2020

# Challenging The Core Assumption Of Chronic Absenteeism: Do Excused And Unexcused Absences Equally Contribute To The Effective Early Identification Of Students At Risk For Future Achievement Problems?

Cassandra Henderson
University of Pennsylvania

Follow this and additional works at: https://repository.upenn.edu/edissertations

Part of the Education Policy Commons

# Challenging The Core Assumption Of Chronic Absenteeism: Do Excused And Unexcused Absences Equally Contribute To The Effective Early Identification Of Students At Risk For Future Achievement Problems?

Abstract

In response to the Every Student Succeeds Act (2015), nearly three-fourths of states in the U.S. have adopted chronic absenteeism—defined as missing 10% of the school year—as a measure of school quality and student success (Jordon, Fothergill, & Rosende, 2018). Due to its widespread adoption and the strong predictive relationship between early absences and negative educational outcomes, chronic absenteeism is increasingly being utilized by schools as an early warning indicator of later problems, such as low academic achievement. As such, chronic absenteeism theoretically allows schools to identify academically at-risk students in the early primary grades using readily available attendance data and provide them with additional resources to prevent later difficulties (Balfanz, Herzog, & Mac Iver, 2007). Given its pervasive use as both an accountability metric and an early warning indicator, the need to ensure the scientific integrity of chronic absenteeism is vital. Major theoretical assumptions underlying this indicator, however, have never been empirically validated.

The current study represents the first effort to scientifically test the most basic assumption upon which chronic absenteeism is based—that all absences from school (i.e., both excused and unexcused absences) are equally detrimental to student outcomes and should be utilized to identify at-risk students. The purpose of this study was thus to test whether excused and unexcused absences have comparable diagnostic accuracy in the early identification of academically at-risk students. Using the state-of-the-art receiver operating characteristic (ROC) methodology, this study presented evidence that only unexcused absences provided diagnostic accuracy for academic risk status in math and English achievement for an entire cohort of young students in Philadelphia. This diagnostic accuracy was evident in kindergarten and increased across the early elementary years. Excused absences, on the other hand, provided no diagnostic utility in differentiating between students at risk for academic problems and students on track for success within and across the early elementary grades. The findings presented here indicate that chronic absenteeism could be a more effective early warning indicator for students in large urban school districts by taking absence types into account. These results have further implications for researchers and policymakers, surfacing the need to prioritize additional empirical studies testing the underlying assumptions of chronic absenteeism.

Subject Categories
Education Policy

CHALLENGING THE CORE ASSUMPTION OF CHRONIC ABSENTEEISM:

DO EXCUSED AND UNEXCUSED ABSENCES EQUALLY CONTRIBUTE TO THE

EFFECTIVE EARLY IDENTIFICATION OF STUDENTS AT RISK FOR FUTURE

ACHIEVEMENT PROBLEMS?

Cassandra M. Henderson

A DISSERTATION

in

Education

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2020

Supervisor of Dissertation:

_____

John W. Fantuzzo, Albert M. Greenfield Professor of Human Relations

Graduate Group Chairperson:

_____

J. Matthew Hartley, Professor of Education

Dissertation Committee:

John W. Fantuzzo, Albert M. Greenfield Professor of Human Relations

Paul A. McDermott, Professor of Education

Michael A. Gottfried, Professor of Education

CHALLENGING THE CORE ASSUMPTION OF CHRONIC ABSENTEEISM: DO
EXCUSED AND UNEXCUSED ABSENCES EQUALLY CONTRIBUTE TO THE
EFFECTIVE EARLY IDENTIFICATION OF STUDENTS AT RISK FOR FUTURE
ACHIEVEMENT PROBLEMS?

ACKNOWLEDGMENTS

While my name stands alone below the word copyright, this dissertation represents the invisible work of so many others. I am incredibly humbled and grateful to have the support of each person that has helped me achieve this milestone.

To my dissertation chair, mentor, and friend—John Fantuzzo. How can I say a mere "thank you" for all you have done to contribute to my development as a scholar and a human? Your work may be the reason I came to Penn, but your mentorship is the reason I have stayed. In your child development course that I took during my first year, you told our class that we can't expect great things from people without first caring for them unconditionally and respecting their individual needs—that people can't thrive without being nurtured and seen for who they are. Thank you for providing me with the unconditional support and respect that I needed to be able to expect great things from myself. Your model of mentorship recognizes that the scholar and the person are one and must develop in tandem—that knowledge of an academic subject remains an unfulfilled contribution if you have never connected it to your understanding of your own purpose in life. At the start of my journey, I knew only a portion of myself; here at the end, I can confidently say that I am an expert. I have connected my purpose and passion in life to the work I get to do every day. What an incredible gift you have given me, for which a "thank you" will not suffice. I know the greatest thanks I can give you is to live all that you have taught me and pass on that wisdom to the next generation of scholars. I promise I will. Excelsior!

To the other members of my committee—Michael Gottfried and Paul McDermott. Michael, I reached out to you as an expert in the field, hoping I might get a minute of your time over email. Within a week, I was sitting with you over coffee talking for hours about every research idea in my head. Your joy in mentoring and collaborating with students was so apparent from that first meeting and is clearly part of the reason you are so enthusiastic about your work. Thank you for your expertise, insight, openness, and humor throughout this process. I look forward to many collaborations in the future. Paul, I learned about ROC analyses in your class. Without that introduction, I would not have known this methodology existed (and should be more widely used in education!). I credit your teaching with my understanding of all the advanced methodology (and SAS code) I learned during my graduate training. Thank you for these skills and thank you for always providing me with methodological guidance and consultation when I needed support.

To my colleagues and friends. Katie Barghaus—for almost seven years, you have been my model of not only thoughtful and meaningful scholarship, but of how that rigorous scholarship can exist alongside kindness, humility, and integrity. You expertly balance so many competing demands and still find time to be a caring mentor to everyone on your path. You have also been a model for me of how to balance a career and a family, of how to try to exist fully and presently in the role of both researcher and mother. While I know (now more than ever) that it isn't easy, you are an example of how it can be done with grace. Thank you for being my role model and guide. I can't wait to continue building great things alongside you at Penn Child. Kristen Coe—my fellow traveler on the long, strange trip that is earning your Ph.D., we have shared many of the

same experiences throughout this journey—both the highs and the lows. Thank you for

always being an emotional support and thought partner for me during a process that is so

emotionally and intellectually difficult. I don't know how I'll survive without our

marathon chats about work, life, and everything in between. While I am sad to be losing

you as a coworker, I am so excited to watch you make major contributions to both

research and practice. Your many talents and kind heart are desperately needed to make

the world a better place. And something tells me you won't be too far away, especially

when it comes time for me to repay you with postpartum groceries! Whitney LeBoeuf—

while I no longer get to work with you on a daily basis, you helped shape me into the

person I am today during the time I was fortunate enough to spend with you. Interactions

with you were always learning opportunities, whether you were explicitly teaching me

research skills (seriously, HUD was like a second dissertation) or whether I was

observing and internalizing the many implicit ways you are so successful in all that you

do. Your candor, warmth, and humor made me feel like I had found my home at Penn

Child. Thank you for helping to blaze the trail that I am now walking. To my broader

academic and CUCMC family—I feel privileged to be among you. I am incredibly

grateful to be connected to so many wonderful, supportive scholars that share similar

values not only in the kind of work they do but in the way they do that work. Thank you

for embracing me as part of your tribe. And to my fellow Ed. Policy students—I have

watched you all transform the division for the better. You have made it your mission to

connect with and support one another to grow our collective sense of community. I am

proud to be associated with so many talented people. Thank you for your support and friendship.

To my teacher and friend—Peter McVeigh. You have changed generations of people for the better, through your care, compassion, and desire to effect positive change in this world. I know I am only one of many whose life was indelibly altered by knowing you. It seems incredibly unfair that I won't be able to share this acknowledgment with you, just as I shared my application essay for the program with you seven years ago. That essay outlined how you helped shape my life and the type of work that I want to do every day. Those sentiments are more true now than ever. You are missed and loved by so many, who will continue to work to make the world a better place because of what you taught them. Thank you for everything, buddy.

To my parents. What can I possibly say? There are no people in the world that have done more to get me to the place I am today. And while I know that can be said by most about their parents, I also know that mine have done so much more than most. Whether it was proofreading my papers in high school or driving me to and from the Trenton train station every morning for an internship, you have always put my success above your own. I know my dreams have not always been linear—how does one go from South Africa, an English degree, movie production in Manhattan, and nonprofit work to a Ph.D. in education?—but you have done everything in your power to support them, even when they were strange or unrealistic. You have always been my biggest fans, more convinced than I ever was that I had a contribution to make to the world. Knowing what I know now about how children grow and develop, I know that the only reason I have been

able to spread my proverbial wings and fly is because you convinced me long ago that I could. My gratitude feels such a meaningless gift to offer you in return for all that you've done. But it's all children have to give their parents. That and the pride of seeing the person your child has become. I hope I make you proud of the person that I am, especially as I become a parent to my own child. I hope you see all that I am willing to sacrifice for my daughter and know that is your legacy. I love you.

To my older sister—Alicia. I truly would never have applied for this program had it not been for your encouragement. During a walk on the beach, when the thought of getting a Ph.D. was still just a vague notion, I expressed doubt about my ability to even get into a doctoral program, convincing myself that I didn't have the background or experience necessary to be considered a worthy applicant. You stopped my nervous rambling and looked at me incredulously saying, "Are you serious? You know you can do this." You have always loomed so large for me, not just because you are my older sister but because you have accomplished so much in your life. There was something about hearing your confidence in me that day that made me have confidence in myself. Thank you for always setting the bar so high and for making me believe that I could reach it too. I love you and can't wait to see our children become the second wave of awesome Henderson girls!

To my grandparents, Grandmom and Poppie. Grandparent doesn't adequately capture what you both have been to me. Most people see their grandparents a few times a year, if they are lucky. Mine were a daily presence. And when you weren't busy cooking dinner, folding clothes, or driving us around, you were encouraging us to do our

homework, praising us for reading, and emphatically asserting that we were the smartest children in the world. While I know every grandparent believes that of their grandchildren, I was fortunate enough to be exposed to such unconditional confidence in my abilities every day. You sacrificed any semblance of a retirement to be there for us and not only did it with a smile but still assert that it was the best time of your life. I know that no one could possibly be prouder that I now get to call myself Dr. Henderson. I love you both.

To Colin—my husband and the best guy I know. We met at the end of my first year as a graduate student, and even though I hate the idea of "favorites", the day we met was one of the best days of my life. You changed everything from the second you walked into my world, so much so that after only a few months I couldn't imagine it without you. And I cannot imagine having completed my Ph.D. without you. Actually, I don't know how anyone could possibly earn a Ph.D. without you by their side. My mentor is fond of saying that unconditional positive regard is like psychological oxygen—that without love and acceptance, human beings can't breathe. Thank you for being my breath for these last six years, for seeing me and valuing me unconditionally since the moment we met. I am enough in my own eyes because I am enough in yours. You embody kindness, patience, and positivity. I know there were times during this process when I did not display these traits at all let alone to the same degree that you display them every day. And yet, you continue to bestow them upon me, however unworthy of them I may be. You are also the most self-motivated person I know, teaching yourself an entirely new career in your free time and then excelling in that profession in such a short period of time. Your work ethic

and commitment to bettering yourself helped to push me on days when I would have rather stayed in bed. You are an equal partner in this accomplishment, just as you are my partner and equal in everything we do. I promise to continue to be sickeningly in love with you for the rest of my life. I love you so much.

Finally, to my daughter—Jillian Elizabeth McNamara (also known as Jilly Bean, Beansie Boop, Boopsie, Boop, or The Booper). Nothing and no one in this world could have inspired me to finish my dissertation more than you. And that isn't just because I defended at 36 weeks pregnant. Wanting to make you proud was truly my best motivation. As monumental as earning my Ph.D. was, crossing the threshold into becoming your mother is by far the most significant and transformative step I have ever taken in my life. I know that our relationship as parent and child will be the most rewarding journey I will ever experience. Helping you become the person you are meant to be is the most important manifestation of my purpose on this earth. I loved you before you even existed, when you were only a dream and a hope of what would someday be. And now that you are here, I love you so much more than words could ever convey.

ABSTRACT

CHALLENGING THE CORE ASSUMPTION OF CHRONIC ABSENTEEISM: DO

EXCUSED AND UNEXCUSED ABSENCES EQUALLY CONTRIBUTE TO THE

EFFECTIVE EARLY IDENTIFICATION OF STUDENTS AT RISK FOR FUTURE

ACHIEVEMENT PROBLEMS?

