ability is at risk of truancy, time out of class, and dropping out of school. Teachers often use behavior modification techniques to address problematic behaviors, but fail to investigate reading-specific academic deficits manifesting as poor performance on standardized tests and lower grades (Oliver et al., 2011).

The purpose of this quantitative investigation was to investigate the relationship between seventh-grade reading scores on standardized tests and office discipline referrals and indicators of grade retention among seventh grade students in Northwest Suburban Chicagoland school district. As such, the main research question to be addressed in this current study was: What degree of correlation, of any, exists between seventh-grade standardized reading assessment scores from Illinois Standards Achievement Test (ISAT) and future occurrences of office discipline referrals during students freshman and sophomore years in Northwest Suburban Chicagoland school district? The null hypothesis to be tested was: There is no significant correlation between seventh-grade standardized reading assessment scores from Illinois Standards Achievement Test (ISAT) and future occurrences of office discipline referrals during students freshman and sophomore years in Northwest Suburban Chicagoland school district.

The predictor variable was reading scores on standardized tests while the criterion variables are office discipline referrals and grade retention occurrence. The predictive analysis between variables spanned three academic years: 2012-2013, 2014-2015, and 2015-2016. Reading scores of a cohort of seventh-grade students were gathered from the 2012–2013 school

year. Office discipline referrals and grade retention occurrence of the cohort were gathered from the 2014–2015 and 2015–2016 school years.

Chapter 3 presents details of the current study's methodology. The rationale for choosing a quantitative method with correlational research design is discussed first. A discussion about target population and sample selection is presented next. Data collection procedures, ethical issues, data analysis plan, and reliability and validity are discussed thereafter. A summary of important details of the proposed methodology conclude Chapter 3.

Research Design and Rationale

A quantitative method was employed for the current study. Quantitative methods have been described as a research methodology requiring the use of mathematical techniques to provide statistical inferences about relationships or differences on numerically measured variables (Camm, 2012; Hancock & Mueller, 2010; Wisniewski, 2016). Quantitative methodology is normally used on studies having research questions pertaining to 'who,' 'what,' and 'how many' (Leavy, 2017). The purpose of this quantitative investigation was to investigate the relationship between seventh-grade reading scores on standardized tests and office discipline referrals and indicators of grade retention among seventh grade students in Northwest Suburban Chicagoland school district.

The study's research questions and hypotheses are directed towards determining a predictive relationship between predictor and criterion variables. The predictor variable, reading scores on standardized tests, and criterion variables, office discipline referrals and grade retention occurrence were both measured numerically as gathered from student records kept by selected middle schools. Correlation and regression analysis were used to test all identified

hypotheses. Based on all aforementioned considerations, a quantitative method was appropriate for the current study.

A qualitative method or mixed method was not appropriate for the study. In qualitative studies, interviews, observations, and case studies are used to gather information about a certain phenomenon from identified individuals or group of people under study (Barczak, 2015; Park & Park, 2016). Qualitative methods make use of inductive logic to arrive with explanations and insights from different sources of information such as interview transcripts, recordings, documents, case studies, and/or observations (Barczak, 2015; Park & Park, 2016). Qualitative analysis emphasizes to answer "how" and "why" questions and subsequent interpretation of data as collected in natural circumstances (Peters & Halcomb, 2015). A mixed method is a methodology involving collecting, analyzing, and integrating of quantitative and qualitative techniques (Halcomb & Hickman, 2015; Terrell, 2012). A mixed method study makes use of qualitative analysis to support quantitative results. Data in this study were collected from school records of students to investigate relationships between variables without a need to interview or gather experiences, opinions, and qualitative observations from participants (Johnson & Christensen, 2012). Qualitative and mixed method approaches would not be as effective as the quantitative method in addressing research questions and hypotheses of the current study.

A correlational research design was employed for this study. Correlational research design seeks to determine relationships between numerically measured variables (Curtis, Comiskey, & Dempsey, 2016; Goodwin & Goodwin, 2013). The use of correlational research design provides opportunity to evaluate both the magnitude and behavior of relationships between variables (Leedy & Ormrod, 2012; Whitley, Kite, & Adams, 2013). Through use of correlational research design, relationships between reading scores (predictor) and office

discipline referrals and grade retention occurrence (criterion) variables by correlation analysis was determined and addressed research questions and hypotheses of this study. A correlational research design suits the objective of this current study.

Other research designs such as causal comparative and experimental were deemed to be inappropriate for the study. A causal-comparative research design primarily aims in explaining differences of means of a dependent variable across two or more groups (Babbie, 2013; Rottman & Hastie, 2014). This current study only focused on one group of participants tracked over three years. An experimental approach was not be appropriate for the study as experimental approach uses hypothesis (or several hypotheses) affirming whether or not a treatment or experiment affect a variable or variables (Babbie, 2013; Hoe & Hoare, 2012). The study did not conduct any treatment or experiment with selected participants and only focused on existing student characteristics. Causal comparative and experimental research designs were inappropriate for the objective of this study.

Research Procedures

In any research, clear and detailed research procedures are important to effect replicability of the study. The following subsections discuss research procedures for the current study. Specifically, population and sample selection were detailed. Instruments used to gather data and data collection procedures are presented along with an explanation of data preparation procedures.

Population and Sample Selection

The target population for this study were seventh-grade students in Northwest Suburban Chicagoland school district. The high school selected for data collection was a high school located within Northwest Suburban Chicagoland school district. Northwest Suburban

Chicagoland school district is ranked among top 10 largest school districts in Illinois, based on student enrollment, and reports 42% of students are considered low-income compared to 50% of student state-wide (U.S. Department of Education, n.d.). Student participants were randomly chosen from existing student data of three selected middle schools to ensure equal representation among all student cohort participants. All students are identified with a number and were selected randomly using Microsoft Excel's random number generator. To be eligible for this current study, student participants must have been continually enrolled in the same middle school starting in seventh grade, continuing enrollment through tenth grade in high school.

A priori power analysis was conducted to determine required minimum sample size for the current study. Four factors were considered in the power analysis: significance level, effect size, power of test, and statistical technique. The significance level, known as Type I error, refers to the chance of rejecting a null hypothesis given the null hypothesis is true (Haas, 2012). Most quantitative studies make use of a 95% significance level because 95% significance level adequately provides enough statistical evidence of a test (Creswell, 2013).

The effect size refers to an estimated measurement of the relationship between variables being considered (Cohen, 1988). Cohen (1988) categorized effect size into small, medium, and large. Berger, Bayarri, and Pericchi (2013) purported a medium effect size is better by striking balance between being too strict (small) and too lenient (large). The power of test refers to probability of correctly rejecting a null hypothesis (Sullivan & Feinn, 2012). In most quantitative studies, an 80% power is usually used (Sullivan & Feinn, 2012).

The statistical test used for this study was correlation analysis. Using G*Power 3.1.9.2 (Faul et al., 2009), the computed required minimum sample size with a 95% significance level, medium effect size, 80% power of test, and correlation analysis as the statistical test is 84 (see

Appendix A). In order to account for missing data and number of possible participants available for recruitment, a total of 100 student records were retrieved.

Instrumentation

The main data collection instruments or sources used were Illinois Standards Achievement Test (ISAT) Reading assessment scores, discipline records, and grade retention occurrence of student participants. Illinois Standards Achievement Test Reading assessment is a state assessment test administered by schools to measure reading achievement of students. Illinois State Board of Education (ISBE) published reliability and validity of ISAT Reading assessment in the 2013 Technical Manual. Feldt and Brennan (as cited in Illinois State Board of Education [ISBE], 2013a) described the coefficient as a range from 0.00 to 1.00 with a highly reliable interpretation occurring with coefficients at or above 0.90. Illinois State Board of Education's ISAT seventh-grade reading assessment has a coefficient of 0.90 and is regarded as being highly reliable (ISBE, 2013a).

Kane (2013) described "validation of test score interpretation and use" as gathering of evidence by different means with the purpose of examining various aspects of an assessment to determine the degree or success of the assessment to measure what the assessment was intended to (p. 5). Illinois Standards Achievement Test's validity is supported and interpreted by an examination of components supporting three recognized validity rules: content validity, construct validity, and criterion-related validity (ISBE, 2013a; Kane, 2013). Discipline records of students were formally requested from school district administrators. Validity and reliability of discipline records rely solely on data collection, reporting, and storage methods within Northwest Suburban Chicagoland school district.

Data Collection

Data were collected from school district databases and school records of seventh-grade students. Four types of data were collected: reading scores on standardized tests, office discipline referrals, grade retention status, and demographical information. This information is confidential and formal permission were requested from school district administrators before formal retrieval and use of requested data. Data were requested by formal submission of Application to Conduct Research administrative procedures form (see appendix B).

Reading scores from standardized tests were collected in the form of a scaled score from 2012-2013 school year's administration of ISAT Reading assessment. During 2012-2013 school year, student participants were in the seventh grade cohort. Each score was examined and categorized as outlined by ISBE as: academic warning, below standards, meets standards, or exceeds standards (ISBE, 2013b).

Data were collected in the form of office discipline referral category and occurrence for 2014-2015 and 2015-2016 school years for the same cohort of students. This cohort of students was in ninth and tenth grades respectively. For purposes of data collection, sample size, and to insure consistency, only students who consistently attended Northwest Suburban Chicagoland school district schools for seventh through sophomore year were included.

The demographical information of students was collected to characterize the sample. Demographical information collected included gender and age. All data were in measured categorical form. Gender was categorized into male and female, and age categorized into 11, 12, 13, and 14 years old based on participants' age at time of assessment administration. Frequency and percentages were conducted on collected data.

Data Preparation

Data were requested of the school district by submission of Application to Conduct Research administrative procedures form (see appendix B). Raw data were exported from Northwest Suburban Chicagoland school district's student information management system to Microsoft Excel. After which, data cleaning and screening procedures were conducted to ensure all valid and complete data sets were included in final analysis. Only students with complete records of reading test scores, office discipline referrals, and demographical information were considered. After arriving at a cleaned final data set, data were exported to Statistical Package for the Social Sciences (SPSS).

Data Analysis

Descriptive analysis was conducted first in order to characterize demographics of participants. Descriptive statistics such as frequency, percentage, mean, and standard deviation were computed. Charts and figures were generated to accompany descriptive analysis.

Correlational studies aim to answer research questions pertaining to whether a direct relationship exists between two numerically measured variables (Gardner & Neufeld, 2013). Specifically, bivariate correlation is usually utilized to analyze underlying relationships between two variables (Sedgwick, 2012). Pearson's correlation and Spearman's Rank correlation are among two most common correlation tools (Eisinga, Grotenhuis, & Pelzer, 2013). The former is used when a dataset follows parametric assumptions while the latter is used when a dataset violates parametric assumptions. Pearson's correlation and Spearman's Rank correlation are used to examine both the strength and direction of a linear relationship between two variables. Pearson *r* is more appropriate for explaining positive, negative, or null correlation between two variables based on the direction scatter plot results take as compared to ranking of data (Eisinga

et al., 2013). Spearman's rank correlation is more appropriate for examining variables when data violates parametric assumptions (Sedgwick, 2012). The linear direction is easier to present to a lay audience who may not be familiar with how ranks associated with two variables sheds light on a relationship, should one exist. If data prove to violate assumptions of a parametric analysis, the non-parametric counterpart of Pearson's correlation analysis, which is Spearman's rank correlation analysis, may be utilized (Walls, 2012).

Following are assumptions needing to be tested for parametric tests: normality, linearity, and homogeneity assumptions. The normality assumption assumes distribution of the test is normally distributed with a mean of zero, one standard deviation, and a symmetric bell-shaped curve (Finch, 2005). A normal probability plot was generated to examine if a violation of normality assumption exist. The assumption of linearity indicates the relationship between variables (i.e., the predictor and criterion variables) follows a straight line (Bücher, Dette, & Wieczorek, 2011). A scatter plot with standard regression output was generated to examine if a violation of the linearity assumption exist. The assumption of homoscedasticity refers to equal variance of all values of independent variables around the regression line (Finch, 2005). A residual scatter plot was generated to examine if a violation of linearity assumption exist.

A correlation value indicates the magnitude and behavior of relationships between two variables. A positive correlation value means as one variable increase in value, the other variable increases as well (Sedgwick, 2012). A negative correlation value means as one variable increase in value, the other variable decreases (Sedgwick, 2012). A correlation value of 0.70 and higher is considered high correlation between two variables while a correlation value of 0.30 and below is considered low correlation between two variables (Gardner & Neufeld, 2013).

The generation of statistical analysis is often conducted using statistical analysis software. One of the most recognized software used to generate correlation coefficient analysis is IBM SPSS software (Alversia, 2011). IBM SPSS software provides an advantage in how scatter plot graphs are rendered after data is analyzed (Alversia, 2011). All analysis were conducted using 95% significance level. A *p*-value greater than the significance level indicates the predictor variable (reading scores on standardized tests) has no impact on the criterion variable (office discipline referrals). A *p*-value less than (or equal to) the significance level indicates a significant relationship between variables.

Reliability and Validity

Reliability and validity of this study heavily lies on gathered data. All data for this current study were retrieved from student records gathered and kept by the district's student information management system. Data collected from student records were ISAT Reading assessment scores, office discipline referral event records, and records indicating whether or not participants were retained in grades 9 or 10. Data were not manipulated in any way once provided. Data reliability and validity rested upon the integrity of data collection methods of Northwest Suburban Chicagoland school district practices. Northwest Suburban Chicagoland school district administrators were asked to provide reliability and validity tests or supporting documentation if any were conducted. Data were examined to ensure only current and correct data were retrieved. No manipulation was done with raw data.