Cassandra M. Henderson

John W. Fantuzzo

In response to the Every Student Succeeds Act (2015), nearly three-fourths of states in the U.S. have adopted chronic absenteeism—defined as missing 10% of the school year—as a measure of school quality and student success (Jordon, Fothergill, & Rosende, 2018). Due to its widespread adoption and the strong predictive relationship between early absences and negative educational outcomes, chronic absenteeism is increasingly being utilized by schools as an early warning indicator of later problems, such as low academic achievement. As such, chronic absenteeism theoretically allows schools to identify academically at-risk students in the early primary grades using readily available attendance data and provide them with additional resources to prevent later difficulties (Balfanz, Herzog, & Mac Iver, 2007). Given its pervasive use as both an accountability metric and an early warning indicator, the need to ensure the scientific integrity of chronic absenteeism is vital. Major theoretical assumptions underlying this indicator, however, have never been empirically validated.

The current study represents the first effort to scientifically test the most basic assumption upon which chronic absenteeism is based—that *all* absences from school

(i.e., both excused and unexcused absences) are *equally* detrimental to student outcomes and should be utilized to identify at-risk students. The purpose of this study was thus to test whether *excused and unexcused absences have comparable diagnostic accuracy in the early identification of academically at-risk students*. Using the state-of-the-art receiver operating characteristic (ROC) methodology, this study presented evidence that *only unexcused abse*nces provided diagnostic accuracy for academic risk status in math and English achievement for an entire cohort of young students in Philadelphia. This diagnostic accuracy was evident in kindergarten and increased across the early elementary years. Excused absences, on the other hand, provided *no* diagnostic utility in differentiating between students at risk for academic problems and students on track for success within and across the early elementary grades. The findings presented here indicate that chronic absenteeism could be a more effective early warning indicator for students in large urban school districts by taking absence types into account. These results have further implications for researchers and policymakers, surfacing the need to prioritize additional empirical studies testing the underlying assumptions of chronic absenteeism.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

CHAPTER 1: INTRODUCTION

Background

The educational system in the United States guarantees that all children have access to a public education. This tenet—that everyone is entitled to an education—was foundational to the vision of an American democracy in which all citizens were sufficiently educated to vote for their leaders and engage in the process of self-government. It also served as the basis for an American cultural vision and deeply held national beliefs about "the American dream"—that everyone has equal opportunity to benefit from their education and, thus, equal opportunity to succeed in life. Since the release of the Coleman Report in 1966, however, this vision has been challenged. As Coleman et al. most notably revealed, the United States faces considerable disparities in the educational opportunities and outcomes of children from low-income families and children from racial and ethnic minority groups (Coleman, 1966). This troubling finding that certain groups of children are experiencing considerable educational gaps has been confirmed by decades of research from across the nation (Fryer & Levitt, 2006; Phillips, Crouse, & Ralph, 1998). Even more alarming, research suggests that these educational gaps are evident in early grades and persist or grow over time (Duncan & Magnuson, 2011). For instance, there is evidence to suggest that students from low-income families generally enter kindergarten behind their more economically advantaged peers on measures of academic skills and other school readiness behaviors (Entwisle, Alexander, & Olson, 1997; Gottfried & Gee, 2017; Lee & Burkam, 2002).

The exposure of these educational gaps has led to decades of efforts to minimize or eliminate them. In fact, the U.S. has an established policy goal to decrease these achievement gaps (Berliner, 2009; Blankenship, 2015). Over the course of the last 50 years, many reform efforts have attempted to achieve this policy goal through a variety of mechanisms: school choice initiatives, such as charter schools and voucher programs; a focus on smaller class sizes and lower student-teacher ratios; top-down accountability systems and increased national attention on standardized testing outcomes; and school finance reform efforts and scrutiny of per pupil expenditures (Williams, 2003; Manning & Kovach, 2003). Despite the proliferation and variety of reform efforts, however, these educational gaps persist (Gershenson, Jacknowitz, & Brannegan, 2017). Some national research even indicates that these disparities have grown, with one study estimating that the gap between students from high- and low-income families has increased by roughly 40% in the last 25 years (Reardon 2011).

The pervasiveness and persistence of these educational gaps in spite of the cascade of efforts to ameliorate them has led researchers, practitioners, and policymakers to search for other potential contributing factors and underlying causes. As Gershenson et al. (2017) note, "Understanding the source(s) of the achievement gap is crucial to devising an appropriate policy response" (p. 138). It is thus imperative that the underlying causes of the achievement gap are elucidated so that suitable policy solutions can be crafted.

Until recently, attendance was understudied as an important educational input, was not used as a benchmark for accountability purposes, and was not often considered a

contributing factor to the achievement gap (Gershenson et al., 2017; Lamdin, 1996; Sheldon, 2007). As such, consistent student attendance was often an assumption upon which research, school reform efforts, and accountability benchmarks were based. As Balfanz (2016) states, "The U.S. education system is based on a comforting assumption. Absent an illness or the occasional family event, we tend to think that preK-12 students are in school every day. School district instructional pacing guides assume this, as do accountability systems and program evaluations" (p. 8). Nauer (2016) similarly asserts that "Educators and policy makers have historically overlooked absenteeism—an irony, given how much effort goes into improving schooling on the assumption that students are actually attending regularly" (p. 30). Many policy mandates and costly educational reform efforts devised in the last 50 years have thus relied on the supposition that students are generally in school every day.

The Attendance Gap

Despite the ubiquity of this assumption, recent national research has challenged it with the discovery that a large number of children are missing school. In the 2013-2014 school year, for instance, about 6.8 million students, or roughly 14% of the school-age population, missed 15 days (or three weeks) of school or more in the United States (U.S. Department of Education, Office for Civil Rights, 2016). Research from the 2015-2016 school year indicates that about 8 million students in the United States (U.S.) missed 15 days of school or more (Bauer, Liu, Schanzenbach, & Shambaugh, 2018), an increase of about 15% over the number reported in 2013-2014. The sheer number of students missing three weeks of school each year and the increase in students experiencing these absences

across years is cause for concern and suggests that the assumption that students are generally in school every day is inaccurate.

Of even greater concern, the subgroups of students missing the most school are the same subgroups of students being disproportionately affected by the academic achievement gap. While estimates differ by locality, national estimates indicate that students from low-income families, students in urban areas, Black and Latino students, English-language learners, and students with disabilities tend to miss more school days than their peers (Applied Survey Research, 2011; Chang & Davis 2015; Chang & Romero, 2008; Chen & Rice (2016); Gottfried & Gee, 2017; Gottfried, Stiefel, Schwartz, & Hopkins, 2019; Jacob & Lovett, 2017; London, Sanchez, & Catrechini, 2016; Spencer, 2009; U.S. Department of Education, Office for Civil Rights, 2016). Thus, the same demographic groups of children negatively affected by the achievement gap are also experiencing an "attendance gap".

In addition, this attendance gap seems to appear early in the educational lives of children, much as the achievement gap does. In fact, national attendance rates by grade show that students in the early grades have some of the highest frequencies of absence, with at least 10% of kindergarteners missing 15 days or more of school each year nationally (Chang & Davis, 2015). As Balfanz (2016) notes, "absenteeism rates are high in preK, kindergarten, and 1st grade" (p. 9). The attendance gap is not only evident at school entry much as the achievement gap is, it also persists across school years, as the achievement gap does (Bauer et al., 2018; Ehrlich, 2014; Mac Iver, 2010). Thus, students who exhibit attendance problems early are more likely to have attendance problems in the

future, just as children who exhibit lower academic achievement and school readiness behaviors in kindergarten tend to remain behind their peers as they progress through school.

Because this attendance gap seems to align with the achievement gap in terms of the groups of students it affects, how early it appears in the educational lives of students, and how persistent it is across time, researchers have begun to theorize that the attendance gap may underlie broader gaps in educational well-being, such as academic achievement. As Chang and Romero (2008) state,

> Student absences potentially contribute to the achievement gap in two ways. First, absence rates are higher among socioeconomically disadvantaged students, so such students are exposed to the potentially harmful effects of absences more often. Second, absences may cause greater harm to students who reside in socioeconomically disadvantaged households because such households may be less able to compensate for lost instructional time than their more advantaged counterparts. (p. 138)

The attendance gap may thus contribute to the achievement gap as students with high rates of absence, particularly students from low-income families, may not be exposed to the educational benefits of time spent in school and may not have the resources at home to make up for this loss of instructional time (Blazer, 2011; Chang & Romero, 2008; Claessens, Engel, & Curran, 2015; Gershenson et al., 2017; Hocking, 2008; Railsback, 2004; Sparks, 2010). Because this attendance gap may contribute to the achievement gap and has been largely ignored until recently, it has potentially "undermined school reform efforts of the past quarter century and negated the positive impact of future efforts" (Balfanz & Byrnes, 2012, pp. 3-4). It is therefore crucial for future educational policies to address the attendance gap if efforts to close the achievement gap are to succeed. As

Balfanz (2016) suggests, "one of the most effective strategies for closing the achievement gap will be a concerted effort to ensure that high-poverty students attend school regularly from preK through 12th grade" (p. 10).

In order to address the achievement gap in new, more cost-effective, more readily scalable ways, it is therefore necessary to understand how attendance may be utilized as a lever to reduce educational inequities. Before potential policy responses related to attendance are discussed, the research evidence linking the attendance gap to student academic performance must be reviewed. Because children in the earliest grades have the lowest attendance rates, patterns of attendance are established early, and these patterns tend to persist across time, it is critical to focus on attendance in early elementary school as a potential cause of the achievement gap (Alexander, Entwisle, & Kabbani, 2001; Spencer, 2009). The next section will review the theory of why early attendance is important for student success and why early absences may be detrimental to student outcomes. The following section will summarize the empirical literature relating both school attendance and absences to educational outcomes.

A Conceptual Framework Linking Early Attendance and Absences to Educational Outcomes: A Review of the Theoretical and Empirical Research

This section will review the conceptual framework linking early student attendance and early student absences to educational outcomes. Because attendance and absences are complementary events, it is important to consider the effect of each on student outcomes. Attendance and absences are often treated and studied as separate phenomena, however, and much of the research literature reflects this bifurcation. This section will explore the theoretical and empirical research base for both attendance and

absences as they relate to educational outcomes in order to bridge this gap and present a more comprehensive understanding of how these complementary events simultaneously contribute to student well-being.

Early Attendance as a Protective Factor: Theoretical Framework

There is relatively little theoretical dialogue around the importance of attendance. In their recent paper, Gottfried and Gee (2017) note that research about student attendance "has been largely atheoretical" (p. 2). The authors, however, invoke a useful theoretical framework in understanding how attendance may be important in positively affecting student outcomes—the developmental-ecological model of human development. The developmental-ecological model states that human development is driven by the interaction of person, context, and time. Intra-individual characteristics transact with various contexts at various levels of proximity, and these transactions occur across time to create individual developmental pathways (Bronfenbrenner, 2015; Bronfenbrenner & Morris, 1996). The most influential drivers of human development are the proximal processes or direct interactions that occur between the developing child and other people, places, and things in his/her immediate environment.

At school, for example, the classroom represents a direct context or microsystem in which the child engages in critical proximal processes. As Kearney and Graczyk (2014) note, "attendance provides youth a setting for academic development, a language-rich environment, opportunities to develop social competence and relationships, and experiences that nurture work-related skills such as persistence, problem-solving, and the ability to work with others" (p. 2). Within the school classroom, the child is exposed to

range of inputs—from direct educational instruction to social interactions to norms and routines.

Furthermore, the information and skills children are exposed to within the classroom context are cumulative and progressive. Children must have consistent contact with the classroom context in order for knowledge and skills to develop and accrue. For example, children must learn number recognition, counting, and cardinality before they can learn addition and subtraction. A student who is routinely present in the classroom will be consistently exposed to more basic skills, like number recognition, first and will master them before moving on to more complex skills, like subtraction. Additionally, consistent attendance in the early grades and the resulting exposure to classroom content can create a foundation for later educational success (Coelho et al., 2015). Early grades act as a portal to the public education system and are therefore crucial for creating lasting routines around school and for establishing the foundational skills upon which later learning will be based (Gottfried & Gee, 2017; Hickman & Heinrich, 2011; Kagan & Kauerz, 2006). For instance, if a child consistently attends school in the early primary years, he/she will get into the "habit" of attending, and this behavior will become routinized and be more likely to endure in the future; similarly, a child exposed regularly to more basic knowledge and skills because of his/her consistent attendance will be more likely to have a firm grasp of those skills and be better equipped able to learn complex skills in the future. Routine attendance in the early elementary years can, therefore, help to establish positive educational trajectories that persist as the child develops.