Illinois State Board of Education published reliability and validity of ISAT Reading assessment in the 2013 Technical Manual. Feldt and Brennan (as cited in ISBE, 2013a) described coefficient as a range from 0.00 to 1.00 with a highly reliable interpretation occurring with coefficients at or above 0.90. Illinois State Board of Education ISAT seventh-grade reading

assessment had a coefficient of 0.90 and is regarded as being highly reliable (ISBE, 2013a). Illinois Standards Achievement Test assessment's validity is supported and interpreted by an examination of components supporting three recognized validity rules: content validity, construct validity, and criterion-related validity (ISBE, 2013a; Kane, 2013).

Ethical Procedures

This current study followed all ethical procedures outlined in American College of Education's Institutional Review Board (IRB). Approval of American College of Education's IRB was secured before data collection and analysis commenced. Approval on an application to conduct research was requested of Northwest Suburban Chicagoland school district administrators prior to retrieval of any student records for the current study.

All data were anonymous and confidential. No personal identifying information was collected. Pseudo codes, such as Student #1, were used to tag all participants. Hard copies of raw data and other documents pertinent to the current study are securely kept in a locked filing cabinet. Soft copies of raw data and other documents are saved in a password-protected flash drive. All data and documents related to the current study will be destroyed five years after completion. Hard copies will be shredded while soft copies will be deleted.

Chapter Summary

The purpose of this quantitative investigation was to investigate the relationship between seventh-grade reading scores on standardized tests and office discipline referrals and indicators of grade retention among seventh grade students in Northwest Suburban Chicagoland school district. The predictor variable was reading scores on standardized tests while the criterion variables were office discipline referrals and indicators of grade retention among participants. The predictive analysis between variables spanned three academic years. Illinois Standards

Achievement Test reading scores of a cohort of seventh-grade students was gathered from 2012–2013 school year. Office discipline referrals and indicators of grade retention of the same cohort were gathered from two school years after, specifically 2014–2015 and 2015–2016 school years. The target population for this study was seventh-grade students.

A total of 100 seventh-grade student records were retrieved from the student information management system. Both descriptive and inferential statistics were conducted to analyze data. Descriptive statistics included frequency, percentage, mean, and standard deviations. Inferential statistics included correlation analysis – either Pearson's correlation or Spearman's rank correlation. Chapter 4 presents results and Chapter 5 provides discussion and conclusion for the current study.

Chapter 4: Research Findings and Data Analysis Results

Disciplinary responses to students who demonstrate disruptive and inappropriate behavior in school frequently involves removing those students from classroom settings to speak with an administrator and as a restrictive disciplinary action. Student disciplinary-related classroom removals have an effect of reducing exposure to instructional time removed students receive, which may translate into lower academic achievement, and lower academic achievement often leads to further disruptive classroom and student behavior (Oliver et al., 2011). This quantitative correlational study analyzed archival records of a sample of 100 students who were in seventh grade in three middle schools Northwest Suburban Chicagoland school district in 2012-2013 school year and followed those students through ninth and tenth grades of high school. The purpose of the quantitative investigation was to investigate the relationship between seventh-grade reading scores on standardized tests and office discipline referrals and indicators of grade retention among seventh grade students in Northwest Suburban Chicagoland school district. Three research questions were addressed in the study.

Research question 1. What degree of correlation, if any, exists between seventh-grade standardized reading assessment scores from ISAT and future occurrence of office discipline referrals during students' freshman and sophomore years?

Research question 2. What degree of correlation, if any, exists between student office discipline referrals and average to above average standardized reading assessment scores?

Research question 3. What degree of correlation, if any, exists between seventh-grade standardized reading assessment scores from ISAT and grade retention throughout students' freshman and sophomore years in high school?

Chapter 4 describes the process by which data were gathered. The sample is then described in terms of its demographic characteristics, ISAT seventh-grade reading test performance, disciplinary characteristics and grade retention. Chapter 4 further describes analyses used to address the study's research questions, results of those analyses, and answers to the study's research questions. Chapter 4 concludes with a summary.

Data Collection

Upon receipt of approval from Northwest Suburban Chicagoland school district and American College of Education's Institutional Review Board, archived student records were downloaded from Northwest Suburban Chicagoland school district's student information management system student record database. Records were gathered for a random sample of 100 students who were in seventh grade during the 2012-2013 school year, in ninth grade during 2014-2015 school year, and in tenth grade during 2015-2016 school year. All students maintained continuous enrollment in Northwest Suburban Chicagoland school district from seventh through tenth grades. All records were anonymized. Data were retrieved for the following variables: (a) gender, (b) age when enrolled in seventh grade, (c) Illinois Student Achievement Test (ISAT) seventh-grade reading test scores; (d) reading score grades (*academic warning, below standards, meets standards, exceeds standards*); (e) grade retention records (whether retained or not and, if retained, in which grade); (f) number of office disciplinary referrals; and (g) behavioral events leading to office disciplinary referrals.

Data Analysis and Results

Data were imported to IBM Statistical Package for the Social Sciences (SPSS) (Version 25.0) as an SPSS data file and all subsequent data processing and statistical analysis were performed using IBM SPSS software, except calculations of effect size were performed using

G*Power software (Version 3.1.9.2; Faul, Erdfelder, Lang, & Buchner, 2007). G*Power was used for effect size calculations because SPSS does not provide effect size for all statistical analysis, such as Cohen's *d*. Using Cohen's *d* to measure of effect strength and strength of correlational relationships required G*Power because SPSS does not support this analysis as an output (Faul et al., 2007). Prior to performing any analyses, frequency distributions were generated for all variables to check for out-of-range values and other errors (Tabachnick & Fidell, 2013). No errors were found.

Sample Characteristics

The sample contained 100 randomly selected, anonymous student data from three distinct academic years. Standardized reading test data were from 2012-2013 school year when the sample cohort was in seventh-grade. Office discipline referrals and grade retention occurrence data were provided from school years 2014-2015 and 2015-2016, during the cohorts' high school freshman and sophomore years respectively. To be considered for the sample, students must have been continually enrolled in a Northwest Suburban Chicagoland school district school starting with 2012-2013 school year through completion of 2015-2016 school year.

Demographic characteristics. The sample of 100 students consisted of 38 females (38.0%) and 62 males (62.0%). When in the seventh grade, 66 (66.0%) of the students were 12 years of age and 34 (34.0%) were 13 years of age. No programmatic data, such as special education or Section 504 plan eligibility, or non-native English speaker status, was collected or considered as part of the study's sample.

Illinois Standards Achievement Test seventh-grade reading test performance. Scores on the ISAT seventh-grade reading test ranged from 187 to 312, with a mean of 249.46 and standard deviation of 27.14. Since the total sample size was 100 (n = 100), frequency

counts were the same as percentages. The scores provided a close visual approximation to the normal curve, with neither skewness = 0.11 nor kurtosis = -0.19 approaching the values ±1.0 Hair, Black, Babin, and Anderson (2010) have suggested serve as criteria for identifying markedly non-normal distributions. Figure 1 shows a frequency and percentage histogram of ISAT reading test scores.

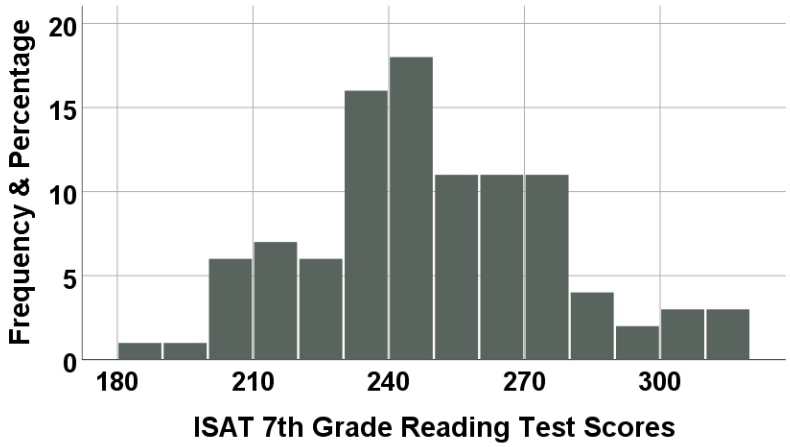


Figure 1. Frequency and percentage histogram of ISAT seventh-grade reading test scores.

The near-normality of distribution was confirmed by a non-significant Shapiro-Wilk test of distribution normality, $SW = 0.99$, $df = 100$, $p = .505$ (Tabachnick & Fidell, 2013). Students' ISAT seventh-grade reading test scores were accompanied by reading performance grades: AW = *academic warning,* BS = *below standards,* MS = *meets standards,* and ES = *exceeds standards*. Table 2 provides counts and percentages of the sample in each reading performance category.

Table 2

Distribution of ISAT Seventh-Grade Reading Test Scores into Reading Performance Categories

Performance Category	*f*	%
Academic Warning	5	5.0%
Below Standards	29	29.0%
Meets Standards	43	43.0%
Exceeds Standards	23	23.0%
Total	100	100.0%

Disciplinary characteristics. A total of 214 behavioral events were recorded resulting in office disciplinary referrals in this sample of 100 students during students' ninth and tenth grade years. Office discipline referrals were relatively rare and 68 (68.0%) students were never referred for disciplinary action. The mean number of referrals associated with any single student in the sample was 2.14 events (SD = 6.62), but the distribution was extremely skewed. The median value of 0.00 provides a better indication of the number of times a typical student in this sample received an office discipline referral because the median is a better indicator of the typical score in a skewed distribution (Tokunaga, 2019). The typical student, in the statistical sense, was never referred. Figure 2 provides a frequency and percentage histogram for office discipline referrals.

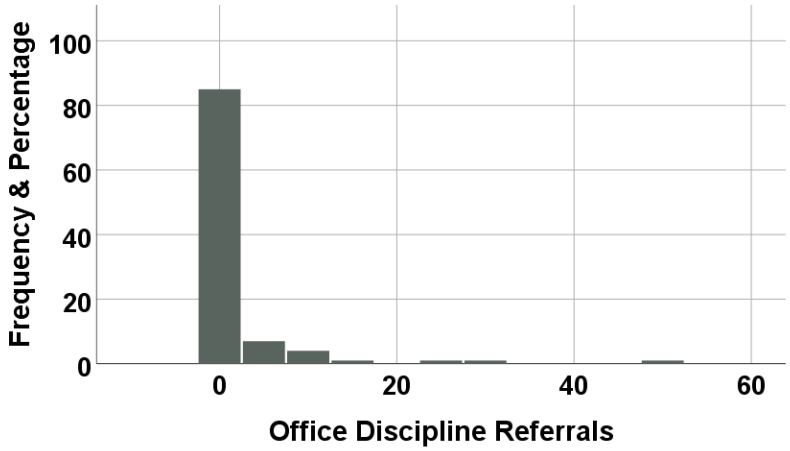


Figure 2. Distribution of office discipline referrals per student in the sample.

A minority of only 32 students (32.0%) were referred for some type of disciplinary action. In this minority subset of students, another minority was responsible for most behavioral events leading to office discipline referrals. Nearly 56.0% of events leading to an office discipline referral were attributed to only four students. The majority of referred students (53.1%) were referred only once or twice. Figure 3 captures the distribution of office discipline referrals in the subset of 32 students who were referred one or more times.

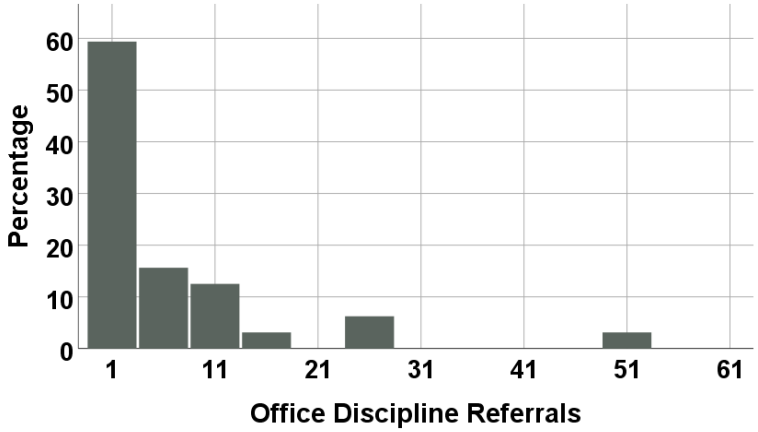


Figure 3. Distribution of office discipline referrals in the subset of 32 students who were referred one or more times.

In addition to tracking behavioral events leading to office discipline referrals, data were available on the nature of these events. Sixteen varieties of referred behavior were reported. Table 3 lists referred behaviors from most frequent to least frequent.

Table 3

Frequency of Behavioral Events Leading to 214 Office Discipline Referrals

Behavior	*f*	%
Class cuts	75	35.0%
Failure to report for detention	50	23.4%
Insubordination-Defiance-Disrespect	22	10.3%
Electronic devices/cell phones	14	6.5%
Tardies	11	5.1%
Assault/Battery/Fighting-Without injury	10	4.7%
Inappropriate behavior	10	4.7%
Parking violation	5	2.3%
Tobacco	3	1.4%
Truancy	3	1.4%
Unauthorized exit from building/grounds	3	1.4%
Drugs	2	0.9%
Forgery	2	0.9%
Inappropriate Conversation	2	0.9%
Cheating/Academic dishonesty	1	0.5%
Profanity	1	0.5%
Total	214	100.0%

Note. Percentages do not sum to 100.0% due to rounding error.

Grade retention. Only four students (4.0%) in this sample of 100 experienced grade retention. The lack of variability on grade retention precluded any correlational analyses involving said variable. A variable which does not vary cannot covary with other variables and variables which cannot covary cannot be correlated (Tokunaga, 2019).

Research Question 1: What degree of correlation, if any, exists between seventh-grade standardized reading assessment scores from ISAT and future occurrence of office discipline referrals during students' freshman and sophomore years?

As previously established, ISAT seventh-grade reading test scores provided a close approximation to the normal curve (Figure 1), and findings show frequency of office discipline referrals was strongly skewed with some high-scoring outliers (Figure 2). Tests of the significance of the Pearson correlation are based on an assumption both variables in the analysis are normally distributed and neither distribution includes outliers such as extremely low or high scores (Tabachnick & Fidell, 2013). Violations of those assumptions cause inflation of the Type I error rate (Bishara & Hittner, 2012).

In contrast, Spearman rank-order correlation is robust to outliers and violations of the normality assumption but does demand relationships between variables be monotonic (Sheskin, 2011). The monotonicity of a relationship between ISAT seventh-grade reading scores and office discipline referrals was evaluated by examining a scatterplot of those variables. The relationship was somewhat more curvilinear ($R^2 = 0.20$) than linear ($R^2 = 0.13$) as assessed using R^2 as a measure of goodness-of-fit, but the relationship did not show any indication of strong non-monotonicity. This study concluded Spearman correlation was appropriate for use in analyzing the relationship between reading scores and discipline referrals. The monotonicity of a relationship between ISAT seventh-grade reading scores and office discipline referrals scatterplot is shown in Figure 4.

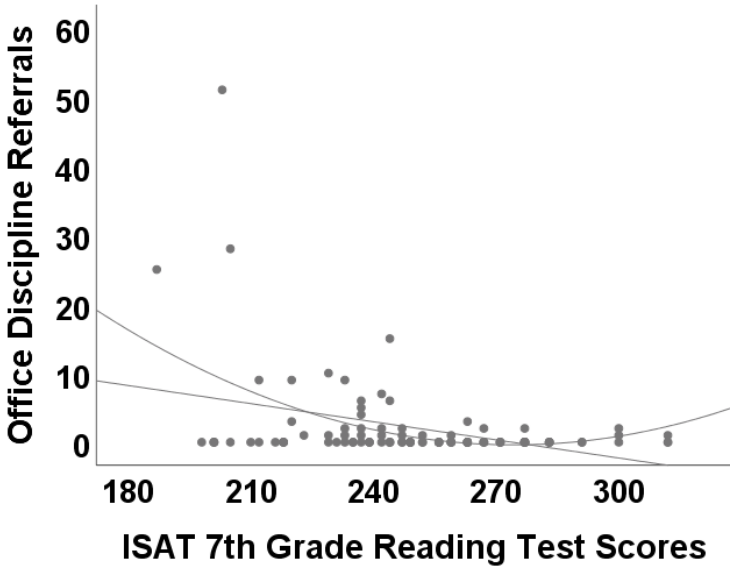


Figure 4. Scatterplot of the relationship between ISAT seventh-grade reading test scores and office discipline referrals.

The Spearman correlation between seventh-grade reading scores and office discipline referrals was of moderate strength by Cohen's (1988) standards, negative, and statistically significant, $\rho(98) = -.29$, $p = .003$ (two-tailed). As reading scores improved, the number of office discipline referrals declined. The null hypothesis ($H_0 1$), there is no significant correlation between seventh-grade standardized reading assessment scores from ISAT and future occurrence of office discipline referrals during students' freshman and sophomore years, was rejected. As is apparent in Figure 4, elevated numbers of office discipline referrals in high school were confined to students whose seventh-grade reading test scores were in the bottom half of the distribution of reading test scores. No students in the top half of ISAT reading test score distribution showed large numbers of discipline referrals.

Research Question 2: What degree of correlation, if any, exists between student office discipline referrals and average to above average standardize reading assessment scores?

Illinois Standards Achievement Test seventh-grade reading test scores were graded into four performance categories. In ascending order, performance categories were *academic warning* (*n* = 5), *below standards* (*n = 29*), *meets standards* (*n* = 43), and *exceeds standards* (*n* = 23). Since the small number of students in the bottom performance category precluded any meaningful analysis of *academic warning* category, the bottom two categories were combined into a single *below average* category (*n* = 34). Table 4 provides descriptive statistics on office discipline referrals (the dependent variable in this analysis) as a function of the independent variable, ISAT seventh reading performance (defined by the three levels of reading performance).

In this study, correlation between office discipline referrals and reading performance was evaluated by determining if a number of discipline referrals differed significantly from one reading performance category to the next. Correlational statistics (e.g., Pearson & Spearman correlations) are not the only statistical procedures which can establish correlational relationships between variables. Significance difference tests, like those used in this study, can be used for said purpose. If groups representing levels of an independent variable (such as three groups representing reading performance in this study) are found to differ significantly on a continuous dependent variable (such as office discipline referrals), an appropriate conclusion among the independent and dependent variables are the appearance of being significantly related or correlated (Gravetter & Forzano, 2016). Statistical measures of strength of those significant differences, i.e., measures of effect strength, like Cohen's *d* statistic used in this study, measure strength of correlational relationships.

Distributions of scores on the office discipline referral dependent variable were strongly skewed and leptokurtic in all reading performance groups, with values of skewness and kurtosis

well beyond the criterion values ±1.0 used to identify non-normal distributions (Hair et al., 2010). Shapiro-Wilk tests of normality were strongly significant in all groups, indicating distributions were strongly non-normal. Both the *below average* and *meets standards* groups contained high scoring outliers, defined in this study as individuals with office referral *z*-scores exceeding ±3.30 ($p < .001$ in a normal distribution; Tabachnick & Fidell, 2013). Further, data variability was much greater in *below average* reading performance than in other groups. These characteristics of data—non-normal distributions containing outliers and groups with heterogeneous variances—precluded the use of analysis of variance (ANOVA) in comparing group means on office discipline referrals. The Kruskal-Wallis *H* test provides a nonparametric alternative to the one-way between-subjects ANOVA and the Kruskal-Wallis procedure is robust to non-normal distributions, outliers, and heterogeneous variances (Corder & Foreman, 2009; Sheskin, 2011).

The Kruskal-Wallis procedure combines scores on the dependent variable from all groups and ranks those scores from lowest (ranked 1) to highest (ranked *N* out of *N* cases in the analysis; Field, 2013). The mean of ranks assigned to participants in each of the groups are then compared. If one group has scored generally higher than another group, the higher scoring group may have generally higher ranks and a higher mean rank. If the groups' distributions on the dependent variable are similarly shaped, Kruskal-Wallis *H* test is considered to be a test of differences between group medians (Conover & Iman, 1979). If the groups' distributions are differently shaped, the procedure is interpreted as a test of differences in the group mean ranks (Sheskin, 2011).

Figure 5 provides frequency line graphs for office discipline referrals for each reading performance group. All three distributions were positively skewed but differed substantially in

dispersion of scores (Table 4). In the *below average* reading performance group, distribution was highly dispersed, with frequency of office discipline referrals ranging from 0 to 51. Frequency of office discipline referrals by reading level are displayed in Figure 5.

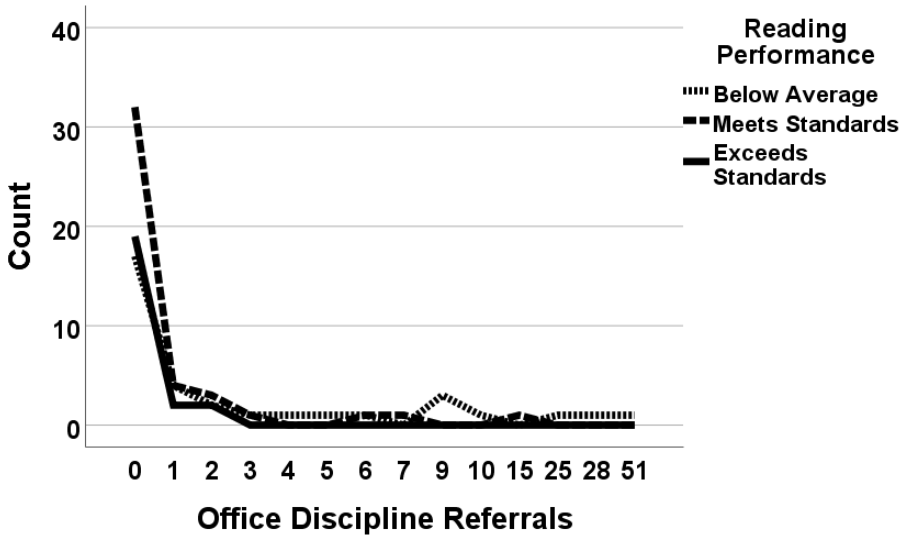


Figure 5. Counts of office discipline referrals in three levels of seventh-grade reading performance as measured by ISAT.

In the *meets standards* group, data dispersion was less, with office referrals ranging from 0 to 15 referrals. In *exceeds standards* group, office referrals showed little dispersion and ranged only from 0 to 2. Important differences in data distributions of three groups, particularly involving data dispersion, were concluded from this comparison of the groups' distributions (Hart, 2001). The Kruskal-Wallis *H* test was treated in this study as a method of comparing the groups' mean ranks as shown in Table 4.

Table 4

Office Discipline Referrals as a Function of Seventh-Grade Reading Performance Groups

								Shapiro-Wilk Tests		
Reading Performance	*n*	*M*	*Mdn*	*SD*	Mean Ranks	Skewness	Kurtosis	*SW*	*df*	*p*
Below Average	34	4.91	0.50	10.49	60.99	3.25	11.64	0.53	34	<.001
Meets Standards	43	0.95	0.00	2.65	46.74	4.15	19.44	0.41	43	<.001
Exceeds Standards	23	0.26	0.00	0.62	42.02	2.31	4.26	0.48	23	<.001

Note. The mean rank for any given group will differ between Table 4 and Table 5 because the analyses reported in the two tables are based on different total sample sizes (*N*). In this table, mean ranks are based on a total $N = 100$ cases. In Table 5, which reports pairwise comparisons, there are fewer cases in each analysis and so mean ranks are lower.

The *H* test was statistically significant, $H(2) = 10.42$, $p = .005$, indicating groups' mean ranks on the office referral dependent variable differed significantly as a function of the independent variable, reading performance. Another way of expressing this finding is to say the number of office discipline referrals students received in high school was significantly related to (or correlated with) students' reading performance category in seventh grade. The null hypothesis ($H_0 2$), there is no significant correlation between student office discipline referrals and average to above average standardized reading assessment scores, was rejected.

Three post hoc pairwise comparisons were performed to identify the source(s) of the significant *H* test by determining which reading performance groups differed significantly on office discipline referrals. Mann-Whitney *U* tests were used in performing these comparisons as recommended by Dunn (1964). The significance levels shown in Table 5 are two-tailed and were adjusted using a Bonferroni correction to hold the familywise Type I error probability at .05 (i.e., obtained significance levels from the Mann-Whitney tests were multiplied by 3—the number of pairwise comparisons; Tabachnick & Fidell, 2013). The series of comparisons

showed students in the *below average* reading performance group (*M* = 4.91, *Mdn* = 0.50, *SD* = 10.49) were referred for disciplinary matters significantly more often than students in the *meets standards* (*M* = 0.95, *Mdn* = 0.00, *SD* = 2.65) or *exceeds standards* groups (*M* = 0.26, *Mdn* = 0.00, *SD* = 0.62). The results of the post hoc comparisons are summarized in Table 5.

Table 5

Results of Bonferroni-Adjusted Mann-Whitney U Tests Used to Compare Office Discipline Referral Medians Among Reading Performance Groups

Groups Compared	First Group Mean Ranks	Second Group Mean Ranks	Mann-Whitney *U*	Mann-Whitney *z*	Mann-Whitney *p*
Below Average (*n* = 34) vs. Meets Standards (*n* = 43)	45.14	34.14	522.00	-2.49	.039
Below Average (*n* = 34) vs. Exceeds Standards (*n* = 23)	33.34	22.59	243.50	-2.78	.015
Meets Standards (*n* = 43) vs. Exceeds Standards (*n* = 23)	34.60	31.43	447.00	-0.87	.999

Note. Significance levels (*p*) are two-tailed and have been adjusted using the Bonferroni procedure so the α = .05 across the series of three comparisons. Any given group's mean rank will be different from one comparison to the next because the numbers of cases in the three comparisons differed (*N* = 77, *N* = 57, and *N* = 65, respectively. The group mean ranks in this table differ from those in Table 4 for the same reason, i.e., the total *N* in Table 4 was 100.

There was no significant difference in the number of office discipline referrals received by students in those latter two groups. Measured using Cohen's *d* statistic, and evaluated by standards proposed by Cohen (1988), the strength of the difference between *below average* group and *meets standards* group was moderate, *d* = 0.52. The strength of the difference between the *below average* group and the *exceeds standards* group was moderate, *d* = .63. Figure 6's graph helps visualize the general nature among differences in office discipline referrals in those groups, but worth noting is neither Kruskal-Wallis *H* nor Mann-Whitney *U* tested differences between means, but rather, differences between groups' mean ranks. Figure 6

graphs the mean number of office referrals among *below average, meets standards,* and *exceeds standards* reading performance groups.