In sum, attendance provides the developing child with a host of developmental opportunities within the microsystem of the classroom, and these opportunities tend to build upon each other in complexity. The amount of time the developing child is actually present in school, thus, directly affects the extent to which that child is able to benefit from the important proximal processes that occur within the classroom and the degree to which the child can establish positive educational trajectories that can persist into the future (Bauer, et al., 2018; Dougherty, 2018).

Early Absence as a Risk Factor: Theoretical Framework

The developmental-ecological model outlined above is similarly useful for understanding absences as a potential risk factor to educational success. Because the classroom is a vital developmental microsystem for children, school absences have potentially negative consequences (Gershenson et al., 2017). The proximal processes to which children are exposed in the classroom provide the basis for academic learning, social growth, and the acquisition of important learning behaviors (Huston & Bentley, 2010). It follows that being absent from this environment lessens the degree to which the developing child is exposed to those important processes and able to gain these vital competencies. In cases where absences are infrequent and/or the home context is able to provide similarly rich experiences and resources as the classroom environment, the developmental risks of missing school are relatively low. In cases where absences are frequent and/or the home environment does not have resources available to compensate for lost time in the classroom, the effects of missing school may be more pronounced (Chang & Romero, 2008). Considering that many of the children missing the largest

number of school days nationally are from low-income families, it is possible that they face disproportionate risk from school absences due to fewer resources at home (Gonzales, Richards, & Seeley, 2002; Spencer, 2009; Teasley, 2004). This could also account for the effects observed in Ready's study (2010), where attendance seemed to have more beneficial effects for students from low-income families.

Furthermore, within a developmental-ecological model, the negative effects of absences are cumulative over time and may thus be particularly detrimental to young students. These cumulative effects can build and negatively alter developmental pathways especially during the early years of schooling. Coelho et al. (2015) state:

> Throughout the early elementary years, students gain the social and academic skills that are essential to their educational achievement. The learning and attainment of these skills occurs during a critical period of development in the child's life. Disturbances or delays in a child's learning in early years can ripple across their progress, as they attempt to build new knowledge and skills upon more basic iterations, and ultimately alter their life course trajectories on several measures of well-being. (p. 8)

As the developmental competencies learned within the classroom accumulate and grow progressively more complex, prolonged or habitual absence can have substantial negative impacts on children's educational outcomes, especially during the early years of school. For example, if a child has missed a significant number of school days during which the teacher has engaged students in mastering alphabet recognition, the child may struggle more with that competency than children who have consistently attended. If, once the child returns to school, the teacher has moved onto word recognition and other early literacy skills that rely on a firm grasp of letter recognition, the absent student may struggle to acquire this new competency because of a lack of exposure to and mastery of

the previous competency. This could then engender a sense of discouragement within the child as they are unable to acquire the new skills and start to fall behind their peers. As Hickman and Heinrich (2011) state, "When [habitually absent] students do return to school, they find themselves academically behind their peers as a result of missing educational guidance and instruction. As a result of their educational gaps, these students begin to disengage further from school as they realize they are simply too far behind to catch up to their peers" (p. 43). Furthermore, early absences can also become habituated in young students; if absences are routine in the early primary years, it is more likely that they will continue in the future. The negative effects of missing school therefore compound over time as the developing child falls further behind his/her peers, feels disengaged in the classroom, and establishes an "absence habit" that can persist in the future.

Thus, from a theoretical perspective, absences remove children from a crucial microsystem in which they are exposed to a variety of vital developmental experiences. The knowledge and skills that children miss when they are not present within this microsystem are cumulative and grow progressively more complex. Therefore absences, particularly among young children, are not singular, contained events but can reverberate into the future and negatively affect young children's developmental trajectories.

Early Attendance as a Protective Factor: Empirical Literature Relating Elementary School Attendance to Positive Educational Outcomes

The empirical, peer-reviewed literature relating early attendance to positive educational outcomes is relatively limited. Traditionally, the research literature has not

recognized student attendance as an important educational input that may affect

children's outcomes and even less attention has been paid to the effects of early school

attendance (Balfanz, 2016; Gershenson et al., 2017; Sheldon, 2007). The first peer-

reviewed, published study attempting to show a relationship between early attendance

and educational success appeared in the 1990s. Caldas (1993) performed an observational

study using administrative records from 737 public elementary schools in Louisiana in

1989. These records were aggregated at the school level to reflect average attendance and

average student achievement scores at each school. The study examined a step-wise

regression model to determine the relationship between average school attendance rates

and average student achievement on a composite index of state standardized achievement

tests across a variety of subjects, while controlling for a number of demographic

covariates (e.g., number of Black students per school, percentage of students receiving

free and reduced-price lunch, etc.). While the observational nature of the study and the

unit of analysis (i.e., data aggregated at the school level) undermine the ability to draw

causal conclusions from this study, the author found that there was a statistically

significant, positive relationship between average school attendance and average school

achievement. In short, higher rates of school attendance predicted higher rates of student

achievement in elementary schools.

Following Caldas, Lamdin (1996) and Roby (2004) each published similar

studies. Lamdin (1996) looked at 97 Baltimore public elementary schools in the 1989-

1990 school year. Again, the data were observational in nature and aggregated at the

school level, limiting the strength of the study and possible causal conclusions that can be

drawn. Utilizing the econometric framework of educational production functions, Lamdin performed a regression analysis looking at the relationship between average school attendance and average performance on the on California Achievement Test (CAT) at the school level, controlling for similar demographic covariates. Lamdin found that average elementary school attendance rate was positively related to average school achievement in math and reading. Similarly, Roby (2004) looked at administrative data for about 2,000 public elementary schools. Roby related average daily attendance rates per school to average school performance on the Ohio State Proficiency Test, a standardized achievement test. Unlike Caldas and Lamdin, Roby only looked at correlations between these two variables and did not control for demographic characteristics, making these findings weaker than the previous two studies. Nevertheless, Roby found that there was a moderate to strong relationship between average school score on the standardized state achievement exam and the average school attendance rate. Furthermore, Roby compared the average achievement scores of schools that had the highest and lowest 10% average daily attendance rates. The achievement differences between these two groups were statistically significant for elementary schools. Overall, the results of these three studies provided preliminary evidence that there is a relationship between early attendance and academic achievement.

Following these studies, Ready (2010) and Gottfried (2010) used more advanced research methods and more detailed data sources to establish the link between early attendance and student achievement. Ready (2010) used a nationally representative dataset—the Early Childhood Longitudinal Study, Kindergarten Cohort (ECLS-K) from

1998-1999—to explore the relationship between academic skills and early attendance. Ready was able to make use of individual, child-level records for about 13,000 students, rather than aggregating results at the school level as the previous three studies did. With this more fine-grained unit of analysis, Ready's study was also able to use a more advanced methodology to determine the effect of attendance on achievement—hierarchical linear modeling (HLM) within a three-level, growth-curve framework. This methodology allows for the removal of potentially extraneous sources of variance, such as the school or classroom context, and better isolates the effect of attendance on the individual. Additionally, these data are nationally representative and thus allow for more generalizability than the first three studies, which focused on specific localities. Controlling for a variety of demographic covariates, Ready showed that there is an effect of attendance on the growth of children's cognitive skills in kindergarten and first grade. While the effects were stronger for literacy than math, both results were statistically significant. Furthermore, Ready found that attendance had a stronger effect on growth in cognitive skills among students from low-income families than students from high-income families. For instance, low-income children with high attendance rates see more gains in literacy skills than their peers from high-income families with similar rates of attendance. This suggests that the protective effects of attendance may be greater for those children who do not have access to resources at home (i.e., children from low-income families).

While Ready's study provided stronger evidence of a relationship between attendance and achievement than the previous aggregate studies, causal inferences about

the relationship of attendance and achievement were still tenuous. Gottfried (2010), however, utilized a quasi-experimental research design to better explore the causal effects of attendance on achievement. Gottfried's study consisted of an administrative dataset of multiple cohorts of elementary children in the School District of Philadelphia from 1994 to 2001. The study examined the relationship between attendance days and two measures of academic skills—The Stanford Achievement Test Ninth Addition (SAT9) and grade point averages (GPAs)—at the individual level. Gottfried accounted for a host of demographic variables and controlled for prior student achievement, which provides a better understanding of how attendance actually influences achievement within a given year by accounting for previous academic abilities. Most significantly, the study utilized an instrumental variable approach to better estimate the causal relationship between attendance and achievement. Gottfried used the geographic distance that the student lived from the school (determined through home address information) as an instrumental variable, as geography could reasonably affect a student's attendance (e.g., students who live closer to school tend to have higher attendance rates) but should not reasonably affect a student's academic performance (e.g., students who live closer to school do not have higher achievement scores). Geographic distance from school may have an indirect effect on achievement, however, through the effects of attendance; for example, if students who live closer to school attend more frequently and attendance causes higher achievement, then geographic distance from school would be related to achievement through the effects of attendance. The results of the study show that there is a statistically significant, positive relationship between attendance and both measures of academic

achievement (i.e., standardized test scores and GPAs) for elementary school students even when controlling for demographic variables and prior academic achievement. Moreover, there is a statistically significant relationship between geographic distance from school and both measures of achievement, suggesting that geographic distance is indirectly related to achievement through the causal effect of attendance on achievement. Gottfried's study thus provides the most compelling evidence that attendance and achievement are not only positively related but that higher levels of attendance may actually cause higher levels of academic achievement.

Though the empirical literature relating early attendance to educational outcomes is relatively sparse, the evidence across these studies is consistent—student attendance is positively related to achievement. Thus, students who attend school more frequently generally have higher levels of achievement and, it appears, higher attendance may be a cause of higher achievement. Additionally, these findings comport with the theoretical framework outlined above and suggest increased exposure to the school and classroom microsystems (i.e., higher attendance) has positive developmental effects on students. In sum, the theoretical and empirical literature support the supposition that the school attendance promotes academic achievement.

Early Absence as a Risk Factor: Empirical Literature Relating Elementary School Absences to Negative Educational Outcomes

The empirical, peer-reviewed literature around absences as a risk factor to educational outcomes is more pervasive than the literature exploring attendance as a potential protective factor for student success. Typically, however, the research around absences and educational outcomes has focused on older children, such as adolescents,

and relates school absences to risky social behaviors and outcomes, such as drug use, school dropout, and teen pregnancy (Kearney, 2008; Kearney & Graczyk, 2014; Rumberger & Thomas, 2000). There has even been some international, longitudinal research linking middle and high school absenteeism to negative outcomes in adulthood, such as unemployment, marital issues, and psychiatric problems (Broadhurst, Patron, & May-Chahal, 2005; Hibbet, Fogelman, & Manor, 1990).

Until recently, less attention has been paid to early absences as they relate to negative educational outcomes. In the last decade, several peer-reviewed studies relating elementary school absences to academic achievement outcomes have emerged. Morrissey et al. (2014) conducted a study utilizing data from the Miami School Readiness Project, a large-scale, longitudinal study of over 40,000 children in Florida from 2003 to 2007. The purpose of the study was to determine if student absences were the main mechanism through which low student socio-economic status affected achievement. Absence and achievement information (i.e., subject-area grades and performance on the reading and math portions of the Florida state achievement test) were obtained from administrative school records for children in kindergarten through fourth grade. The authors utilized a random-effects model to address omitted variable bias and a within-child, fixed-effects model to account for unmeasured characteristics of the child and family. While the study found that absences only partially attenuated the relationship between academic achievement and receipt of free or reduced-price lunch (rather than acting as the main mechanism of low-income status on achievement), results from both models indicate that higher absences within a given year are associated with lower grades and lower scores on

standardized reading and math tests within that same year. While the effects of absence

on achievement are clear, the question of whether absence is the main mechanism

through which family poverty affects achievement remain unclear, as this study found

little association between the number of absence days and the free or reduced-price lunch

status of the child, a finding inconsistent with previous research (Applied Survey

Research, 2011; Chang & Davis 2015; Chang & Romero, 2008; Chen & Rice (2016);

Gottfried & Gee, 2017; Jacob & Lovett, 2017; London et al., 2016; Spencer, 2009; U.S.

Department of Education, Office for Civil Rights, 2016). This discrepancy with previous

research may have been caused by the use of free or reduced-price lunch status as a proxy

for family socioeconomic status. Past research in education has found that the designation

of free or reduced-price lunch in educational records serves as a weak indicator of family

socioeconomic status (Harwell & LeBeau 2010). Furthermore, the study did not look at

the predictive effects of absence across years (e.g., kindergarten absences predicting

fourth-grade math achievement, or kindergarten through third grade absences predicting

fourth-grade reading achievement), which may be a better indication of the strength of

the effects of school absence on achievement. Nevertheless, the authors still found a

relationship between school absence and achievement indicating that increased absences

are associated with lower academic achievement scores from kindergarten through fourth

grade.