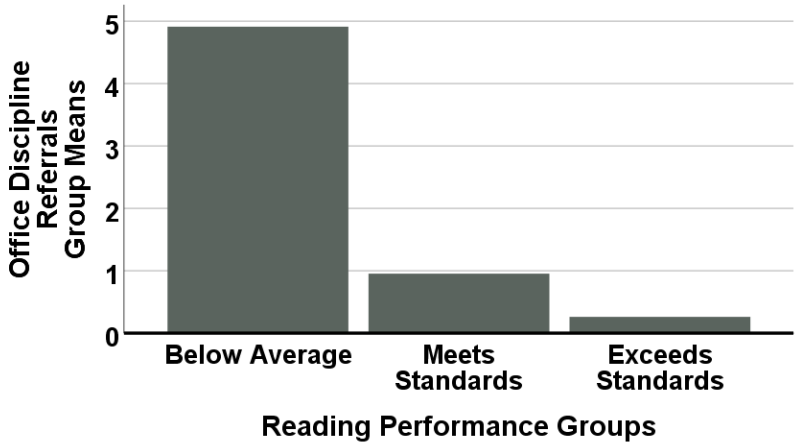


Figure 6. Mean number of office discipline referrals among *below average, meets standards,* and *exceeds standards* reading performance groups.

Research Question 3: What degree of correlation, if any, exists between seventh-grade standardized reading assessment scores from ISAT and grade retention throughout students' freshman and sophomore years in high school? As described previously, the variable of grade retention was virtually a constant in this study, with only 4 (4%) of 100 students in the sample having been retained a grade. As was noted previously, a variable which does not vary cannot covary with other variables, and variables which cannot covary cannot be correlated. Consequently, a correlational evaluation between ISAT seventh-grade reading test scores and the occurrence of grade retention in ninth and tenth grades was impossible in this study. A descriptive profile of students who were grade retained in high school, including information about students' performance on ISAT seventh-grade reading test was generated. The descriptive profile is summarized in Table 6.

Table 6

Characteristics of Grade Retained Students

Case ID	Gender	Grade Retained	ISAT 7th Grade Reading Score	ISAT 7th Grade Reading Score (*z*)	Reading Performance Category	Discipline Referrals	Discipline Referrals (*z*)
16	male	10	229	-0.75	below standards	10	1.19
23	male	9	247	-0.09	meets standards	0	-0.32
93	male	10	203	-1.71	below standards	51	7.38
97	male	10	244	-0.20	meets standards	15	1.94

All 4 grade retained students (100%) were males in comparison to a sample in which 62.0% were males and 38% were females. Three retained students were retained in tenth grade and one was retained in ninth grade. Illinois Standards Achievement Test seventh-grade reading scores of retained students were all below average, but students' standardized scores on ISAT reading test (*z*-scores) showed only one student's score was remarkably below average (the raw score of 203 had a *z*-score equivalent of -1.71). Two grade retained students were classified on ISAT reading test as *below standards*, and two were classified as *meets standards.* None of the retained students' reading test scores were low enough as to have caused students to be classified as *academic warning.* In fact, none of the five students whose seventh-grade reading test scores were classified as *academic warning* subsequently experienced grade retention, at least not in ninth or tenth grades. The median number of disciplinary referrals in the sample of 100 students was zero. In comparison, one retained student had no record of office discipline referrals, but the other three were all well above average with regarding to office discipline referrals as shown by standardized (*z*-score) office discipline referral scores which were above +1.00.

Reliability and Validity

Data analyzed were existing data provided by a Northwest Suburban Chicagoland school district's student information management system. Illinois Student Achievement Test was not created specifically for this study but collected and analyzed student assessment result data is considered both reliable and valid (ISBE, 2013a). Office discipline referral and grade retention occurrence data were provided directly from the school district's student information management system. Threats or potential threats to reliability and validity of data were not detected or evident.

Reliability

Results of any research study are only as reliable as analyzed data. An unreliable measuring procedure can produce two different scores on two different occasions even when the attribute being measured is unchanged. Consequently, a study using unreliable measures can produce two different findings on two different occasions, even when facts have remained unchanged. If a study using unreliable measures were to happen, determining which, if either, of the findings was accurate would be challenging. Reliability and validity work hand in glove; a measuring procedure should be reliable in order to be valid. Quinino, Reis, and Bessegato (2013) described a reliable instrument as one in which most variability in scores reflects variability in the attribute being measured and little score variability is related to other attributes or is random. Variability of scores of an unreliable measuring procedure is unrelated to the intended attribute; consequently, an unreliable measuring procedure does not provide a valid measure of the intended attribute.

In this study, reliability of analyzed data was assumed to be quite high based upon school district policies and procedures provided clear governance on matters of student discipline, grade

retention, standardized testing, and record keeping. Data analyzed in this study were retrieved directly from archived student records kept by Northwest Suburban Chicagoland school district's student information management system. No data were eliminated or altered subsequent to the download. The reliability of ISAT reading assessment is addressed in the 2013 Technical Manual published by Illinois State Board of Education (ISBE, 2013a). As measured by Cronbach's alpha coefficient, reliability of ISAT's seventh-grade reading test was reported by those authors as $\alpha = 0.90$, considered to be an excellent level of reliability by commonly accepted standards (Miller & Lovler, 2016).

Validity

While reliability of a measure is essential to said measure's validity, there is more to validity than reliability. The event for a measuring procedure to reliably measure an attribute other than what is intended is quite possible. Accordingly, the measure is reliable but not valid. In the current study, reliability of measures provided a foundation for validity, along with additional guarantors of validity of ISAT seventh-grade reading test. These validity measures came in the form of assessments of content validity, construct validity, and criterion-related validity as described in the 2013 Technical Manual of the ISAT (ISBE, 2013a). Validity of data on student disciplinary actions and grade retention are assumed to result from detailed, clearly stated school district policies and procedures governing said matters and student record keeping.

The validity of measures is not the only validity issue in research. There is a matter of external validity, defined as the degree to which the findings from a study can be generalized to other situations, including other people, places, and times (Gravetter & Forzano, 2016). The use in this study of a quantitative methodology employing inferential statistical data analyses contributed importantly to the study's external validity. The most fundamental meaning of a

statistically significant sample finding is the finding is large or strong enough to be considered reliable and replicable beyond the sample at hand (Tokunaga, 2019). In other words, a statistically significant sample finding is a finding one can generalize to other samples drawn from the same population.

In this study, students' ISAT seventh-grade reading scores were found to be significantly correlated with numbers of office discipline referrals in ninth and tenth grades (Research Question 1). The statistical significance of the correlation means it was not just a fluke of the particular examined sample, but rather, is a characteristic of the larger population from which the sample was drawn. This being the case, one would expect the correlation to generalize well to other samples from the same population. This study found the number of office discipline referrals students received in high school was significantly related to students' reading performance category in seventh grade (Research Question 2). Statistical significance of the finding means observed relationships may not be confined to the particular sample selected for examination in this study, but is a characteristic of the larger population from which the sample was drawn. This being the case, one would expect the same variable relationship to generalize well to other samples beyond the one observed in this study; statistical significance speaks directly to external validity. The small number of students in this study who were grade retained prevented analyses of using said variable involving any inferential statistical procedures. As a result, findings related to grade retention in this study cannot be assumed to generalize beyond the observed sample.

Evaluating external validity of this study is a matter of identifying the population to which statistically significant sample findings can be generalized—which students, in which places, at which times. This study was limited to students who were enrolled in four schools in

Northwest Suburban Chicagoland school district. Strictly speaking, findings of this study can be generalized only to those schools Northwest Suburban Chicagoland school district. To the degree other students in other schools in other school districts are similar to those examined in this study, one would expect to observe the same findings seen in this study.

Students in this study were examined at specific times in students' lives. Seventh-grade reading scores were collected to see if those reading scores were related to subsequent grade retentions and discipline office referrals in ninth and tenth grades. Variables were related and significant findings can be generalized to other students in seventh, ninth, and tenth grades in Northwest Suburban Chicagoland school district. A conclusion cannot be determined from this study relating to reading scores in grades other than seventh grade are related to grade retention and disciplinary referrals in other grades than those examined in this study.

Chapter Summary

This quantitative correlational study of school records contained 100 students' records randomly sampled from archived student data of a large school district in suburban Chicago. Students examined in this study were 38 (38.0%) female and 62 (62.0%) male. Reading performance in seventh grade was assessed using ISAT seventh-grade reading test. Scores were normally distributed in the sample, with a mean of 249.46 and standard deviation of 27.14. The majority of students (66.0%) met or exceeded state standards for reading, with 29% below standards and 5% receiving an academic warning. In ninth and tenth grades, a minority subset of these students (32.0%) engaged in a total of 214 recorded behavioral events resulting in office discipline referrals. Nearly 56.0% of these referrals were attributed to only four students, and a majority of students with office discipline referrals were referred only once or twice. The

variable of grade retention was found to be a virtual constant, with only four students in the sample having been grade retained.

A Spearman correlation was used to address the study's first research question, which asked what degree of correlation, if any, exists between seventh-grade standardized reading assessment scores from ISAT and future occurrence of office discipline referrals during students' freshman and sophomore years. Spearman correlation was selected for the analysis because of the characteristics of the office discipline referral variable. Specifically, distribution of scores on office discipline referral variable was non-normal and included several extremely high-scoring outliers (Field, 2013; Lehman, O'Rourke, Hatcher, & Stepanski, 2013). The Spearman correlation was negative, of moderate strength, and was statistically significant, $\rho(98) = -.23$, $p = .003$ (two-tailed). The null hypothesis ($H_0 1$), there is no significant correlation between seventh-grade standardized reading assessment scores from ISAT and future occurrence of office discipline referrals during students' freshman and sophomore years, was rejected. As reading scores increased, office referrals declined.

A Kruskal-Wallis *H* test and Bonferroni-corrected post hoc pairwise comparisons using Mann-Whitney *U* tests were used to address the study's second research question (Conover & Iman, 1979). Research question 2 asked what degree of correlation, if any, exists between student office discipline referrals and average to above average standardized reading assessment scores. These nonparametric procedures were used because of the characteristics of the data. Specifically, distributions of scores on the office discipline referral dependent variable were non-normal, included outliers, and showed strong differences in data variability from one reading performance group to the next (the independent variable). Three reading performance groups were compared in the analysis: (a) *below average* (consisting of a combination of the *academic*

warning and *below standards* reading performance groups), (b) *meets standards,* and (c) *exceeds standards.* The Kruskal-Wallis test was statistically significant, $H(2) = 10.42$, $p = .005$, indicating the number of office discipline referrals students received in high school was significantly related to students' reading performance category in seventh grade (Dunn, 1964; Kruskal & Wallis, 1952). The *below average* group had the highest number of office discipline referrals ($M = 4.91$, $Mdn = 0.50$, $SD = 10.49$), followed by *meets standards* group ($M = 0.95$, $Mdn = 0.95$, $SD = 2.65$), with *exceeds standards* group showing the fewest office referrals ($M = 0.26$, $Mdn = 0.00$, $SD = 0.62$). Mann-Whitney *U* tests were used to determine which reading performance groups differed significantly on the office discipline referral variable. The *below average* group was found to have moderately, but significantly more office discipline referrals than either the *meets standards* group, $U = 522.00$, $z = -2.49$, $p = .039$ (two-tailed), Cohen's $d = 0.52$, or the *exceeds standards* group, $U = 243.50$, $z = -2.78$, $p = .015$ (two-tailed), Cohen's $d = 0.63$. The latter two groups did not differ significantly, $U = 447.00$, $z = -0.87$, $p = .999$ (two-tailed).

The small number of students in this sample who were grade retained meant no meaningful inferential statistical analyses could be performed involving said variable. Several descriptive analyses were used to provide a profile of the retained students. All were male and all showed ISAT seventh-grade reading scores below average. Only one retained student's seventh-grade reading score was remarkably below average. Further, none of the five students with the lowest reading scores in seventh grade were subsequently grade retained. One retained student had no disciplinary referrals, but the other three were noticeably above average with respect to no office disciplinary referrals. In sum, the profile of the limited number of retained

students in this sample was: male, below-average reader, with a history of office discipline referrals.

Chapter 4 has provided results of analyses directed at addressing the study's three research questions. Chapter 5 will present more interpretation of those findings and will explain the study's strengths and limitations. Implications of the study's findings on educational practice, suggested areas for further research, and implications on leadership will conclude Chapter 5.

Chapter 5: Discussion and Conclusion

The purpose of this quantitative correlational study was to investigate the predictive relationship between reading scores on standardized tests and office discipline referrals and indicators of grade retention among seventh-grade students in Northwest Suburban Chicagoland school district. Students' behavioral conduct and reproductions of behavioral conduct can impede the acquisition of foundational academic skills due to time out of class for both brief and long periods. Time out of class can be attributed to talking with a school administrator about a behavioral incident, restrictive discipline or interventions, or grade retention. Results from this quantitative correlational study can be used by school leaders to develop academic and behavioral intervention programs at earlier grade level milestones, address school improvement plans, and serve as a means of collaborative data-driven decision making.

Chapter 2 reviewed existing literature surrounding student reading achievement abilities, student discipline and time out of class, and student grade retention as pertaining to student academic success and progress toward graduation. Students may be removed from class, missing or causing a disruption in reading instruction, as a result of disruptive behavior within a classroom setting. A student who struggles with academics or finds academics challenges building, may experience frustration and exhibit disruptive behaviors which may lead to office discipline referrals. Teachers, when confronted with disruptive student behavior, turn to behavior modification strategies aimed to address immediate situations and may never return to student's behavior incidents to discover the behavior's root cause.

The study intended to provide educational practitioners with practical data predictors which can be used when planning to address both academic and behavioral challenges among students. Educational leaders and school staff can use discovered results to plan lessons, address

in class interventions, and work communally with school leaders to develop school improvement plans. For example, once ISAT reading scores are furnished to schools, educators can immediately identify, through data-based inquiry, students who could benefit from academic interventions. Identifying early and providing interventions may have a residual effect on students' future office discipline referral occurrence.

Results from the study provided data trends showing correlational relationships among student reading scores and behavior, and common descriptive profiles for high school students who were grade retained. As detailed by literature in Chapter 2, there is minimal empirical research addressing student discipline as a function of student reading performance using a valid and reliable measure of reading performance. The study aimed to contribute to existing literature which already shows trends in and among school-level academic and behavioral data. Existing literature does not address situations where school leaders would like to know a student-specific likelihood of an outcome based on identified data.

Chapter 3 detailed methodology, research questions, and hypotheses for the study. This quantitative correlational study explored predictive relationships between predictor and criterion variables. Student reading scores on standardized tests acted as the predictor variable, and office discipline referrals and grade retention occurrence acted as criterion variables. Statistical analysis was needed for this study to provide inferences about relationships or differences among said variables.

Data analysis supported H_a1 indicating a significant correlation between seventh-grade standardized reading assessment scores from ISAT and future occurrence of office discipline referrals during students freshman and sophomore years. Higher occurrence of office discipline referrals was evident in student test scored in the bottom half of ISAT score distribution models.

Data analysis supported H_b2 indicating a significant correlation between student disruptive behaviors and average to above average standardized reading assessment scores. Analysis showed mean ranks of office discipline referrals differed significantly as a function of ISAT scores. Using discovered information, school leaders can use the correlational relationship as a guide when reviewing data and determining student-specific academic interventions. Timely academic interventions aimed at increasing reading ability during students' middle school years can have lasting effects, such as decreased occurrence of office discipline referrals, when students are in high school.

Data gathered, analyzed, and summarized in Chapter 4 provided both demographical student data as well as addressed research questions within the study. Of 100 student data used for analysis, data consisted of 38 females and 62 males, all of which were in seventh grade during the 2012-2013 school year. Of 100 students, 66 were 12 years of age, and 34 were 13 years of age. All students were continually enrolled in Northwest Suburban Chicagoland school district schools starting with school year 2012-2013 through 2015-2016.

Chapter 5 provides a discussion around research questions and how further investigation can be done based on data interpretations. Overall findings, limitations, recommendations, and leadership implications are presented in Chapter 5. Chapter 5 begins by providing a summary of findings and data interpretation.

Findings, Interpretations, Conclusions

Chapter 4 detailed results from data analysis relating to three research questions and respective hypotheses. Interpretations and conclusions are drawn based on statistical analysis and within the context of the study's theoretical framework among existing literature. Interpretation of statistical analyses allows for support or rejection of a research question's

hypothesis. This study contained three research questions and corresponding hypotheses.

Research Questions and Hypotheses

Research questions were designed to guide this quantitative correlational research study. Data included an anonymous group of 100 students reading scores, behavioral data, and grade retention status. This current study was designed to address the following research questions:

Research question 1. What degree of correlation, if any, exists between seventh-grade standardized reading assessment scores from ISAT and future occurrence of office discipline referrals during students' freshman and sophomore years?

Research question 2. What degree of correlation, if any, exists between student office discipline referrals and average to above average standardized reading assessment scores?

Research question 3. What degree of correlation, if any, exists between seventh-grade standardized reading assessment scores from ISAT and grade retention throughout students' freshman and sophomore year in high school?

Office discipline referral data and grade retention occurrence frequency from 2014-2015 and 2015-2016 school years served as the study's dependent variables with scores from ISAT assessment from 2012-2013 school year served as independent variable. The hypotheses for this current study's research questions are:

$H_0$1: There is no significant correlation between seventh-grade standardized reading assessment scores from ISAT and future occurrence of office discipline referrals during students freshman and sophomore years.

H_a1: There is significant correlation between seventh-grade standardized reading assessment scores from ISAT and future occurrence of office discipline referrals during students freshman and sophomore years.

$H_0$2: There is no significant correlation between student office discipline referrals and average to above average standardized reading assessment scores.

H_b2: There is significant correlation between student disruptive behaviors and average to above average standardized reading assessment scores.

$H_0$3: There is no significant correlation between seventh-grade standardized reading assessment scores from ISAT and future occurrence of office discipline referrals during students' freshman and sophomore years.

H_c3: There is significant correlation between seventh-grade standardized reading assessment scores from ISAT and future occurrence of office discipline referrals during students' freshman and sophomore years.

For Research Question 1, the Spearman correlation between seventh-grade reading scores and office discipline referrals was of moderate strength by Cohen's (1988) standards, negative, and statistically significant, $\rho(98) = -.29$, $p = .003$ (two-tailed). Elevated numbers of office discipline referrals occurred in the bottom half of ISAT reading test score distribution as shown in Figure 4. After data analysis and interpretation, the null hypothesis ($H_0$1), there is no significant correlation between seventh-grade standardized reading assessment scores from ISAT and future occurrence of office discipline referrals during students freshman and sophomore years, was rejected.

For Research Question 2, ISAT reading test scores were categorized into four categories: *academic warning* ($n = 5$), *below standards* ($n = 29$), *meets standards* ($n = 43$), and *exceeds standards* ($n = 23$). The bottom performing category, *academic warning*, contained few students which did not allow for meaningful analysis specific to *academic warning* category. To conduct meaningful analysis, *academic warning* was combined with *below standards* to create one *below*

average (*n* = 34) category. Descriptive statistics on office discipline referrals and ISAT seventh-grade reading scores are provided in Table 4. Because of differences among data distributions within the three groups, Kruskal-Wallis *H* test was implemented for this study as a statistically appropriate means of comparing the three groups' mean ranks (Hart, 2001). Kruskal-Wallis *H* test proved statistically significant, $H(2) = 10.42$, $p = .005$, indicating mean ranks of office discipline referrals differed significantly as a function of students' ISAT reading test scores. After data analysis and interpretation, the null hypothesis ($H_0 2$), there is no significant correlation between student office discipline referrals and average to above average standardized reading assessment scores, was rejected.

For Research Question 3, analysis was not performed due to statistically insignificant data. Of 100 student data, only 4 within the sample were grade retained. Grade retention data could not covary with ISAT reading test scores and could not be correlated. A descriptive profile of high school grade retained students was generated and included ISAT reading test scores and is summarized in Table 6. All 4 (100.0%) grade retained students were male compared to 62 (62.0%) from the sample of 100. Each grade retained students' ISAT seventh-grade reading score category was above *academic warning*. Within the sample of 100 student data, the median number of office discipline referrals was zero, but among the four grade-retained students, three were above average and one had none.

Interpretation of Findings

Transformational leadership served as theoretical framework for this study and is described as a process where leaders address subordinates' needs through leaders' actions (Bass & Riggio, 2006). Empowerment of subordinates through transformational leadership methods can lead to improvement with subordinates' motivation leading to overall improvement and

collective organizational success (Bass & Riggio, 2006). By arming subordinates with the study's statistically supported data driven results and decision-making processes, leaders create agents of change leading to organizational improvement and success. Educational leaders should engage and collaborate with professional educational colleagues when creating or refining a school's vision. Leaders should empower colleagues, through transformational leadership theories and by providing intellectual stimulation, to make data driven decisions using existing student data (Lee & Kuo, 2019).

Education leaders rely on implementation and innovation from colleagues or subordinates to enact successful and meaningful organizational change. By empowering colleagues or subordinates, educators can approach existing situations or processes in new, innovative ways (Bass & Riggio, 2006). Sharing of data, engaging in meaningful conversations, and empowerment and delegation with and among school staff may seem uncomfortable or unconventional for some leaders. When leaders model said behavior, actively and genuinely engage with colleagues, leaders share power and attempt to stimulate colleague interest in a common good (Lee & Kuo, 2019; Muchiri, McMurray, Nkhoma, & Pham, 2019). Results from this study can be transferred to school improvement or leadership teams and coordinated in conjunction with teachers' professional learning teams to differentiate instruction and meet the needs of all student learners. School teams can use student reading scores and subsequent ranges as predictors when planning intentional academic or behavioral instructional activities. School teams or specialized support teams plan academic interventions aimed at increasing reading ability which can provide two possible benefits: increased student reading ability and decreased frequency of student behavior referrals.

Without a statistical analysis of grade retention occurrence, school teams should consider multiple data sources until a reliable and valid statistical analysis can be completed. Whether an analysis exists or not, four of 100 students within the sample were grade retained during high school. Existing data indicate any student two or more years behind like peers in reading is at risk of truancy, time out of class, and dropping out of school (Kane County Regional Office of Education, 2016). Although this current study could neither support nor reject Research Question 3's hypothesis, existing literature and data show general relationships among reading achievement and grade retention or dropping out of school. Further study and analysis are needed to investigate Research Question 3 and its implication on student success.

Limitations

This study was limited to data collected from one high school within a school district containing three high schools and three middle schools within a school district containing five middle schools. The computed minimum sample size of 84 was increased to 100 to account for missing or erroneous data but is a narrow sample group compared to actual enrollment among students in seventh grade (Creswell, 2013; Sullivan & Feinn, 2012). Attendance records for participants were not examined as part of this study. Although student attendance may provide supporting information on academic effects of time out of class, attendance is an external factor schools cannot necessarily control as is encompasses short-term illness, hospitalization, truancy, court appearance, incarceration, family issues, and vacation (Kane County Regional Office of Education, 2016).

This study was limited to a sample size of 100, of which only four students' data contained grade retention status. As an unfortunate consequence, formal statistical analysis was not possible. A profile of grade retained students was created and is summarized in Table 6.

Reliability and validity for the current study are discussed in Chapter 3. Data analyzed were provided directly from Northwest Suburban Chicagoland school district's student management information system. Data can be considered reliable given clearly defined school district policies and procedures relating to student discipline, grade retention, standardized testing, and student record management. Having only four student data relating to grade retention prevented statistical analysis of said variable meaning findings described in Chapter 4 cannot be assumed to generalize beyond the sample of 100. Illinois Standards Achievement Test reading assessment data came in the form of content validity, construct validity, and criterion-related validity (ISBE, 2013a).

Recommendations

Findings and conclusions from this study form the basis for recommendations for further research. Educators share an ongoing responsibility to review student data when discussing school improvement and instructional design. The study's review of literature revealed a need for additional research with and among student reading achievement and ability and office discipline referrals. Further research is suggested to investigate relationships among variables as predictors of student behaviors including but not limited to grade retention occurrence.

The study attempted to discover relationships between variables derived from student data and was limited to a sample of 100. Of 100 pieces of student data, only 4 students were grade retained. Not having enough data on grade retention status, appropriate statistical analysis was unachievable for purposes of this study. Further research is necessary to determine if relationships exist among reading ability and student grade retention status, specifically addressing reading ability as a future predictor of grade retention status. Using data other than reading ability as a predictor of grade retention status is recommended. By studying predictors

and potential causes of grade retention, educators can be supported in classrooms by differentiating instructional plans to meet individual student's needs and strengths.

To further support school improvement and classroom instruction, educational leaders should examine relationships between and among student data pieces at various other milestones within a student's school career. Educators should consider valid and reliable progress monitoring methods, such as iReady or iStation, when examining student data. Intensifying further research to include multiple grade levels, assessments, and schools within a school district is advised. Further research should address whether student academic indicators, other than reading abilities, predict office discipline referrals. For starters, educators can consult data surrounding student progress monitoring or attendance data trends.

Implications for Leadership

Results from this study provide educational leaders with data driven prediction methods employable by school staff during school improvement processes. School leaders should implement practices and activities involving teachers in decision making and instructional design. Through transformational leadership, sharing or transferring power to subordinates or colleagues fosters collaboration, increased motivation, ongoing professional development, and new ideas. A shared approach provides leaders with insight to innovation, use of internal staff members' talents, and a likely increase of environmental work climate.

With increasing access to data available to classroom teachers, school leaders should provide meaningful professional development focused on use of data and data-based decision making. Many student management information systems include data visualization tools, such as Tableau, which allow customization of data dashboards for end user use. School communities as a whole should engage in data examination on a regular basis as part of professional learning

communities as a means of embedding data-driven conversations and methodology into professional practice.

With some data analysis, the use of data as predictors for future student behavioral events can be transferrable to other grade levels and academic indicators. Educational leaders can transfer results of this study to other milestones such as transition from elementary to middle school or middle to high school. Data contains essential student information necessary for informed decision making. School leaders can use student data when scheduling interventions for students which may provide timely interventions for students in need rather than rely on conjecture which may not provide accurate or appropriate data.

Conclusion

Educators and parents alike have mutual interest in student success. This study provided data relating to student academic progress, discipline behavior, and grade retention and relationships between and among these data. Since data-based exploration and decision making does not require acquisition of new information, the study's data results can be immediately implemented within established school systems and used equally by school leaders and teachers. Innovative and effective means of school improvement are desired by all parties involved with students benefitting from either or both academic or behavioral interventions. Transformational leadership is grounded in sharing of responsibility and power through empowerment (Bass, 1999). New ideas and approaches to educational settings and classroom instruction are beneficial byproducts of transformational leadership and data-based decision making.