Using a more sophisticated research design, Gottfried (2011b) studied the effects

of absences on student achievement using a sub-sample of about 7,000 siblings in public

elementary schools within the School District of Philadelphia from 1994 to 2001. The

author identified students living in the same households within a given academic year through home address information in administrative records. The study also included a number of rich covariates (i.e., teacher level of education, teacher race, class size, prior standardized test scores, and community variables like percent living in poverty and percent home vacancy) in addition to the typical demographic variables used in educational research (e.g., student race, gender, etc.). Gottfried assessed the effects of absenteeism on both the reading and math portions of the SAT-9 (Stanford Achievement Test-Ninth Edition) using a family fixed-effects model to account for unobserved family factors that may affect both absences and achievement. The study found that absences have a negative effect on both reading and math achievement and, more importantly, that the magnitude of the effect actually increased when accounting for unobserved family characteristics. Thus, previous studies that have not employed a family fixed-effects approach may actually be underestimating the negative effects of absences on achievement. Despite these promising findings, there are potential issues associated with relying on home address information in school district records. This information can be somewhat unreliable, especially given the high level of residential mobility experienced by children in urban settings like Philadelphia (Chang & Romero, 2008; Ready, 2010; Romero & Lee, 2008). This could have potentially caused some misidentification of the sibling subsample, as address information from school records was the only criteria used to create a sibling "match." Nevertheless, Gottfried's findings are consistent with other studies on student absences and indicate that the effects of student absences on achievement may be greater than previously estimated.

In a second paper, Gottfried (2013) leveraged this same dataset to perform a quasi-experimental study. Similar to the analytic design utilized to assess the effect of attendance on achievement, Gottfried employed an instrumental variable approach to better estimate the causal effects of absence on achievement (i.e., SAT-9 math and reading scores). The instrumental variable utilized in this study was number of school nurses per school; because there is substantial research indicating that students with health issues miss more days of school and that schools with health-related services and professionals have better rates of attendance, it is reasonable to assume that there is a relationship between number of school nurses and student absences (Allen, 2003; Guttu, Engelke, & Swanson, 2004; Romero & Lee, 2008). Conversely, there does not seem to be a direct relationship between number of school nurses and student achievement; however, there may be an indirect effect of school nurses on achievement through the effect on student absences. Using this quasi-experimental design along with a school fixed-effects model to account for unobserved heterogeneity among schools, this study showed that there is an effect of student absences on standardized test scores and that the effect may be causal. This is consistent with Gottfried's previous quasi-experimental findings about the seeming causal effects of attendance on student achievement (Gottfried, 2010).

In addition to these peer-reviewed studies examining the effects of early absence on achievement, there are two non-peer-reviewed studies worthy of note. The first was published in a working paper of the National Bureau of Economic Research. In this study, Goodman (2014) also utilized a quasi-experimental analytic design to estimate the causal effects of absence on standardized reading and math test scores using data about

1.5 million students from the state of Massachusetts from 2003 to 2010. Similar to

Gottfried, Goodman employed an instrumental variable approach, using snowfall

amounts as the instrument (e.g., heavy snowfalls may induce school closures which affect

all students equally in that all students lose a day of school, whereas moderate snowfalls

may only affect certain students' ability to attend school and cause differential absences).

Goodman showed that each absence induced by moderately inclement weather reduces

elementary school math achievement by .05 standard deviations. By comparison, heavy

snowfalls, which induced school closures, showed no relationship to student

achievement. By showing that there is a relationship between snowfall amounts and

student achievement, presumably through the effects of student absence, Goodman

reaffirmed Gottfried's findings suggesting that there is a causal relationship between

absences and lower achievement.

The other non-peer-reviewed study of note comes from Attendance Works, a

national leader in attendance research and advocacy. This study bears mentioning as none

of the peer-reviewed studies on early absence have utilized nationally representative data,

which limits claims of generalizability. Ginsburg, Jordan, and Chang (2014) utilized

national data from the 2013 National Assessment for Educational Progress (NAEP), a

state-by-state analysis of achievement that is generally considered to be an important

indicator of national educational achievement. Absence data came from student self-

report and only consisted of absences that had been incurred in the month prior to the

exam. The authors found that fourth-grade students missing three or more days in the

month prior to the NAEP exam scored 11 points lower in reading and 13 points lower in

math than students who missed no days in the prior month. While the self-reported nature of the data and the limited timespan for which absence data was collected weakens the study, the findings still show a significant relationship between absences and academic achievement and suggest that the findings from similar peer-reviewed studies may be applicable at the national level.

In addition to the research exploring the effect of early absence on academic achievement, there have been two peer-reviewed studies relating early absences to another important educational outcome—school dropout or failure to graduate on time. Schoeneberger (2012), for instance, analyzed administrative data from a large, urban school district in the southeastern U.S. in 2007 to 2008. While the purpose of the study was not explicitly to relate early absences to educational outcomes but rather to determine if specific profiles of early attendance related to high school dropout, the findings still have important implications about the relationship of absence to student outcomes. Schoeneberger looked at the absence data of about 15,000 children across elementary school. Using group-based trajectory modeling, a form of structural equation modeling combined with random coefficient modeling, the author discovered that children could be categorized into several groups based on their attendance patterns. While the bulk of the students (roughly 82%) fell into the category of "consistent attenders", another group, called "chronic truants" (roughly 4%) was identified. This group exhibited the highest rates of absence across all grades. About 21% of the "chronic truant" group ultimately dropped out of high school compared to only 4% in the "consistent attender" group. While this study did not attempt to find a predictive

relationship between early absence and dropout, the results do suggest that consistently high levels of early absence seem to be associated with increased likelihood of dropout in high school.

In a more direct exploration of early absences as they relate to dropout, Ou and Reynolds (2008) utilized data from the Chicago Longitudinal Study to follow 1,286 low-income, minority students from birth to age 20 (from 1985 to 2005). The study contained a variety of variables across this age range, from demographic information about the family (e.g., mother's education level, parental involvement in school, number of children in the family, reports of abuse or neglect, etc.) to child variables (e.g., low birth weight, preschool enrollment, early literacy skills, student expectations about education, etc.). The data also included information about the average number of absences each student experienced per year before age 12. Using this rich set of almost 50 predictors, the multiple regression model was able to predict 73% of students who ultimately graduated from high school versus students who ultimately dropped out. The strongest, statistically significant predictors within this model were maternal educational attainment, school mobility, educational expectations at youth, and early school absences. In fact, each additional absence day per year was associated with a 7% decrease in the likelihood of high school graduation. While this study suggests that there is a relationship between early absences and dropout even after accounting for a host of other variables, much of the data relied on student, teacher, and parent report. Specifically, the absence data was based on teacher and parent report rather than on administrative school records, which could potentially undermine the validity of the findings. The findings are,

however, consistent with Schoeneberger's study in that students experiencing more absences are more likely to dropout and less likely to graduate on time.

Overall, the empirical literature confirms that there is a relationship between early student absences and two important educational outcomes—academic achievement and school dropout. The research is consistent with the theoretical model of how absences affect achievement and, additionally, comports with the literature findings around early attendance as a protective factor. Furthermore, there is some evidence that the relationship between absences and academic achievement is causal, meaning that absences have a direct, negative effect on student achievement, just as there is some evidence that attendance has a direct, positive effect on student achievement. In sum, the literature around early student attendance and absences supports the notion that the attendance gap may underlie the achievement gap and, moreover, that the attendance gap must be addressed in order to ameliorate the achievement gap.

The Mutability of Attendance: Mitigating Absences as a Risk Factor and Promoting Student Attendance

Because student attendance and absences can affect educational outcomes and because absences may be a cause of the achievement gap, it is necessary to determine whether student attendance is an intervenable behavior. Despite the seemingly negative effects of early absences on educational outcomes, there is empirical evidence that attendance is malleable and that habitual student absence can be improved. Research suggests that when attendance issues are monitored and intervened upon, especially at an early age, negative outcomes can be mitigated and educational trajectories changed.

Research suggests that attendance problems "become more difficult as students age, suggesting that the earlier intervention occurs, the more likely it is to succeed" (Blazer, 2011, p. 1). Thus, absence patterns are mutable, especially at the start of children's educational trajectories.

Furthermore, there are already existing interventions that have showed an impact in reducing absences, especially among low-income, minority students. There are a number of programs and strategies that have been reviewed in the empirical literature and were found to have a positive effect on attendance patterns (Balfanz & Byrnes, 2013; Chang & Romero, 2008; Faria et al., 2017; Jordan, Fothergill, & Rosende, 2018; Kearney & Graczyk, 2014; Lehr, Sinclair, & Christenson, 2004; Sheldon, 2007, Sutphen, Ford, & Flaherty, 2010; Teasley, 2004). For instance, a large-scale effort to reduce absences in New York City public schools by providing students with a variety of supports showed that students with severe attendance challenges gained nearly two additional weeks of school per year when given these supports; in addition, students living in emergency shelters, who are particularly vulnerable to school absences, were about 30% less likely to have significant absence challenges than students not receiving support services (Balfanz & Byrnes, 2013). Another study randomly assigned schools in the Midwest to utilize a system of indicators, including early absence information, to identify students at risk of developing educational issues. Simply by monitoring attendance and with limited implementation of targeted support services, schools in the intervention group were able to reduce the number of students experiencing significant attendance challenges compared to control schools (Faria et al., 2017).

Among the myriad interventions, programs, and strategies identified in the literature two common themes emerge in terms of how to effectively target absence issues: (1) absence data must, first and foremost, be tracked and monitored as early as possible to determine which children are having attendance problems and more precisely target supports; (2) it is essential to engage families and promote home-school communication in the process of mitigating absences, especially in the case of elementary school children experiencing frequent absences (Balfanz, 2016; Chang & Romero, 2008; Faria et al., 2017; Jordan et al., 2018; Kearney & Graczyk, 2014; Lehr et al., 2004; Sheldon, 2007; Sutphen et al., 2010; Teasley, 2004). School-based efforts to reduce absences and promote student attendance must utilize student-level attendance data to identify at-risk students and should work with children's families to identify barriers to attendance and connect families to resources that eliminate these barriers. Because problematic student absences are intervenable, reducing student absences should be considered a major focus of U.S. educational policy, especially as it relates to and may underlie the achievement gap and negatively impact important educational outcomes.

As student attendance, especially in the early grades, is mutable and may have significant impacts on important educational outcomes such as academic achievement, it is necessary to review national educational policy responses that attempt to address this issue. The next section will explore historical and current educational policy efforts to mitigate absence issues and promote student attendance. This section will review the history of attendance policies in the U.S., the legal standards established for student

attendance, and current national efforts and policies around attendance. Critiques of current policies will then be discussed.

The Educational Policy Response to Student Attendance and Absences

The History of Attendance Policy in the U.S.: Compulsory Student Attendance

School attendance is one of the most fundamental aspects of the U.S. educational system. Part and parcel of the founding democratic vision for America was an educated citizenry able to make informed decisions about self-government (Tyack, 1974). For many years, however, this vision did not uniformly apply to all people living in the U.S. While the majority of children attended at least some school during the colonial period, long-term schooling was traditionally reserved for the children of wealthy, white Americans (Howard, 2010; Tyack, 1974). Many occupations during this time did not require formal schooling and often children were needed, for at least part of the school year, to help the family with duties at home or as sources of additional labor (Tyack, 1974).

With the advent of the industrial era, the growth of cities, and rising numbers of immigrants flooding into the country, the need for a formal educational system became more apparent (Howard, 2010). The impetus to establish mandatory schooling laws came from three driving forces: (1) many people were concerned that the rising number of immigrants would erode American culture and that immigrants needed some form of institutional assimilation into American language and cultural norms; (2) progressive reformers, who had a new understanding of childhood as a sacred developmental period and were concerned about the number of children living in urban poverty and entering the

industrial workforce; and (3) workforce and labor unions that worried about competition from child laborers and thus sought to curtail their ability to work through mandatory schooling legislation (Kotkin & Aikman, 1980; Tyack, 1974).

With the exertion of these cultural and political forces, adoption of attendance laws grew steadily over the 19th and 20th centuries but varied by locality. Falling under the 10th amendment, compulsory attendance laws were left to the discretion of states, rather than the federal government (Howard, 2010). By 1885, 16 out of 38 states had compulsory attendance laws, and by 1900, 31 states had compulsory attendance laws for children ages eight to fourteen (Tyack, 1974). Of those states without compulsory attendance laws, many still had a large percentage of children attending school, as attitudes around the benefits of schooling continued to grow in popularity. However, the impetus for mandatory schooling laws continued even as more children entered the school system. Children who were not attending school, especially children of immigrants living in urban poverty, were regarded as social misfits, delinquents, and potential criminals. Tyack (1974) notes that at this time, many Americans believed that the children of poor immigrants "should be compelled to attend school, for it was precisely such children that needed training the most….In the arguments of many advocates of compulsory attendance...schooling became a form of preventive detention— and often the intermediate step on the way to more total institutionalization" (pp. 68- 69). Thus, school attendance laws were viewed as a way to curb delinquency and inculcate American values in young people, particularly from immigrant families and families living in poverty.

By 1918, all states had adopted compulsory attendance laws in conjunction with the advent of child labor laws (Kotkin & Aikman, 1980). The pervasiveness of these laws belies their effectiveness, however, as there was little infrastructure to actually enforce them in any way (Howard, 2010; Landes & Solmon, 1972). These laws were commonly ignored in the early 1900s with few legal consequences (Tyack, 1974). While the creation of positions like school attendance officers were aimed at promoting enforcement of compulsory attendance laws, these positions were woefully understaffed. For instance, by 1925 there was only one school attendance officer for every 7,500 children in schools nationally (Tyack & Berkowitz, 1977). As such, there were still few prosecutions related to non-attendance in the early 1900s. Compulsory attendance laws were, at this point, more symbolic than practical or realistic. As the legal mechanisms for enforcing compulsory attendance laws became more robust and the enforcement infrastructure grew through the early- and mid-1900s, so too did legal standards and policy strategies for reducing absence and promoting student attendance.

Current Legal Standards for School Attendance

Like the majority of educational policies, attendance laws continue to be controlled at the state and local level. It is thus difficult to provide a national picture of attendance laws as they vary by locality. As Kotkin and Aikman (1980) note: "While the basic structure of the compulsory attendance and child labor provisions is similar from state to state, the details of those provisions are substantially different so that nationally, the compulsory attendance and child labor laws present a dense network of laws which are not easily susceptible to classification" (p. 5). For instance, state laws that mandate

the age range at which compulsory school attendance is required vary considerably.

Some states require attendance beginning at five, while others begin at age seven; some

states mandate attendance until the age of eighteen, while others only require attendance

until age sixteen (Howard, 2010, pp. 372-374). Even within a given state, there are often

differences in attendance policies and practices from district to district and even from

school to school. As Sutphen et al. (2010) state, "Attendance policies and procedures

followed by individual schools are most often set locally, either be state departments of

education or school districts even if there are state statutory definitions…exemplifying

the 'localized' nature of the problem" (p. 161). Thus, there are no federal laws stipulating

student attendance; regional differences preclude specific statements about attendance

policy nationally, but, in general, students in all states are required to attend some form of

schooling from middle childhood into adolescence.

The Use of Policy Indicators Related to Attendance

As with legal standards around attendance, policy indicators around attendance

and absence are quite variable by region and are difficult to discuss at the national level.

Two common indicators do emerge, however, when surveying the national educational

policy landscape. Most states and districts require public schools to track two key

indicators related to attendance: average daily attendance and truancy. Additionally, these

indicators have been brought to national prominence in response to federal reporting

mandates under the No Child Left Behind Act (No Child Left Behind [NCLB], 2002),

which required states to report average daily attendance for elementary and middle

schools and truancy rates for high schools (Bauer, Jordan, Chang, & Balfanz, 2018**)**.

Average daily attendance (ADA) is the average number of students present in a school on any given day. There are multiple calculations and different measurement decisions tied to this indicator (e.g., when is a student considered "present?" Are tardy students counted as present?), but the indicator is typically expressed as a percentage of students in school on an average school day (Christie, 2005). Thus, schools who report a 90% daily attendance rate are indicating that on average, 90% of their students are present on any given school day. While there is no legal standard for ADA (e.g., there are no legal statutes dictating that all schools in a given state must maintain an ADA of 95%), it serves an important legal function in terms of its connection to school funding formulas. As Nauer (2016) notes, "Average daily attendance is the measure used nationwide to evaluate attendance for school funding and accountability" (p. 33). Because many states allocate money to districts based on enrollment and attendance rates, ADA has become an important educational indicator to determine distribution of funds. Guare and Cooper (2003) argue that "failure to attend may mean loss of revenues (as state aid per pupil is allocated based on 'average daily attendance' or ADA in many systems)" (p. 9). As a policy indicator, ADA is an aggregate measure of attendance at the school, district, or state level. It does not provide information about individual student attendance and is not related to any systematic follow-up actions or interventions (i.e., if a school has below 90% ADA, all students and families receive an attendance flyer reminding them how important school attendance is). While individual schools or districts may institute certain policies or practices related to ADA, it does not have

repercussions at the individual level and is therefore not directly tied to efforts to curb individual student absences.

The other most common legal standard around attendance is truancy. While ADA provides aggregate information related to funding allocation and does not have implications for individual students, truancy is an indicator of absences at the individual level and is generally tied to actions related to a single student or family. As Sutphen et al. (2010), "Truancy is a legal term that is generally defined by each state as a specified number of unexcused absences from school over a designated period of time" (p. 161). What constitutes an unexcused absence and the number of unexcused absences that must be reached before a child is deemed truant is highly variable by state and even by district (London et al., 2016). The authors go on to note that "there is no uniform national definition of truancy and, therefore, no estimate of the national prevalence of the problem" (Sutphen et al., 2010, p. 161).

While it is difficult to provide an overarching definition of truancy given these regional differences, it is crucial to have some basic understanding of the term "unexcused absence", as the concept of truancy is built around this designation. An "unexcused" absence typically refers to an absence that is not recognized as legitimate by the school. With some form of parental acknowledgement (e.g., a phone call or note to the school), most absences are considered excused by the school (e.g., absence for illness, absence for family reasons; Teasley, 2004). There are also absences that can be sanctioned by the school (e.g., a field trip or athletic game), which would also be

considered excused. Thus, an "unexcused absence" typically refers to an absence that occurs without parental and/or school sanction.

Though students of any age may accrue unexcused absences, the idea of adolescents "cutting class" or "skipping school" is heavily associated with truancy, as truancy is generally understood to refer to surreptitious absences that are happening without the knowledge of the family or school (Kearney, 2008b). Despite the public perception of truancy as an adolescent issue, however, students of all ages may be deemed truant, as truancy centers around the idea of a minor who is unaccounted for by adults (i.e., a child who is not under the supervision of the family or the school). Truancy, therefore, relates to a broader societal concern—the need for the adult supervision of children. The idea that children need constant supervision and protection to ensure their safety and well-being is well established as both a legal and ethical principle (Brazelton & Greenspan, 2009; Maughan & Moore, 2010). Truancy thus provides a mechanism for adults from both home and school to be held accountable for children's care and encourages adults from both home and school communicate about children's whereabouts (e.g., the school calls the parent to inform them that the child did not show up for class or the parent calls the school and to inform them that the child is ill and will be staying home).

When children are not being supervised by adults at home or school (i.e., when the student begins to accrue unexcused absences), a designation of truancy is tied to multi-system ramifications to reflect the seriousness of this need for adult supervision of youth. A designation of truancy triggers involvement from multiple public service

sectors—education, law enforcement, child protective service systems, and the justice system. Once a child begins to accumulate unexcused absences, schools typically activate a cascading series of steps in order to prevent the involvement of other public service systems and their more serious consequences. For instance, one unexcused absence may trigger an automated call to the home, three unexcused absences may trigger a formal letter to the home, and five unexcused absences may trigger a conference among family members, school personnel, and the student to address the attendance problem (Sutphen et al., 2010). If the student continues to accumulate unexcused absences despite these interventions and meets the criteria for truancy, other public service systems are triggered to take action. For instance, law enforcement or truancy officers may be called to the student's home to compel him/her to attend school or child protective services may be called to assign a caseworker to the family of a young student to help them troubleshoot issues preventing the student from getting to school. The most severe consequence of truancy is imposed by the justice system in the form of legal action taken against the student or the student's family, depending on the age of the child. Legal sanctions typically involve: monetary fines, orders to accompany the minor to school, counseling, probation, parenting classes, etc. In extreme cases, adolescents who are habitually truant may be sentenced to jail time and families of younger students may be charged with child neglect or abuse and face accompanying jail sentences (Smink & Heilbrunn, 2006). Truancy thus involves multiple public service systems beyond education and is tied directly to legal repercussions for students and their families.

Because truancy definitions are so variable among states, the legal ramifications, school policies, and interventions to prevent truancy are similarly difficult to look at broadly and defy unidimensional classification. In general, though, truancy can be thought of as an individual indicator of attendance status tied to a series of cascading responses from multiple public service systems aimed at curbing absence issues through punitive measures.

Critiques of Current Indicators: Problems with Average Daily Attendance and Truancy

Despite their widespread use as indicators of attendance, there are several issues with using average daily attendance and truancy to monitor and improve student attendance behaviors. ADA, as described above, is an aggregate indicator of a school's overall attendance rate. While it can signal school-wide attendance challenges and thus trigger school-wide responses (e.g., a school with a low ADA might put up posters around campus encouraging better attendance), it does not allow a school to detect which specific students are having attendance issues and provide a targeted response to those at risk (Chang & Romero, 2008). ADA is thus more appropriately used as a policy tool in combination with enrollment numbers to allocate resources and is not as useful in terms of providing actionable information about individual students. Moreover, ADA might actually be misleading in its portrayal of overall school attendance. As Balfanz (2016) notes, "commonly used attendance measures can mask attendance challenges. A school could have an average daily attendance rate of 92% and still have 20% of its students missing a month or more of school" (p. 9). Thus, while its ADA rate may seem high, a school might still have pockets of students experiencing incredible attendance challenges.

ADA does nothing to identify those students at risk and may ultimately paint a more optimistic picture of school-wide attendance than is warranted.

On the other hand, truancy provides information at the individual level and can thus be used to identify and target supports for at-risk students. The major critiques of truancy, in contrast, stem from its predominant focus on older students, which limits its utility in identifying younger children in need of supports around attendance, and the fact that it does not account for all missed time in school but focuses exclusively on unexcused absences. While students of any age can accumulate enough unexcused absences to be deemed truant by their schools, there is a general conception of truancy as an adolescent syndrome of delinquency, one that does not apply to younger students (Chang & Romero, 2008). Much of the literature and rhetoric around truancy is focused on adolescent misbehavior related to drugs, sexual activity, and crime (Maynard, Salas-Wright, Vaughn, & Peters, 2012; Mogulescu & Segal, 2002; Mueller, Giacomazzi, & Stoddard, 2006; Sheldon & Epstein, 2004; Smink & Heilbrunn, 2006). Gaure and Cooper (2003) conclude that "truancy (unlike absenteeism) involves an unjustified absence in which students themselves are the cause. Hence, technically students are not truant if their parents or guardians keep them home for various reasons" (p. 8). Chang and Romero (2008) reiterate this rationale, suggesting that young students are unlikely to be absent without the knowledge of their primary caregiver, making truancy an inappropriate designation for young children. Thus, an elementary school student would rarely be deemed a "truant" regardless of the number of unexcused absences he or she has accumulated; this limits the utility of a policy indicator such as truancy for young

children (Nauer, 2016; Ready, 2010). Moreover, some suggest that unexcused absences are a relatively infrequent occurrence for younger students, as they connote that the parent or guardian is unaware of the student absence and most young students do not willfully skip school. If most absences experienced by younger students are excused, this limits the utility of an indicator like truancy for elementary school children. Balfanz (2016) contends that "In preK, kindergarten, and the elementary grades, reporting only on truancy rates greatly underestimates [total absences] because, at these grade levels, most absences are excused" (p. 9). Although Balfanz does not offer empirical support to substantiate his claim, there is a general belief that unexcused absences are less common among younger children. Thus, truancy would not apply to many young children as, for the most part, their absences from school would be condoned by primary caregivers and therefore would be excused.

Younger students face significant attendance challenges, however, and the need to identify these children remains pressing regardless of whether they fit into conventional understandings of truancy. As the literature surrounding school attendance and absence suggests, attendance issues often appear at school entry (Balfanz, 2016; Balfanz et al., 2007; Chang & Davis, 2015; Chang & Romero, 2008; Faria et al., 2017; Hickman & Heinrich, 2011; Lehr et al., 2004; Sheldon, 2007). Furthermore, it is vital that attendance issues are detected and intervened upon early to increase the likelihood of disrupting negative developmental pathways and promoting successful outcomes (Bauer et al., 2018; Ehrlich, 2014; Ginsburg et al., 2014; Mac Iver, 2010). As Blazer (2011) suggests, "the start of elementary school is the critical time to shape attendance patterns" (p. 1).