The purpose of this quantitative investigation was to investigate the relationship between seventh-grade reading scores on standardized tests and office discipline referrals and indicators of grade retention among seventh grade students in Northwest Suburban Chicagoland school

district. Student success on standardized reading tests, combined with office discipline referrals which take students out of classroom settings, and student grade retention effect students and schools similarly. With a data-driven method of prediction of future office discipline referrals, schools and educators can provide academic interventions at earlier grade levels to enhance students' foundational reading skill levels or concurrently provide academic interventions to students. Early interventions can combat a notion which studies show students who struggle with academic progress and concepts experience frustration while in class which may manifest as disruptive behaviors leading to time out of class (Bruhn & Watt, 2012). Regardless of position or job title, all educators are educational leaders and agents of change with students' best interests in mind.

References

Alexander, K. L., Entwisle, D. R., & Kabbani, N. S. (2001). The dropout process in life course perspective: Early risk factors at home and school. *Teachers College Record, 103*(5), 760-822. doi:10.1111/0161-4681.00134

Allman, K. L., & Slate, J. R. (2011). School discipline in public education: A brief review of current practices. *International Journal of Educational Leadership Preparation, 6*(2). Retrieved from http://cnx.org/contents/104795f8-7143-42eb-8f1c-5aa3f6054fff@1

Alversia, Y. (2011). Doing quantitative research in education with SPSS. *Evaluation & Research In Education, 24*(4), 305. doi:10.1080/09500790.2011.596379

Andersen, L. B., Bjørnholt, B., Bro, L. L., & Holm-Petersen, C. (2018). Achieving high quality through transformational leadership: A qualitative multilevel analysis of transformational leadership and perceived professional quality. *Public Personnel Management, 47*(1), 51-72. doi:10.1177/0091026017747270

Aziz, T., Mahomed, F., & Mason, D. (2016). A unified compatibility method for exact solutions of nonlinear flow models of Newtonian and non-Newtonian fluids. *International Journal of Non-Linear Mechanics, 78*, 142-155. doi:10.1016/j.ijnonlinmec.2015.01.003

Babbie, E. R. (2013). *The practice of social research*. Belmont, CA: Wadsworth Cengage Learning.

Bailey, B., & Stegelin, D. A. (2003). Creating a sense of place: Anchoring at-risk students within K-12 classrooms. *The Journal of At-Risk Issues, 9*(2), 17-26. Retrieved from www.dropoutprevention.org/resources/journals

Barczak, G. (2015). Publishing qualitative versus quantitative research. *Journal of Product Innovation Management, 32*(5), 658. doi:10.1111/jpim.12277

Bass, B. M. (1999). Two decades of research and development in transformational leadership. *European Journal of Work and Organizational Psychology*, *8*(1), 9-32. doi:10.1080/135943299398410

Bass, B. M., & Riggio, R. E. (2006). *Transformational leadership* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Berger, J., Bayarri, M. J., & Pericchi, L. R. (2013). The effective sample size. *Econometric Reviews*, *33*(1-4), 197-217. doi.org/10.1080/07474938.2013.807157

Bernard, M. (2017). Impact of teaching attitudes and behaviors for learning on the reading achievement of students falling behind. *International Journal of Learning, Teaching and Educational Research*, 16(8), 51-64. Retrieved from https://www.ijlter.org

Berninger, V., Abbott, R., Cook, C. R., & Nagy, W. (2017). Relationships of attention and executive functions to oral language, reading, and writing skills and systems in middle childhood and early adolescence. *Journal of Learning Disabilities*, *50*(4), 434-449. doi:10.1177/0022219415617167

Bishara, A. J., & Hittner, J. B. (2012). Testing the significance of a correlation with non-normal data: Comparison of Pearson, Spearman, transformation, and resampling approaches. *Psychological Methods, 17,* 399-417. doi:10.1037/a0028087

Bottiani, J. H., Bradshaw, C. P., & Mendelson, T. (2017). A multilevel examination of racial disparities in high school discipline: Black and white adolescents' perceived equity, school belonging, and adjustment problems. *Journal of Educational Psychology*, *109*(4), 532. doi:10.1037/edu0000155

Bruhn, A., & Watt, S. (2012). Improving behavior by using multicomponent self-monitoring within a targeted reading intervention. *Behavioral Disorders*, *31*(1), 3–17. doi.org/10.1177/019874291203800102

Bücher, A., Dette, H., & Wieczorek, G. (2011) Testing model assumptions in functional regression models. *Journal of Multivariate Analysis*, *102*(10), 1472-1488. doi:10.1016/j.jmva.2011.05.014

Camm, J. D. (2012). *Quantitative methods for business* (12th, International ed.). Andover, Mason, Ohio: South-Western.

Chávez, C. L., Yáñez, G., Catroppa, C., Rojas, S., Escartin, E., Hearps, S. J., & García, A. (2016). Adolescents with vascular frontal lesion: A neuropsychological follow up case study. *Neurocirugía, 27*(3), 136-143. doi:10.1016/j.neucir.2015.09.001

Chung, K. K. H. (2015). Behavioural self-regulation and its contribution to reading among Chinese poor readers. *Asia Pacific Journal of Developmental Differences, 2*(1), 5-25. doi:10.3850/S2345734115000174

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Conover, W. J., & Iman, R. L. (1979). *On multiple-comparisons procedures* (LA-7677-MS). Retrieved from Los Alamos Scientific Laboratory website: https://permalink.lanl.gov/object/tr?what=info:lanl-repo/lareport/LA-07677-MS

Corder, G. W., & Foreman, D. I. (2009). *Nonparametric statistics for non-statisticians: A step-by-step approach* [Paper]. doi:10.1002/9781118165881

Council, M. R., Cartledge, G., Green, D., Barber, M., & Gardner, R. (2016). Reducing risk through a supplementary reading intervention: A case study of first- and second-grade urban students. *Behavioral Disorder*s, 41(4), 241-257. doi:10.17988/bedi-41-04-241-257.1

Creswell, J. W. (2013). *Qualitative inquiry and research design: Choosing among five traditions*. Thousand Oaks, CA: Sage Publications.

Croft, S. J., Roberts, M. A., & Stenhouse, V. L. (2015). The perfect storm of education reform: High-stakes testing and teacher evaluation. *Social Justice, 42*(1), 70-92. Retrieved from https://www.socialjusticejournal.org/

Curtis, E. A., Comiskey, C., & Dempsey, O. (2016). Importance and use of correlational research. *Nurse Researcher, 23*(6), 20-25. doi:10.7748/nr.2016.e1382

DeMatthews, D. (2016). Effective leadership is not enough: Critical approaches to closing the racial discipline gap. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas, 89*(1), 7-13. doi:10.1080/00098655.2015.1121120

DeMatthews, D., Carey, R. L., Olivarez, A., & Moussavi Saeedi, K. (2017). Guilty as charged? Principals' perspectives on disciplinary practices and the racial discipline gap. *Educational Administration Quarterly, 53*(4), 519-555. doi:10.1177/0013161X17714844

Dunn, O. J. (1964). Multiple comparisons using rank sums. *Technometrics*, 6, 241-252. doi:10.1080/00401706.1964.10490181

Eide, E. R., & Showalter, M. H. (2001). The effect of grade retention on educational and labor market outcomes. *Economics of Education Review*, *20*(6), 563-576. doi:10.1016/s0272-7757(00)00041-8

Eisinga, R., Grotenhuis, M., & Pelzer, B. (2013). The reliability of a two-item scale: Pearson, Cronbach, or Spearman-Brown? *International Journal of Public Health, 58*(4), 637-642. doi:10.1007/s00038-012-0416-3

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*(4), 1149-1160. doi.org/10.3758/BRM.41.4.1149

Faul, F., Erdfelder, E., Lang, A., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175-191. doi.org/10.3758/BF03193146

Field, A. (2013). *Discovering statistics using IBM SPSS statistics* (4th ed.). Thousand Oaks, CA: Sage.

Finch, H. (2005) Comparison of the performance of nonparametric and parametric MANOVA test statistics when assumptions are violated. *Methodology, 1*(1), 27-38. doi:10.22237/jmasm/1383278580

Fine, J. G., & Davis, J. M. (2003). Grade retention and enrollment in post-secondary education. *Journal of School Psychology*, *41*(6), 401-411. doi:10.1016/j.jsp.2003.07.001

Follmer, D. J. (2018). Executive function and reading comprehension: A meta-analytic review. *Educational Psychologist, 53*(1), 42-60. doi:10.1080/00461520.2017.1309295

Ford, A. D., Olmi, D. J., Edwards, R. P., & Tingstrom, D. H. (2001). The sequential introduction of compliance training components with elementary-aged children in general education classroom settings. *School Psychology Quarterly*, *16*(2), 142–157. doi.org/10.1521/scpq.16.2.142.18702

Friedman, N. P., & Miyake, A. (2017). Unity and diversity of executive functions: Individual differences as a window on cognitive structure. *Cortex, 86*, 186-204. doi:10.1016/j.cortex.2016.04.023

Gage, N. A., Sugai, G., Lunde, K., & DeLoreto, L. (2013). Truancy and zero tolerance in high school: does policy align with practice? *Education and Treatment of Children, 36*(2), 117-138. doi:10.1353/etc.2013.0011

García, C., García, J., López Martín, M., & Salmerón, R. (2015). Collinearity: Revisiting the variance inflation factor in ridge regression. *Journal of Applied Statistics, 42*(3), 648-661. doi:10.1080/02664763.2014.980789

Gardner, R. C., & Neufeld, R. W. (2013). What the correlation coefficient really tells us about the individual. *Canadian Journal of Behavioural Science*, *45*(4), 313-319. doi:10.1037/a0033342

Garwood, J. D., Varghese, C., & Vernon-Feagans, L. (2017). Internalizing behaviors and hyperactivity/inattention: consequences for young struggling readers, and especially boys. *Journal of Early Intervention, 39*(3), 218-235. doi:10.1177/1053815117706524

Gesell, A. (1925). The mental growth of the preschool child. *American Journal of Nursing*, *25*(7). doi:10.1097/00000446-192507000-00030

Glover, T. A., Reddy, L. A., Kettler, R. J., Kurz, A., & Lekwa, A. J. (2016). Improving high-stakes decisions via formative assessment, professional development, and comprehensive educator evaluation: The school system improvement project. *Teachers College Record, 118*(14), 1-26. Retreived from http://www.tcrecord.org/

Goodwin, C. J., & Goodwin, K. A. (2013). *Research in psychology: Methods and design* (7th ed.). Hoboken, NJ: Wiley.

Gravetter, F. J., & Forzano, L. B. (2016). *Research methods for the behavioral sciences* (5th ed.). Stamford, CT: Cengage.

Haas, J. P. (2012). Sample size and power. *American Journal of Infection Control*, *40*(8), 766 – 767. doi:10.1016/j.ajic.2012.05.020

Hair, J., Black, W., Babin, B., & Anderson, R. E. (2010). *Multivariate data analysis* (3rd ed.) Upper Saddle River, NJ: Prentice-Hall.

Halcomb, E., & Hickman, L. (2015). Mixed methods research. *Nursing Standard*, *29*(32), 41-47. doi:10.7748/ns.29.32.41.e8858

Hancock, G. R., & Mueller, R. O. (2010). *The reviewer's guide to quantitative methods in the social sciences*. New York: Routledge.

Hart, A. (2001). Mann-Whitney test is not just a test of medians: Differences in spread can be important. *BMJ*, *323*, 391-393. doi:doi.org/10.1136/bmj.323.7309.391

Heilbrun, A., Cornell, D., & Lovegrove, P. (2015). Principal attitudes regarding zero tolerance and racial disparities in school suspensions. *Psychology in the Schools, 52*(5), 489-499. doi:10.1002/pits.21838

Henderikus, S. (2010). Theory. In N. J. Salkind (Ed.), *Encyclopedia of research design* (pp. 1498-1502). Thousand Oaks, CA: Sage.

Herman, W. E. (2011). *Motivational correlates of academic success in an educational psychology course*. Paper presented at Conference on the Teaching of Psychology: Ideas & Innovations, Potsdam, New York. Retrieved from https://www.semanticscholar.org

Herszenhorn, D. M. (2004, March 25). In reversal, Chicago eases promotion. *New York Times*.

Hickman, G. P., Bartholomew, M., Mathwig, J., & Heinrich, R. S. (2008). Differential development pathways of high school dropouts and graduates. *The Journal of Educational Research*, *102*(1), 3-14. doi:10.3200/joer.102.1.3-14

Hoe, J., & Hoare, Z. (2012). *Understanding quantitative research: Part 1*. Nursing Standard, 27(15-17), 52. doi:10.7748/ns2012.12.27.15.52.c9485

Hoffman, S. (2014). Zero benefit: Estimating the effect of zero tolerance discipline polices on racial disparities in school discipline. *Educational Policy, 28*(1), 69-95. doi:10.1177/0895904812453999

Hong, H., Pradhan, B., Xu, C., & Bui, D. T. (2015). Spatial prediction of landslide hazard at the Yihuang area (China) using two-class kernel logistic regression, alternating decision tree and support vector machines. *Catena, 133*, 266-281. doi:10.1016/j.catena.2015.05.019

Hughes, J. N., West, S. G., Kim, H., & Bauer, S. S. (2017). Effect of early grade retention on school completion: A prospective study. *Journal of Educational Psychology*, 110(7), 974-91. doi:doi.org/10.1037/edu0000243

Illinois State Board of Education. (2013a). Illinois standards achievement test - 2013 technical manual. Retrieved from https://www.isbe.net/documents/isat_tech_2013.pdf

Illinois State Board of Education. (2013b). Interpretive guide 2013. Retrieved from https://www.isbe.net/documents/interp-guide13.pdf

Illinois State Board of Education. (2017). *2017 Illinois report card glossary of terms*. Retrieved from https://www.isbe.net/Documents/2017_Report_Card_Glossary.pdf#search=cohort%20definition

Indera, N., Yassin, I., Zabidi, A., & Rizman, Z. (2017). Non-linear autoregressive with exogeneous input (NARX) Bitcoin price prediction model using PSO-optimized parameters and moving average technical indicators. *Journal of Fundamental and Applied Sciences, 9*(3S), 791-808. doi:10.4314/jfas.v9i3s.61

Jimerson, S. R., Woehr, S. M., & Kaufman, A. M. (2007). *Grade retention and promotion: Information for parents*. Retrieved from National Association of School Psychologists website: http://www.w.naspcenter.org/

Johnson, B., & Christensen, L. B. (2012). *Educational research: Quantitative, qualitative, and mixed approaches* (4th ed.). Thousand Oaks, CA: SAGE Publications.

Johnson, S. B., Sudhinaraset, M., & Blum, R. W. (2009). Neuromaturation and adolescent risk taking: Why development is not determinism. *Journal of Adolescent Research, 25*(1), 4-23. doi:10.1177/0743558409353339

Kalpakci, A., Ha, C., & Sharp, C. (2018). Differential relations of executive functioning to borderline personality disorder presentations in adolescents. *Personality and Mental Health*, *12*(2), 93-106. doi:10.1002/pmh.1410

Kane County Regional Office of Education. (2016). *Truancy prevention manual*. Retrieved from http://www.kaneroe.org/wp-content/uploads/2017/02/Truancy-Manual-Public-v2.pdf

Kane, M. T. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement*, *50*(1), 1-73. doi:10.1111/jedm.12000

Kenneady, L. M. (2004). *Good for nothing: In-grade retention*. Retrieved from Intercultural Development Research Association website: https://www.idra.org/resource-center/good-for-nothing-in-grade-retention/

King, K. R., Gonzales, C. R., & Reinke, W. M. (2018). Empirically derived subclasses of academic skill among children at risk for behavior problems and association with distal academic outcomes. *Journal of Emotional and Behavioral Disorders, 2*(1), 16-23. doi:10.1177/1063426617754082

King, P. E., & Finn, A. N. (2017). A test of attention control theory in public speaking: Cognitive load influences the relationship between state anxiety and verbal production. *Communication Education, 66*(2), 168-182. doi:10.1080/03634523.2016.1272128

Klapproth, F., Schaltz, P., Brunner, M., Keller, U., Fischbach, A., Ugen, S., & Martin, R. (2016). Short-term and medium-term effects of grade retention in secondary school on academic achievement and psychosocial outcome variables. *Learning and Individual Differences*, 50, 182-194. doi:10.1016/j.lindif.2016.08.014

Kruskal, W. H., & Wallis, W. A. (1952). Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association, 47*(260), 583-621. doi:10.1080/01621459.1952.10483441

Leavy, P. (2017). *Research design: Quantitative, qualitative, mixed methods, arts-based, and community-based participatory research approaches*. New York: Guilford Press.

Lee, Y. D., & Kuo, C. T. (2019). Principals' transformational leadership and teachers' work motivation: Evidence from elementary schools in Taiwan. *International Journal of Organizational Innovation, 11*(3), 90-113. Retrieved from http://www.ijoi-online.org

Leedy, P. D., & Ormrod, J. E. (2012). *Practical research: Planning and design* (10th ed.). Boston: Pearson.

Lehman, A., O'Rourke, N., Hatcher, L., & Stepanski, E. (2013). *JMP for basic univariate and multivariate statistics: Methods for researchers and social scientists* (2nd ed.). Retrieved from https://www.sas.com/

Lewis, S., & Hardy, I. (2015). Funding, reputation and targets: The discursive logics of high-stakes testing. *Cambridge Journal of Education, 45*(2), 245-264. doi:10.1080/0305764X.2014.936826

Li, J., & Valliant, R. (2015). Linear regression diagnostics in cluster samples. *Journal of Official Statistics, 31*(1), 61-75. doi:10.1515/jos-2015-0003

Ma, L., Yan, X., Wei, C., & Wang, J. (2016). Modeling the equivalent property damage only crash rate for road segments using the hurdle regression framework. *Analytic Methods in Accident Research, 11*, 48-61. doi:10.1016/j.amar.2016.07.001

Mahjoub, M. (2017). The treatment effect of grade repetitions. *Education Economics*, 25(4), 418-432. doi:10.1080/09645292.2017.1283006

Maier, M. P., Pate, J. L., Gibson, N. M., Hilgert, L., Hull, K., & Campbell, P. C. (2016). A quantitative examination of school leadership and response to intervention. *Learning Disabilities Research & Practice*, *31*(2), 103-112. doi:10.1111/ldrp.12100

Marchbanks, M., Blake, J., Booth, E. A., Carmichael, D., Seibert, A. L., & Fabelo, T. (2013). The economic effects of exclusionary discipline on grade retention and high school dropout. Retrieved from University of California Los Angeles: The Civil Rights Project website: https://escholarship.org/uc/item/7zc0c135

Marshall, K. (2016). Rethinking differentiation — Using teachers' time most effectively. *Phi Delta Kappan, 98*(1), 8-13. doi:10.1177/0031721716666046

Martorell, P., & Mariano, L. T. (2018). The causal effects of grade retention on behavioral outcomes. *Journal of Research on Educational Effectiveness, 11*(2), 192-216. doi:10.1080/19345747.2017.1390024

Mathys, C., Véronneau, M., & Lecocq, A. (2017). Grade retention at the transition to secondary school: Using propensity score matching to identify consequences on psychosocial adjustment. *The Journal of Early Adolescence*, 39(1), 97-133. Retrieved from doi.org/10.1177/0272431617735651

Mazyck, D. (2013). Beneath the surface: Addressing health-related chronic absenteeism. *Principal Leadership*, *14*(4), 12-15. Retrieved from http://www.nassp.org

McIntosh, K., Horner, R. H., Chard, D. J., Boland, J. B., & Good, R. H. (2006). The use of reading and behavior screening measures to predict nonresponse to school-wide positive behavior support: A longitudinal analysis. *School Psychology Review*, *35*(2), 275–291. Retrieved from http://www.nasponline.org

Meltzer, E. P., Kapoor, A., Fogel, J., Elbulok-Charcape, M. M., Roth, R. M., Katz, M. J., Lipton, R. B., & Rabin, L. A. (2017). Association of psychological, cognitive, and functional variables with self-reported executive functioning in a sample of nondemented community-dwelling older adults. *Applied Neuropsychology: Adult, 24*(4), 364-375. doi:10.1080/23279095.2016.1185428

Merriam-Webster. (n.d.-a). Freshman. Retrieved from http://www.merriam-webster.com/dictionary/freshman

Merriam-Webster. (n.d.-b). Sophomore. Retrieved from http://www.merriam-webster.com/dictionary/sophomore

MetaMetrics. (n.d.). Linking DIBELS oral reading fluency with the Lexile framework for reading. Retrieved from https://lexile.com/

Micai, M., Kavussanu, M., & Ring, C. (2015). Executive function is associated with antisocial behavior and aggression in athletes. *Journal of Sport and Exercise Psychology, 37*(5), 469-476. doi:10.1123/jsep.2015-0021

Miller, L. A., & Lovler, R. L. (2016). *Foundations of psychological testing: A practical approach* (5th ed.). Thousand Oaks, CA: Sage.

Muchiri, M. K., McMurray, A. J., Nkhoma, M., & Pham, H. C. (2019). How transformational and empowering leader behaviors enhance workplace safety: A review and research agenda. *The Journal of Developing Areas, 53*(1), 257-265. doi:10.1353/jda.2019.0015

Mundy, L. K., Canterford, L., Tucker, D., Bayer, J., Romaniuk, H., Sawyer, S., Lietz, P., Redmond, G., Proimos, J., Allen, N., & Patton, G. (2017). Academic performance in primary school children with common emotional and behavioral problems. *Journal of School Health, 87*(8), 593-601. doi:10.1111/josh.12531

National Center for Learning Disabilities. (n.d.). What is Response to Intervention (RTI)? Retrieved from http://www.rtinetwork.org/learn/what/whatisrti

Ness, M. K. (2016). Reading comprehension strategies in secondary content-area classrooms: Teacher use of and attitudes towards reading comprehension instruction. *Reading Horizons*, *55.1*, 58-84. doi:10.1177/003172170708900314

Nitschke, K., Köstering, L., Finkel, L., Weiller, C., & Kaller, C. P. (2017). A meta-analysis on the neural basis of planning: Activation likelihood estimation of functional brain imaging results in the Tower of London task. *Human Brain Mapping, 38*(1), 396-413. doi:10.1002/hbm.23368

Nunes, L. C., Balcão Reis, A., & Seabra, C. (2018). Is retention beneficial to low-achieving students? Evidence from Portugal. *Applied Economics*, 50(40), 4306-4317. doi:10.1080/00036846.2018.1444261

Oliver, R. M., Wehby, J. H., & Reschly, D. J. (2011). Teacher classroom management practices: Effects on disruptive or aggressive student behavior. *Campbell Systematic Reviews, 7*(1), 1-55. doi:10.4073/csr.2011.4

Özek, U. (2015). Hold back to move forward? Early grade retention and student misbehavior. *Education Finance and Policy, 10*(3), 350-377. doi:10.1162/EDFP_a_00166

Özkanal, Ü., & Arıkan, N. (2011). The relation between success and absenteeism at Esogu English Preparatory School. *Journal of Language Teaching and Research, 2*(1), 68-72. doi:10.4304/jltr.2.1.68-72

Park, J., & Park, M. (2016). Qualitative versus quantitative research methods: Discovery or justification? *Journal of Marketing Thought, 3*(1), 1-7. doi:10.15577/jmt.2016.03.01.1

Payne, A. A., & Welch, K. (2015). Restorative justice in schools: The influence of race on restorative discipline. *Youth & Society, 47*(4), 539-564. doi:10.1177/0044118X12473125

Peters, K., & Halcomb, E. (2015). Interviews in qualitative research. *Nurse Researcher, 22*(4), 6. doi:10.7748/nr.22.4.6.s2

Pimperton, H., & Nation, K. (2014). Poor comprehenders in the classroom: Teacher ratings of behavior in children with poor reading comprehension and its relationship with individual differences in working memory. *Journal of Learning Disabilities, 47*(3), 199-207. doi:10.1177/0022219412454172

Plank, S. B., & Condliffe, B. F. (2013). Pressures of the season: An examination of classroom quality and high-stakes accountability. *American Educational Research Journal*, *50*(5), 1152-1182. doi:10.3102/0002831213500691

Price, J. H., & Murnan, J. (2004). Research limitations and the necessity of reporting them. *American Journal of Health Education*, *35*(2), 66-67. doi:10.1080/19325037.2004.10603611

Quinino, R. C., Reis, E. A., & Bessegato, L. F. (2013). Using the coefficient of determination R^2 to test the significance of multiple linear regression. *Teaching Statistics*, *35*(2), 84-88. doi:doi.org/10.1111/j.1467-9639.2012.00525.x

Rhodes, J., Thomas, J. M., & Liles, A. R. (2018). Predictors of grade retention among children in an elementary school truancy intervention. *Journal of At-Risk Issues*, *21*(1), 1-10. Retrieved from http://dropoutprevention.org/resources/journals

Rottman, B. M., & Hastie, R. (2014). Reasoning about causal relationships: Inferences on causal networks. *Psychological Bulletin*, *140*(1), 109-139. doi:10.1037/a0031903

Ryan, M. (2011). *Early warning indicator systems*. Retrieved from Education Commission of the States website: https://www.ecs.org/research-reports/

Savioja, H., Helminen, M., Fröjd, S., Marttunen, M., & Kaltiala-Heino, R. (2017). Delinquency and sexual experiences across adolescence: Does depression play a role? *The European Journal of Contraception & Reproductive Health Care*, *22*(4), 298-304. doi:10.1080/13625187.2017.1374361

Schargel, F. P., & Smink, J. (2001). *Strategies to help solve our school dropout problem*. Larchment, NY: Eye on Education.

Schwartz, J. A., Connolly, E. J., & Brauer, J. R. (2017). Head injuries and changes in delinquency from adolescence to emerging adulthood: The importance of self-control as a mediating influence. *Journal of Research in Crime and Delinquency, 54*(6), 869-901. doi:10.1177/0022427817710287

Schwerdt, G., West, M. R., & Winters, M. A. (2017). *The effects of test-based retention on student outcomes over time: Regression discontinuity evidence from Florida* (21509). Retrieved from National Bureau of Economic Research website: https://www.nber.org/papers/

Sedgwick, P. (2012). Pearson's correlation coefficient. *BMJ, 345*(1), 4483. doi:10.1136/bmj.e4483

Sheskin, D. J. (2011). *Handbook of parametric and nonparametric statistical procedures* (5th ed.). Boca Raton, FL: Chapman & Hall/CRC Press.

Singh, P., Märtsin, M., & Glasswell, K. (2015). Dilemmatic spaces: High-stakes testing and the possibilities of collaborative knowledge work to generate learning innovations. *Teachers and Teaching, 21*(4), 379-399. doi:10.1080/13540602.2014.976853

Skiba, R. J. (2014). The failure of zero tolerance. *Reclaiming Children and Youth, 22*(4), 27-36. Retreived from http://reclaimingjournal.com/

Slate, J. R., Gray, P. L., & Jones, B. (2016). A clear lack of equity in disciplinary consequences for Black girls in Texas: A statewide examination. *The Journal of Negro Education, 85*(3), 250-260. doi:10.7709/jnegroeducation.85.3.0250

Sullivan, G. M., & Feinn, R. (2012). Using effect size—or why the P Value is not enough. *Journal of Graduate Medical Education, 4*(3), 279–282. doi.org/10.4300/JGME-D-12-00156.1

Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed.). Boston, MA: Pearson.