In addition, truancy's exclusive focus on unexcused absences rather than *all* absences depends on the rationale that schools need a way to identify unsupervised students (i.e., students unaccounted for by the school or the family). While there is certainly a logic and practical need behind identifying students that have not been accounted for by the school or family, there is a competing logic suggesting that any time away from school is detrimental to the child. As the conceptual framework invoked previously indicates, absences from school remove the student from a crucial microsystem that fosters development; any absence from school may harm a child's outcomes, regardless of whether or not it was sanctioned by the school and the family. While both of these competing assertions seem feasible, critics of truancy note that the logic behind the indicator—that unexcused absences are more important to monitor and act upon than excused absences—has not been empirically substantiated. Rather, truancy was adopted to fulfill the important societal need to ensure that adults were accountable for the supervision of children; truancy did not, therefore, undergo any rigorous empirical testing to ensure that unexcused absences were the only absences that mattered for all student outcomes. Thus, truancy lacks evidence-based validation of its exclusive focus on unexcused absences and may obfuscate attendance issues by not taking excused absences into account.

ADA and truancy are, therefore, insufficient indicators of attendance problems. ADA is used only at the aggregate level and is not helpful in identifying students at the individual level, while truancy is limited in its utility to detect young children in need of attendance support services and focuses solely on unexcused absences without empirical

substantiation. Because of these shortcomings, researchers and policymakers have recently called for the adoption of an additional policy indicator related to student attendance.

A New Policy Indicator for Attendance: Chronic Absenteeism

Critiques of both ADA and truancy have catalyzed the creation of a new policy indicator for attendance. This new policy indicator—called chronic absenteeism—emerged in the last decade and has become increasingly popular in the last five years. The federal Every Student Succeeds Act (ESSA) of 2015 requires all states to include "non-academic" accountability measures as part of how they evaluate schools. Almost three-fourths of states have included chronic absenteeism as an accountability indicator in their ESSA implementation plans (Jordan et al., 2018). Bauer et al. (2018) report that ESSA "requires states to hold schools accountable for at least one measure of 'school quality or student success [SQSS]'….36 states, the District of Columbia, and Puerto Rico have chosen chronic absenteeism as either one of or their only SQSS indicator(s)" (p. 5). While there is no consistent definition of chronic absenteeism, most states define the indicator as missing 10% of the school year for any reason, whether the absence is excused or unexcused (Balfanz, 2016; Jordan et al., 2018). Bauer et al. (2018) go on to note that, "There is no consistent definition of chronic absenteeism, either in the academic literature or across states…The majority of states define a chronically absent student as one who misses at least 10% of the school year" (p. 10). Though 10% is the most commonly used threshold at which chronic absenteeism is defined, other states define chronic absenteeism using a 5% threshold, while others utilize a threshold for

absence counts rather than a percentage of days missed. The adoption of the 10%
threshold by the majority of states reflects recommendations by advocates of chronic
absenteeism but, as yet, little empirical research has been conducted to substantiate the
predictive utility of this threshold in identifying at-risk students.

While there is no precise definition of chronic absenteeism and states have
operationalized it in various ways, chronic absenteeism as a policy indicator differs
conceptually from ADA and truancy in four important ways: (1) ADA is an aggregate
measure of student attendance and provides no actionable information at the student
level, while chronic absenteeism is an indicator of attendance for each individual student
and can lead to the identification of students at risk for educational issues; (2) truancy is
primarily associated with adolescents and older students, whereas chronic absenteeism
can be used to identify students with absence issues at any grade level and is particularly
useful for identifying young students for whom truancy is less appropriate; (3) truancy is
tied to a series of legal responses, most of which are punitive (e.g., fines), whereas
chronic absenteeism is used strictly for identification of students at risk for educational
problems (i.e., is used as an "early warning indicator" of later potential problems) and is
not associated with any sort of legal mechanism or punitive response; and (4) truancy is
only defined by unexcused absences or absences that are not condoned by both the school
and the primary caregiver, while chronic absenteeism takes *all* absences from school into
account whether they are unexcused or excused (London et al., 2016).

Because chronic absenteeism is not meant to supplant ADA or truancy as a policy
indicator but serve as a complement to them, understanding the differences between them

is crucial. The difference between chronic absenteeism and ADA is similar to the difference between truancy and ADA; ADA is only useful as an attendance indicator at aggregate levels for determining things like overall attendance in a school or district and is necessary for school funding formulas, whereas chronic absenteeism—and truancy, for that matter—attempt to identify individual students who are experiencing a problematic number of absences. Thus, ADA and chronic absenteeism are used for very different purposes. The difference between chronic absenteeism and truancy, however, is more complicated.

The first two differences between truancy and chronic absenteeism described above relate to how the indicators are *used*. Unlike truancy which is used to monitor and remediate older students for whom attendance—and likely other educational issues—has already become a problem, *chronic absenteeism is meant to be used preventatively* to identify students at risk for educational problems *as early as possible* in order to target supports and interventions to them and potentially disrupt maladaptive educational trajectories. As such, chronic absenteeism has been designated as a so-called "early warning indicator". An early warning indicator is designed to detect students who may be "off track" educationally and are at risk for later educational problems (e.g., failing academic grades, high school dropout, etc.) through easily obtained administrative data (e.g., attendance records, academic grades, etc.; Balfanz et al., 2007; Ginsburg et al., 2014). For example, a school might use a combination of absence information, academic grades, and suspension records in elementary school to determine which students are most likely at-risk for high school dropout; resources and efforts to support these students

could then be administered and might change their educational trajectories and actually prevent them from dropping out of high school. This distinction between truancy and chronic absenteeism—namely, that truancy uses consequential control to try to force older students to attend school while chronic absenteeism uses antecedent control to try and find struggling students early on and prevent them from experiencing more severe problems in the future—is crucial to understanding the utility of chronic absenteeism as a complementary policy indicator to truancy. Thus, these first two differences between truancy and chronic absenteeism relate to its use: truancy is generally applied to older students and tied to legal ramifications, while chronic absenteeism can be applied to any age group and is not linked to punitive action but rather serves as an early warning indicator that the student, especially a young student, may need additional supports to prevent future negative outcomes.

The third, and arguably most crucial, difference between truancy and chronic absenteeism relates to a conceptual divergence in how absences are *understood*. Because truancy has primarily served as a designation of whether the student is supervised (i.e., of whether the absences were sanctioned by both the home and the school), the only absences of relevance are unexcused absences. If absences are excused, it means that some communication between home and school has taken place and the child has been accounted for; these absences are thus irrelevant to truancy. Chronic absenteeism, on the other hand, assumes that all attendance is crucial for the well-being of the child and thus any absences are a potential risk factor to the child's education; it thus *equates absence types* rather than distinguishing between them. In the case of chronic absenteeism, the

reason behind the absence—and thus the absence type—becomes irrelevant. For instance, chronic absenteeism does not distinguish between an absence due to illness in which the parent called to notify the school (i.e., an excused absence) and an absence due to illness in which the parent *did not* call to notify the school or an absence in which the student skipped school without the parent's knowledge (i.e., unexcused absences). Chronic absenteeism, thus, does not make a relevant distinction between excused and unexcused absences and gives each an equal "weight" as it represents valuable instructional time in school that has been missed (Jordan et al., 2018).

In sum, chronic absenteeism is distinct from ADA in that it is used to identify individual students rather than describe schools or school districts in the aggregate. Chronic absenteeism is distinct from truancy in both its use—chronic absenteeism is not reserved solely for older children and is particularly useful as an early warning indicator for younger children who may be at risk for later educational problems rather than as a punitive tool tied to legal repercussions—and the logic upon which it is based—that all absences are equally detrimental to educational outcomes and should be given equal weight when identifying at-risk students.

As a relatively new and widely used policy indicator, it is important to establish an evidence base for the use of chronic absenteeism. The next section will review the literature relating chronic absenteeism to educational outcomes, particularly among younger students to determine its efficacy as a policy indicator for identifying students at risk for educational problems.

Chronic Absenteeism: What Does the Research Say?

Because chronic absenteeism has only emerged as a policy indicator within the last decade, empirical, peer-reviewed research substantiating it as a useful indicator of educational problems is somewhat limited. There are only three peer-reviewed, published studies exploring the predictive association between elementary school chronic absenteeism and important educational outcomes, such as academic achievement. Gottfried (2014) released the first published study using data from the 2010-2011 Early Childhood Longitudinal Study-Kindergarten (ECLS-K) from 2010 to 2011. The ECLS-K is a publicly available dataset following a nationally representative group of children beginning in kindergarten. The dataset includes various indicators of educational well-being—from math and reading scores to social skills. The study sample included over 10,000 kindergarten children and looked at the effects of chronic absenteeism on educational outcomes at the end of the kindergarten school year. Chronic absenteeism was determined by teacher-reported absences and only included ranges of absences (e.g., 1 to 4 absences, 5 to 7 absences, etc.), so the most common definition of chronic absenteeism (i.e., 10% of the school year or 18 days) was not used in this study. Rather, the study defined two categories of chronic absenteeism: moderate chronic absenteeism (11 to 19 absences) and strong chronic absenteeism (20 or more absences). The study utilized a classroom fixed-effects regression model and controlled for a variety of demographic characteristics (e.g., gender, race/ethnicity, etc.) and a variety of household characteristics (e.g., number of siblings, age of mother at first birth of child, etc.). Gottfried found that there were statistically significant negative effects of chronic

absenteeism on reading and math outcomes (e.g., scores on a two-stage, adaptive scale

measuring skills like print familiarity, letter recognition, number and pattern recognition,

spatial sense, etc.). Students experiencing moderate chronic absenteeism scored 0.04

standard deviations below their non-chronically absent counterparts in reading and 0.06

standard deviations below in math; the effects were more pronounced for students

experiencing strong chronic absenteeism, who scored 0.17 standard deviations below

their peers in reading and 0.20 standard deviations below in math. Additionally, Gottfried

found that chronic absenteeism had an effect on social skills related to students'

eagerness to learn (i.e., the child shows eagerness to learn new things) and internalizing

behavior problems (e.g., child appears anxious, lonely, exhibits low self-esteem, etc.).

While eagerness to learn represents only one item within the approaches to learning scale

utilized in the ECLS-K and therefore should not be considered a robust scale measuring

student engagement or enthusiasm for learning, the study showed a negative association

between eagerness to learn and chronic absence. Moderate chronic absentees scored 0.08

standard deviations below their non-chronically absent peers on the eagerness to learn

item and strong chronic absentees scores 0.23 standard deviations below. On the scale

measuring internalizing behavior problems, moderate chronic absentees scored .09

standard deviations above their non-chronically absent counterparts, and strong chronic

absentees scored 0.17 standard deviations above, meaning that these students experienced

more behaviors associated with anxiety, sadness, social isolation, etc. While the use of

teacher-report absence ranges rather than administrative records of absence counts

presents a limitation to this study, Gottfried provided evidence using nationally

representative data that chronic absenteeism has a negative effect on academic outcomes and social skills in kindergarten.

Subsequently, London et al. (2016) conducted a study using longitudinal, administrative data from two school districts in the San Francisco area to explore the relationship between chronic absenteeism and achievement on state standardized tests for about 6,000 students from 2003 to 2004. The authors utilized individual longitudinal growth models to estimate the effect of being chronically absent on reading and math scores beginning in elementary school, controlling for a variety of demographic indicators (e.g., receipt of free and reduced-price lunch, race/ethnicity, etc.). The study used the most common definition of chronic absenteeism—children who were absent 10% or more of the school year (or 18 or more days)—and all attendance data was retrieved through district records, rather than from teacher report. The study found that students who were chronically absent for one year in elementary school scored 0.18 standard deviations below their non-chronically absent peers in reading and 0.17 standard deviations below their non-chronically absent peers in math. For students that were chronically absent for multiple years in elementary school, these differences grew to 0.22 standard deviations in reading and 0.32 standard deviations in math. Additionally, the study found that kindergarten students had the highest rates of chronic absence of any students (from first grade to twelfth grade), and that the best predictor of current-year chronic absence is chronic absenteeism in previous years. Both of these findings are consistent with Gottfried's findings from a nationally representative dataset.

Finally, Gottfried (2015) probed the association between chronic absenteeism and educational outcomes even further by exploring whether chronic absenteeism had both a negative effect on the achievement of students who were chronically absent themselves and also on the achievement of non-chronically absent peers in classrooms with a high prevalence of chronically absent students. Using administrative records from over 23,000 third- and fourth-grade students in the School District of Philadelphia from 1994 to 2001, this study used both a baseline model and a school fixed-effects model to determine the influence of chronic absenteeism on student performance in reading and math on the SAT-9 (Stanford Achievement Test-Ninth Edition), while controlling for demographic variables. Like the previous study, chronic absenteeism was defined as missing 10% or more of the school year (18 or more days of school), and absence counts were obtained from district administrative files. This study found that there was a significant, negative effect of chronic absenteeism on achievement, with chronically absent students scoring 0.08 standard deviations below their non-chronically absent counterparts in reading and 0.10 standard deviations below in math. Furthermore, controlling for students' own chronic absenteeism as well as other demographic characteristics, Gottfried found that the percentage of chronically absent classmates also had a significant, negative effect on student reading and math achievement. The effect of percentage of chronically absent classmates was -0.04 standard deviations in reading and -0.05 standard deviations in math. These findings were consistent across the baseline and school fixed-effects models. While the seeming "peer effect" of chronic absenteeism was half the size of the individual effect, these findings still point to an alarming conclusion: that students in

classrooms with high numbers of chronically absent students may still face negative effects on their academic achievement even if they consistently attend school. The need for teacher remediation of chronically absent students and the disruptive learning environment created by frequently absent students may actually have spillover effects on peers in the classroom. Thus, across all three peer-reviewed studies, it appears that chronic absence is negatively associated with achievement in reading and math, is associated with absence problems in future years, may be linked to reduced educational engagement and more internalizing behavior issues, and may have an effect on the academic achievement of non-chronically absent peers.

In addition to this peer-reviewed research, there have been a number of non-peer-reviewed studies in recent years as chronic absenteeism has been increasingly adopted by states and school districts as a relevant policy indicator. Five studies have linked chronic absenteeism in younger students to negative achievement outcomes (Applied Survey Research, 2011; Chang & Davis, 2015; Chang & Romero, 2008; Ehrlich et al., 2014; Spradlin, Cierniak, Shi, & Chen, 2012; Utah Education Policy Center, 2010). Applied Survey Research (2011), Chang & Davis (2015), Chang & Romero (2008), Spradlin et al. (2012), and Utah Education Policy Center (2010), using both national and regional datasets, found that elementary school students experiencing chronic absenteeism (defined in each study as missing 10% or more of the school year) had lower achievement scores than their non-chronically absent peers and that early chronically absenteeism was predictive of chronic absenteeism in later grades. Moreover, Ehrlich et al. (2014) found that students missing 15 days of school or more in preschool (i.e., prior

to elementary school entry) had lower scores on a measure of school readiness in kindergarten, lower reading achievement scores in second grade, and increased likelihood of being absent in future years. These findings are consistent with the peer-reviewed literature showing that students experiencing chronic absenteeism in early grades tend to have lower academic achievement levels and continuing absence problems in later grades. Additionally, there have been several non-peer-reviewed studies linking chronic absenteeism in middle school to later attendance problems and ultimately to high school dropout (Allensworth, Gwynne, Moore, & de la Torre, 2014; Balfanz & Byrnes, 2012; Mac Iver, 2010; Utah Education Policy Center, 2010). While there are no studies linking chronic absenteeism in elementary school to high school dropout, it is noteworthy that chronic absenteeism in elementary school is indicative of continuing attendance problems and, for middle school students, can even be linked to high school dropout (Balfanz et al., 2007; Blankenship, 2015). Thus, the peer-reviewed research linking early chronic absenteeism to educational outcomes provides consistent evidence that this policy indicator is associated with lower academic achievement and future attendance problems for elementary school students.

Despite the need for a policy indicator that will identify young students at risk and the research evidence establishing an associational relationship between chronic absenteeism and negative student outcomes, *there is an important critique of chronic absenteeism that researchers and policymakers have yet to address*. Chronic absenteeism rests upon two key assumptions: (1) *that all absences are equally detrimental to student outcomes* (i.e. that there is no qualitative distinction between excused and unexcused

absences in terms of how they affect outcomes) and (2) *that the 10% threshold of total absences days represents a meaningful cut-off point to signify student risk status*. Unfortunately, *neither of these major assumptions has been empirically tested with appropriate methodologies*. The research presented above provides evidence of an association between chronic absenteeism and negative educational outcomes but does not provide any scientific substantiation of the major assumptions upon which this indicator is based. Because the first assumption relates to the *quality* of the information being used to comprise this indicator, it is necessary to investigate whether the distinction between absence types is differentially related to negative educational outcomes before investigating the second assumption, which has to do with the *quantity* of absences that are being used to operationalize chronic absenteeism. The next section will explore the validity of the first premise upon which chronic absenteeism is based by reviewing the limited empirical evidence regarding the differential effects of absence types on student outcomes.

The Crux of Chronic Absenteeism: Are Excused and Unexcused Absences Equally Detrimental to Student Achievement?

Chronic absenteeism is based on the assumption that instructional time is paramount to educational success and thus absence type (i.e., excused or unexcused absence), and all that is connoted therein, is a meaningless distinction. This theory suggests that there is no meaningful qualitative difference between these two absence types as they relate to important student outcomes, and there is, thus, no need to make relevant distinctions between them. As Balfanz and Byrnes (2012) state:

> Chronic absenteeism is typically based on total days of school missed, including both excused and unexcused absences. This is critical because the evidence indicates that it is how many days a student misses that matters, not why they miss them. In order words, the detrimental impacts of missing school occur if a student misses because of illness, suspension, the need to care for a family member, or any other reason. (p. 7)

Under this conceptual frame, all absences are detrimental to student outcomes and should therefore be used to determine which students are at risk for future educational problems (Coelho et al., 2015).

Despite the logic of this theory and the "evidence" referenced by Balfanz and Byrnes above, the existing empirical research testing this assumption is limited (Sutphen et al., 2010). Chen et al. (2016) note that there is a significant "gap in the empirical literature that makes such research challenging: the lack of differentiation between excused and unexcused absences" (p. 1068). Gottfried (2011a) supports this contention, stating that "The difficulty in relying on the current empirical literature is that most of the studies have not differentiated between unexcused absences and total absences. As a consequence, the findings from these studies may potentially contain confounding issues resulting from not parsing out the effects of [different types of] absence" (p. 1599-1600). Much of the existing literature focuses on total absences or on unexcused absences alone (i.e., truancy), and thus, does not explore the differential effects of absence type (i.e., excused or unexcused absence) on student outcomes.

There are, however, two peer-reviewed studies that attempt to parse the effects of excused and unexcused absences and determine whether they have differential effects on student outcomes. Gottfried (2009) was the first to explore this issue empirically using data from 1994-2001 in the Philadelphia School District. The study included 90,000

second, third, and fourth grade students and used district administrative records of

absences differentiated by type (i.e., excused and unexcused) and reading and math

scores on the SAT-9 (Stanford Achievement Test-Ninth Edition). The study utilized a

fixed-effects regression model with classroom-level clustering to account for between-

classroom differences and also tested a value-added model to account for previous

academic achievement scores. Gottfried found that having a higher proportion of excused

absences to total absences was actually associated *with a positive effect* on reading and

math scores, whereas having a higher proportion of unexcused absences to total absences

had a negative effect on reading and math achievement. As Gottfried notes in this study:

> As students trend toward having an increasingly high ratio of excused absences to total absences, they perform, on average, significantly higher than their reference group— students with an increasingly larger fraction of unexcused absences. For instance, students who have 100% of their absences excused perform higher on the SAT 9 reading exam than do students with 100% unexcused absences, holding all else constant, including number of days absent. Yet, students with 100% of their absences unexcused perform, on average, lower on the SAT 9 reading exam, again holding all else constant, including days absent. (p. 405)

These findings were consistent across baseline, fixed-effects, and value-added models.

Even under the most stringent model (i.e., value-added), proportion of excused absences

to total absences had a small, significant, positive effect size on reading achievement of

0.02 and on math achievement of 0.04; the effect size for the proportion of unexcused

absences to total absences, on the other hand, was significant and negative (-0.02 for

reading and -0.04 for math). These findings led Gottfried to conclude that "Although

much of the literature has focused on absences in the aggregate sense, without any

distinction between excused or unexcused, this study shows that it is just as crucial to

develop a relation between trends in types of absences and subsequent school performance" (p. 409). Thus, this study suggests that there are actually differential effects of absence type on student achievement, with a high proportion of unexcused absences negatively related to reading and math, while a high proportion of excused absences was associated with positive effects on achievement.

The second peer-reviewed study exploring the differential effects of excused and unexcused absences on student outcomes utilized both a national and state dataset. Gershenson et al. (2017) conducted a study using both the ECLS-K (Early Childhood Longitudinal Study-Kindergarten) from 1998 to 1999 and administrative educational data from the state of North Carolina from 2005 to 2010. The ECLS-K sample included about 12,000 students in kindergarten, while the North Carolina sample included about 900,000 students from third to fifth grade. The study utilized a classroom-level, fixed-effects regression model for both samples, controlling for demographic variables based on what was available in the dataset (e.g., both studies controlled for free and reduced-price lunch status, but the ECLS-K contained additional information about family characteristics, like maternal education level). Similar to Gottfried (2009), the study also used a value-added model to account for prior student achievement in estimating the effects of absence types on achievement outcomes. Again, accounting for previous achievement allows for better isolation of the effects of absence type. The ECLS-K study utilized a two-stage assessment of reading and math achievement appropriate for kindergarten students (e.g., letter recognition, beginning sounds, number and shape recognition, addition, etc.), while the North Carolina sample utilized state standardized achievement tests to measure

reading and math ability. While both studies used administrative records of absence data,

about 35% of differentiated absence data were missing from both samples, as roughly

one-third of schools in each sample did not report absence type (i.e., excused or

unexcused) and were thus excluded from the study. The results of the analysis for the

ECLS-K sample found no statistically significant differences in the effect of excused and

unexcused absences on achievement. This finding is tempered, however, by the smaller

sample size of this dataset following the exclusion of students for whom differentiated

absence information was not available (i.e., the sample was reduced from about 12,000 to

7,000). The authors recognize that this smaller sample size may have left the study

underpowered and thus unable to detect a statistically significant difference in the effect

of absence type on achievement. On the other hand, the North Carolina data still

contained about 650,000 student records despite the missing absence information. In this

case, the authors did find a differential effect of unexcused and excused absences on

achievement. They note that "unexcused absences are two to three times more harmful

than excused absences, and these differences are strongly statistically significant" (p.

151). Excused absences were associated with a 0.002 standard deviation decrease in

reading achievement and a 0.005 standard deviation decrease in math achievement;

unexcused absences were associated with a 0.006 standard deviation decrease in reading

and a 0.01 standard deviation decrease in math. While these effects are relatively small, it

is still notable that there is a differential effect of absence types on achievement, with

unexcused absences having a stronger negative effect. Though Gottfried (2009) actually

found a positive effect of the proportion of excused absences on achievement, the

findings of this study still comport with his findings suggesting that unexcused absences have a stronger negative effect on academic achievement than excused absences.

While the findings from both of these studies do not provide sufficient scientific evidence to conclude that unexcused absences are unequivocally more detrimental to student outcomes than excused absences, they do necessitate further inquiry. The necessity for further inquiry has been augmented by the mandates of ESSA; as the majority of states have adopted chronic absenteeism as an early warning indicator of school quality and student success to fulfill these mandates, it is imperative for researchers to question whether absence types have differential effects and whether indicators like chronic absenteeism should reflect those differences. Gottfried (2009) upholds this assertion, noting that: "by differentiating patterns of attendance via types of absences, schools could more efficiently identify at-risk students early in schooling based on proportions of unexcused absences" (p. 411). Thus, a policy indicator, such as chronic absenteeism, might be able to better identify students at-risk for educational problems by accounting for the differential effect of excused and unexcused absences on student outcomes.

Despite the associational evidence provided by these studies, which suggests that there is a differential effect of absence type on student achievement, neither study utilizes an appropriate methodology to test whether this differential effect has implications for a policy indicator like chronic absenteeism. Furthermore, neither study assesses whether this differential relationship between absence type and academic achievement is consistent longitudinally across the early elementary grades. *There are, thus, no existing*

empirical studies that test the first underlying assumption upon which chronic absenteeism is based using a methodology that would have direct implications on its efficacy as an early warning indicator. A study that tests this assumption in a way that is directly applicable to chronic absenteeism would assess the utility of multiple continuous variables (such as excused and unexcused absence days) in classifying people into binary diagnostic categories (e.g., students at risk for low academic achievement vs. students not at risk) and would then test whether this utility was stable across time (for example, across early elementary school).

There are a variety of analytic methodologies that could be applied to answer these questions that have been extensively adopted in certain fields (such as medicine, engineering, and psychology) though they remain little used in educational research (Swets, Dawes, & Monahan, 2000; Wilson, Olinghouse, McCoach, Santangelo, and Andrada, 2015; Youngstrom, 2014). The most widely used and well-tested method for this type of research is called the receiver (or relative) operating characteristic (ROC) (Baker, 2003; Engelbrecht et al., 2002; Hallan, & Åsberg, 1997; Jordan, Glutting, Ramineni, & Watkins, 2010; Lee et al., 2008; Ogilvie et al., 2005). ROC analyses are meant to determine the *diagnostic or classification accuracy* of a test or variable; diagnostic or classification accuracy refers to the ability of a test or variable to distinguish between two binary categories (e.g., people with a disease or condition and people without a disease or condition). For instance, ROC analyses could test the accuracy of a variable like systolic blood pressure in diagnosing a medical condition like heart disease. Such a methodology is well positioned to answer similar questions in education. ROC

analyses would thus be essential in determining whether there is differential diagnostic accuracy between absence types and whether that accuracy is stable longitudinally, as it represents the most well tested and state-of-the-art methodological approach to this issue. A study applying ROC methods to this area of inquiry would constitute the first investigation to assess whether a major assumption of chronic absenteeism is scientifically sound.