Taillardat, M., Mestre, O., Zamo, M., & Naveau, P. (2016). Calibrated ensemble forecasts using quantile regression forests and ensemble model output statistics. *Monthly Weather Review, 144*(6), 2375-2393. doi:10.1175/MWR-D-15-0260.1

Terrell, S. R. (2012). Mixed-methods research methodologies. *The Qualitative Report*, *17*(1), 254-280. Retrieved from https://nsuworks.nova.edu/tqr/vol17/iss1/14

Tokunaga, H. T. (2019). *Fundamental statistics for the social and behavioral sciences* (2nd ed.). Los Angeles, CA: Sage.

U.S. Department of Education. (n.d.). *School and district navigator*. Retrieved from National Center for Educational Statistics website: https://nces.ed.gov/ccd/schoolmap/#

U.S. Department of Health and Human Services. (2004, April 1). *Measures of material hardship*. Retrieved from https://aspe.hhs.gov/execsum/measures-material-hardship#Defining

Walls, G. (2012). Is systematic quantitative research scientifically rigorous? Methodological and statistical considerations. *Journal of the American Psychoanalytic Association, 60*(1), 145-152. doi:10.1177/0003065111435699

Wang, X., Cai, L., Li, L., Yang, Y., Yao, S., & Zhu, X. (2016). Neurological soft signs in Chinese adolescents with antisocial personality traits. *Psychiatry Research, 243*, 143-146. doi:10.1016/j.psychres.2016.04.023

Whitley, B. E., Kite, M. E., & Adams, H. L. (2013). *Principles of research in behavioral science* (3rd ed.). New York: Psychology Press.

Wisniewski, M. (2016). *Quantitative methods for decision makers* (6th ed.). Harlow, Essex, England: Pearson.

Witte, J. F., Wolf, P. J., Cowen, J. M., Carlson, D. E., & Fleming, D. J. (2014). High-stakes choice: Achievement and accountability in the nation's oldest urban voucher program. *Educational Evaluation and Policy Analysis, 36*(4), 437-456. doi:10.3102/0162373714534521

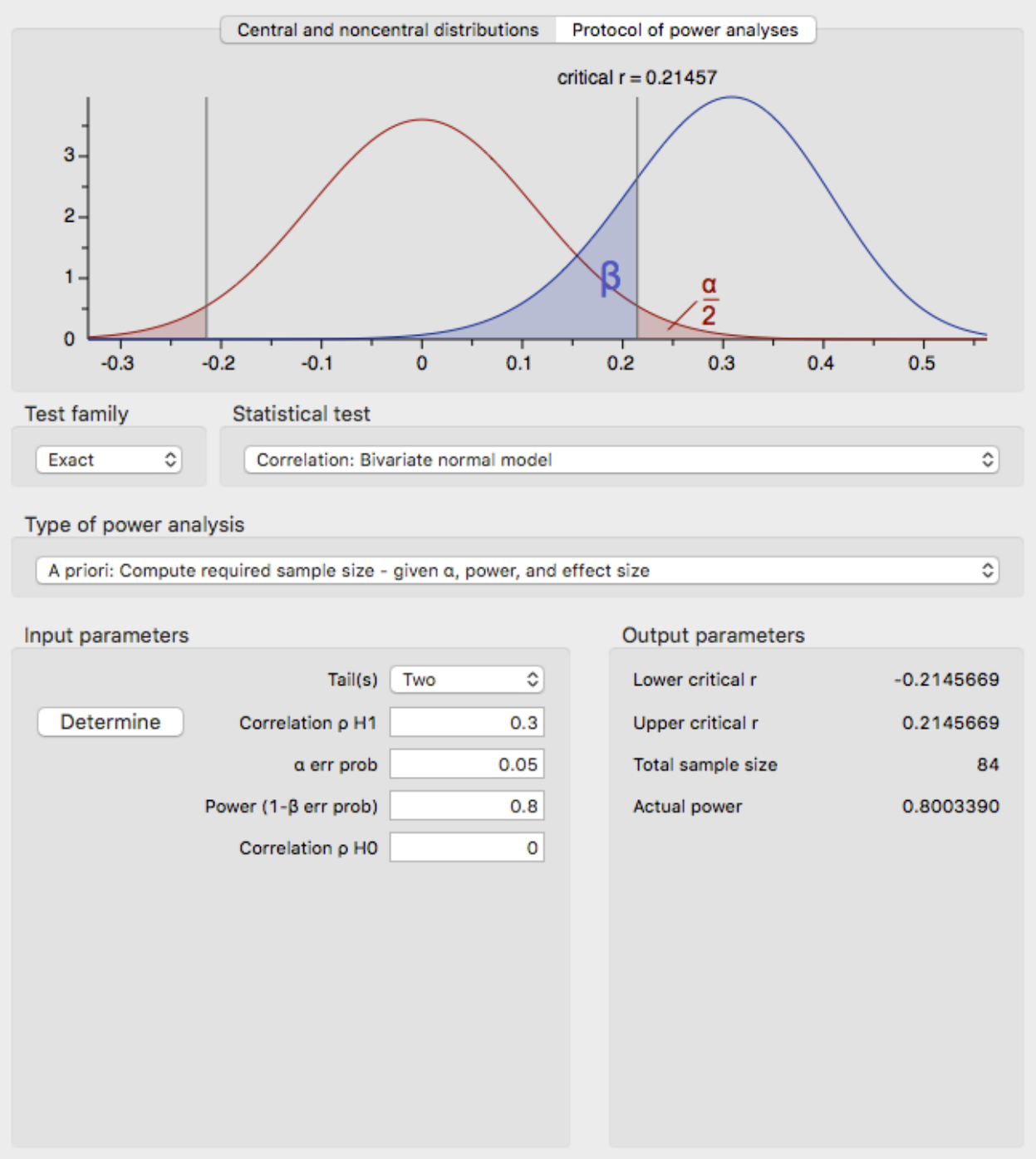
Xia, N., & Kirby, S. N. (2009). *Retaining students in grade: A literature review of the effects of retention on students' academic and nonacademic outcomes: summary*. Retrieved from RAND Education website: http://www.rand.org

Yang, M., Chen, Z., Rhodes, J. L., & Orooji, M. (2018). A longitudinal study on risk factors of grade retention among elementary school students using a multilevel analysis: Focusing on material hardship and lack of school engagement. *Children and Youth Services Review,* 88, 25-32. doi:10.1016/j.childyouth.2018.02.043

Yang, Y., Joshi, S. H., Jahanshad, N., Thompson, P. M., & Baker, L. A. (2017). Neural correlates of proactive and reactive aggression in adolescent twins. *Aggressive Behavior, 43*(3), 230-240. doi:10.1002/ab.21683

Zhang, C., Bingham, G. E., & Quinn, M. F. (2017). The associations among preschool children's growth in early reading, executive function, and invented spelling skills. *Reading and Writing, 30*(8), 1705-1728. doi:10.1007/s11145-017-9746-0

Appendix A: Power Analysis using G*Power



Appendix B: Administrative Procedure – Approval to Conduct Research

ADMINISTRATIVE PROCEDURES MANUAL
Personnel

[redacted]

5:100-E1 Administrative Procedure – Approval to Conduct Research

APPLICATION TO CONDUCT RESEARCH

Applications to conduct research that involves students, parents, or staff of [redacted] will be reviewed by Office of Data and Accountability. Research guidelines incorporated in this application are designed to protect the confidentiality of human subjects and guarantee the integrity and quality of any research conducted in the district. In addition, proposed research cannot be conducted during state testing periods, must not violate state education codes related to privacy, may not create a data burden on teacher or schools, and is entirely voluntary on the part of the participants.

This application will ensure that your proposal is properly aligned with current District policy regarding human subjects and the District research priorities. If the researcher is a graduate student, we require a supporting letter from the graduate advisor.

Please complete the following and attach clearly labeled additional pages as needed. Please allow at least 10 business days for a response from the District. All approved research must be conducted under the supervision of the school principal or other designated administrator.

5:100-E1 Administrative Procedure – Approval to Conduct Research

Part 1 (Study type)
Respond to one of the three depending on your research:

A: External Researcher

Study Title:	Researcher Name(s):
Address:	City, State, Zip:
Main Phone:	Alternative Phone:
Email:	Name of Organization:
Timeline:	
Places where Data will be Collected:	
What tasks/activities will subjects be asked to complete, with an estimate of amount of time it would take to complete (such as filling out a survey, participating in an interview, observing a classroom): **Please attach a copy of an instrumentation to be used. This includes, but is not limited to, surveys, tests, consent forms, and data recording sheets.**	
Significance of Study to [redacted]	Significance of Study to Field of Education:

B: Graduate Study (Thesis or Dissertation)

Graduate Advisor Name:	Name:
Title:	Telephone Number:
Department:	Email:

Please attach the following:
1. Copy of the research proposal.
2. A letter from the graduate advisor denoting approval of the thesis or dissertation.

C: Graduate Study (Course assignment or other related graduate school project)[redacted]

Name: Dan Palombit	School and/or Course: American College of Education; Ed.D. in Educational Leadership
Email: danpalombit@gmail.com	Telephone Number: (847) 769-6746

Please attach the following:
1. Copy of the assignment/project.

[redacted]

5:100-E1 Administrative Procedure – Approval to Conduct Research

Part 2: Subject Information

Reply to the components that are associated with the nature of your study. Write "not applicable" for components that are not associated.

Description of Participants to be involved: Data will be collected in the form of a scaled score from the 2012-2013 school year's administration of the ISAT Reading assessment; during this school year, all students were in the seventh grade cohort. Data will be collected in form of office discipline referral category and occurrence for the 2014-2015 and 2015-2016 school years for the same cohort of students who would be in ninth and tenth grades respectively. Data will be collected in form grade retention occurrence for the 2014-2015 and 2015-2016 school years for the same cohort of students who would be in ninth and tenth grades respectively.	Number of Participants to be Involved: not applicable A total of 100 seventh grade student records will be retrieved from the school database. Data will be representative from each of the three [redacted] feeder middle schools.
Potential Risks/Hazards to Subject: not applicable	Expected Benefits to Subject: not applicable
Proposed Reward/Incentive to Subject: not applicable	How will Subjects be Recruited/Selected: not applicable
How will you assure participation is voluntary? not applicable	What provisions will be made for subjects not willing to participate? not applicable
How will you ensure that the student information remains confidential? All data will be anonymous and confidential. No personal identifying information will be collected. Pseudo codes, such as Student #1, will be used to tag all the participants. Hard copies of raw data and other documents pertinent to the study will be securely kept in a locked filing cabinet inside the personal office of the researcher. Soft copies of raw data and other documents will be saved in a password-protected flash drive. All data and documents related to the study will be destroyed five years after completion. Hard copies will be shredded while soft copies will be deleted. The school district will be referred to as Northwest Suburban Chicagoland school district. No specific school name will be used.	

ADMINISTRATIVE PROCEDURES MANUAL
Personnel

[redacted]

5:100-E1 Administrative Procedure – Approval to Conduct Research

If the research project is associated with an institution that requires a formal human subject's review, a copy of the human subjects review committee approval is required prior to final approval.

I certify that this completed research application is an accurate and complete statement of the nature of my research. I further agree that this research does not involve coercion, deception, or psychological manipulation of any [redacted] participant.

Researcher Signature — April 16, 2019 — Date

ADMINISTRATIVE PROCEDURES MANUAL
Personnel

[redacted]

5:100-E1 Administrative Procedure – Approval to Conduct Research

STATEMENT OF AGREEMENT FOR RESEARCHERS

Title of Research Project: A Quantitative Correlational Investigation of Seventh Grade Reading Scores as a Predictor of High School Office Discipline Referrals

The proposed research activities to be conducted in [redacted] are in compliance with existing legal and ethical codes. The research will not differ significantly from the activities described within the proposal. Any amendments to the original proposal must be submitted to and approved by the Office of the Associate Superintendent. All participation in the study will be voluntary and confidentiality of the data will be maintained. All researchers agree to provide the Office of Data and Accountability of [redacted] with a copy of the final research report. Researchers agree to ensure that all associates, colleagues, and employees assisting in the conduct of the study are informed about their obligations in meeting the district research study commitments.
I understand and agree with the above statement and will follow the guidelines it sets forth.

April 16, 2019
Date

Dan Palombit
Printed Name

Signature

Printed Name (of additional researcher)

Signature (of additional researcher)

FOR OFFICE USE ONLY

4/29/19
Date

[redacted] Research and Assessment
Approved by **Title**

Please mail or email the signed document to:
Office of the Superintendent
[redacted]

[redacted]

Appendix C: School District Approved Approval to Conduct Research Form

Daniel Palombit

From:	Ehrmann, Joseph
Sent:	Monday, April 29, 2019 8:11 AM
To:	Palombit, Dan
Cc:	Patricia Slowik
Subject:	Data Request for Dissertation - Approval
Attachments:	D.Palombit Approval Request.pdf

Dan:
Attached is the signed approval request form. Patty Slowik is ready to go with the anonymous data set when you are. Good luck as you keep progressing,

--
Joseph Ehrmann Ed. D.
[redacted]Research and Assessment Coordinator