Purpose of This Study

The United States educational system sits at an important crossroads. Despite decades of reforms to promote the educational well-being of students from racial/ethnic minority groups and students from low-income families, enormous gaps in educational outcomes still exist. Without the ability to address these gaps, evidence suggests that they will continue to grow and become even more difficult to ameliorate. As such, identifying and understanding the underlying forces causing these gaps is imperative. Recently, the discovery of an "attendance gap" that seems to mirror gaps in educational well-being has led to increased national attention on the importance of attendance as a means by which to address these inequities.

The impact of school attendance on student outcomes is well documented. From both a theoretical perspective and through empirical research, there is consensus that attendance has beneficial effects, while absences can be detrimental to students and have lasting repercussions on their educational trajectories. Due to the strong evidence of the beneficial effects of attending school, the United States has a long history of utilizing laws and policies to both encourage attendance in school and deter student absences. For

almost a century, all states have required children within a certain age range to attend some form of schooling through compulsory education laws. Additionally, many schools and school districts track an aggregate measure of attendance to monitor how many students are in school on an average day and provide a sense of how much funding each school requires. Truancy has also been a popular child-level indicator of attendance for many years, emphasizing the societal concern for children's safety and well-being and the need for constant adult supervision of young people.

In recent years, however, a new policy indicator of attendance has emerged—chronic absenteeism. With the passing of the Every Student Succeeds Act (2015) chronic absenteeism has become increasingly popular, with almost three-fourths of states adopting it as a measure of school quality and student success. Many schools and school districts have found chronic absenteeism to be particularly useful as an early warning indicator for young children at risk for future educational issues, such as low academic achievement. As an early warning indicator, chronic absenteeism allows schools to identify students missing a significant portion of school as early as kindergarten and provide them with additional supports and resources in order to prevent negative outcomes later in their educational trajectories.

In contrast to truancy, which only focuses on unexcused absences, chronic absenteeism rests upon the theory that all absences from school (i.e., both unexcused and excused absences) are *equally* detrimental to student outcomes and should thus be taken into account when identifying at-risk students. Unfortunately, there is no empirical evidence substantiating this theory. Only two studies exist in the peer-reviewed literature

that examine whether excused and unexcused absences in primary school are distinctively associated with academic achievement. The findings from these two studies indicate that unexcused absences are differentially predictive of low achievement. These studies assess whether there is a differential association between excused and unexcused absences and achievement outcomes but do not specifically test whether absence type is differentially accurate in determining future risk status (i.e., low academic achievement) and whether this differential accuracy is consistent across time. *The field, therefore, currently lacks a rigorous empirical assessment of whether absence type can be utilized to discriminate between students at risk for low academic achievement and students on track for academic success and whether that discriminatory ability is consistent across grade levels.* This lack of research is especially troublesome as more states adopt chronic absenteeism in response to ESSA's call for non-academic indicators of student well-being that are vertically aligned and can be tracked across time as students progress through school (Every Student Succeeds Act, 2015).

Thus, there is a pressing need to test the diagnostic accuracy of absence types on achievement status using an appropriate methodology that will assess whether the first major assumption of chronic absenteeism is scientifically sound. This study will test the classification accuracy of absence types on future achievement risk status by applying the most prevalent and extensively tested methodology within this field of statistical analysis—the receiver (or relative) operating characteristic (ROC) (Jordan, Glutting, Ramineni, & Watkins, 2010). While not widely used in education and related social

science fields, ROC analysis has been thoroughly tested and validated as the best means of evaluating the diagnostic accuracy of a test.

The purpose of this study, therefore, is to use appropriate statistical methods to test the first underlying assumption of chronic absenteeism—whether there is a qualitative difference between excused and unexcused absences as they relate to academic achievement status. The study has two major objectives: (1) to determine whether there is differential classification accuracy of excused and unexcused absences within the early primary years in predicting which students will ultimately be at risk for future academic achievement issues; (2) and to determine whether each absence type is consistent in the magnitude of its diagnostic accuracy across years (e.g., whether excused absences and unexcused absences become more or less useful as diagnostic classifiers across the early grades or if they remain the same across time). This study will address the following questions:

Research Objectives and Questions

Objective 1: Investigate the difference in diagnostic accuracies between excused and unexcused absences within the early primary grades.

- Question 1a: Is there a difference in the diagnostic accuracy of kindergarten excused and unexcused absences in classifying students as below basic in third-grade, standardized English and math performance?

- Question 1b: Is there a difference in the diagnostic accuracy of first grade excused and unexcused absences in classifying students as below basic in third-grade, standardized English and math performance?

- Question 1c: Is there a difference in the diagnostic accuracy of second grade excused and unexcused absences in classifying students as below basic in third-grade, standardized English and math performance?

Objective 2: Investigate the consistency of diagnostic accuracy within absence type across the early primary grades.

- Question 2a: Are the diagnostic accuracies of excused absences comparable in magnitude across kindergarten, first grade, and second grade in classifying students as below basic in third-grade, standardized English and math performance?

- Question 2b: Are the diagnostic accuracies of unexcused absences comparable in magnitude across kindergarten, first grade, and second grade in classifying students as below basic in third-grade, standardized English and math performance?

CHAPTER 2: METHODOLOGY

This chapter describes the research methodology that will be used to answer the five research questions outlined in the previous chapter. The research questions seek to determine whether there is a discrepancy in the classification accuracy of excused and unexcused absences in early elementary school and whether the discriminatory accuracy of excused and unexcused absences is stable across grade level. This section will review the data source for the study, the study sample, and the variables and measures used to answer the research questions. The chapter will also present a rationale for and description of the statistical analyses that will be utilized. Finally, the analytic plan for addressing the research questions will be reviewed.

Data Sources

The current study will utilize a subset of data from an existing administrative dataset. The dataset was obtained from a validation study conducted by Drs. John Fantuzzo and Katherine Barghaus. The goal of the first phase of the study was to determine whether a report card-based measure of kindergarten children's classroom engagement skills—called the Classroom Engagement Scale (CES)—exhibited evidence of internal and external validity (Henderson, Barghaus, Fantuzzo, Brumley, Coe, & LeBoeuf, 2018; Penn Child Research Center, 2017). The second phase of this work involves developing resources for kindergarten teachers and families that will help them support the skills measured by the CES in the classroom and at home. To inform the development of these resources, this phase of work involves a study to determine how students' classroom engagement skills are related to another form of student

engagement—school attendance. Because attendance can reflect both a student's and family's level of engagement or disengagement with the educational process, it is crucial to understand how it relates to the demonstration of CES competencies; furthermore, with the district's current emphasis on student attendance and push to reduce numbers of chronically absent students, it is essential to explore the potential negative effects absences and chronic absenteeism can have on the development of young students' engagement. This study emerged as a result of investigating the differential effects of absence types on engagement skills.

The data for the original validation study were obtained from digitized, administrative records directly from the School District of Philadelphia. The district keeps detailed records about all enrolled students and links records across years using a unique student identifier. Direct identifiers—such as name, social security number, and birth date—were removed from the dataset by the district for research use. A research proposal was submitted to and approved by the district for the original validation study. In addition, a separate proposal was submitted to and approved by the district to answer questions about attendance using the existing dataset, as reducing student absences is a high priority in Philadelphia and chronic absenteeism has become the State of Pennsylvania's measure of school quality and student success under ESSA (Bauer et al., 2018).

Both the validity study and attendance study were conducted on an entire cohort of students within the School District of Philadelphia and utilized data from kindergarten through third grade. As such, the variables in the dataset included a wide variety of

information (e.g., demographic characteristics, absence days, suspension information,

report card grades, standardized test performance, etc.). The pertinent variables will be

extracted from the full dataset to answer the study questions.

Study Sample

The study sample was created from an entire cohort of students from the School

District of Philadelphia from kindergarten through third grade (i.e., the study sample

includes data for the same group of students across kindergarten, first grade, second

grade, and third grade). For purposes of analysis, the data will be divided into four

subsamples: subsamples one through three will be used to assess whether there is a

difference in the classification accuracy of excused and unexcused absences within each

grade level; subsample four will assess whether the diagnostic accuracy of excused

absences and unexcused absences is stable in magnitude across grade levels. The data

will be partitioned and analyzed according to the research questions, such that

subsamples one through three will be used to address the first three research questions,

respectively, and subsample four will be used to address the fourth and fifth research

questions.

The first analytic subsample includes all students who were enrolled in

kindergarten full time beginning in academic year 2011-2012 and for whom third grade

standardized test score information was available. Students who entered or exited the

district at some point during the kindergarten year (for whom attendance data was only

available for the months they were enrolled in the district) were excluded from the study;

there were 10,525 children enrolled in the school district for the full academic year in kindergarten.

These kindergarten students, for whom full-year kindergarten absence data was available, were included in the sample only if they: (1) were still enrolled in the district in 2014-2015, (2) had progressed to third grade by that year, and (3) had standardized test score data available. In the 2014-2015 academic year, 8,713 students were enrolled in third grade for the full school year. This decline in enrollment between kindergarten and third grade is typical within the School District of Philadelphia and is related to the large number of student moves to charter schools in the city (which are not a part of the district and for which data are not available). Additionally, between 5% and 10% of children enrolled in kindergarten in 2011-2012 were retained in previous grades or promoted to higher grade levels, so were not enrolled in third grade in 2014-2015. Of the 8,713 students enrolled for the full academic year in third grade, complete standardized test information was available for 7,803 students or about 90% of students enrolled full time in third grade. The final analytic sample for kindergarten includes 6,800 students or about 87% of all students with standardized test score information in third grade in 2014-2015. Figure 1 presents a flow chart for the procedure used to create the first three analytic subsamples (using kindergarten as an exemplar).

A similar sample creation procedure was conducted for the first and second grade analytic samples. In first grade, 10,234 children were enrolled in the district full time and had full-year attendance records during the 2012-2013 academic year. The final analytic

```
┌──────────────┐     ┌──────────────┐     ┌──────────────┐     ┌──────────────┐
│   Starting   │     │ Kindergarten │     │ Kindergarten │     │              │
│ sample: all  │     │ students who │     │ students who │     │    Final     │
│  students    │     │ were not     │     │ were not     │     │  analytic    │
│ enrolled at  │ ──▶ │ enrolled for │ ──▶ │ enrolled in  │ ──▶ │ sample for   │
│ any point in │     │ the full     │     │ 2014-2015,   │     │ kindergarten:│
│ kindergarten │     │ academic year│     │ were placed  │     │              │
│              │     │ in           │     │ in another   │     │  n = 6,800   │
│              │     │ kindergarten │     │ grade other  │     │              │
│              │     │ are removed  │     │ than third   │     │              │
│              │     │ from the     │     │ during that  │     │              │
│              │     │ sample       │     │ year, and/or │     │              │
│              │     │              │     │ did not have │     │              │
│              │     │              │     │ third-grade  │     │              │
│              │     │              │     │ PSSA records │     │              │
│              │     │              │     │ are removed  │     │              │
│              │     │              │     │ from the     │     │              │
│              │     │              │     │ sample       │     │              │
└──────────────┘     └──────────────┘     └──────────────┘     └──────────────┘
```

Figure 1. Flow Chart of Sample Creation for the First Three Analytic Samples

sample for first grade (i.e., students who had full-year attendance data in first grade, had

standardized test score information in third grade, and were in third grade during 2014-

2015) includes 7,453 students. This is about 96% of all students with standardized test

score data in third grade in 2014-2015. In second grade, 9,261 students were enrolled in

the district full time and had full-year attendance data in 2013-2014. The final analytic

sample for second grade includes 7,254 students or about 93% of third grade students

with standardized test score information in the 2014-2015 academic year.

The fourth analytic subsample is longitudinal and included students who were

enrolled full time in kindergarten, first grade, and second grade, were enrolled in third

grade in 2014-2015, and had full standardized test score information in third grade.

Figure 2 presents a flow chart for the creation of the fourth analytic subsample. The final

subsample for the fourth and fifth research questions includes 6,223 students or about

80% of students with standardized test score information in the 2014-2015 academic

year. Figure 2 presents a flow chart for the creation of the kindergarten analytic

subsample.
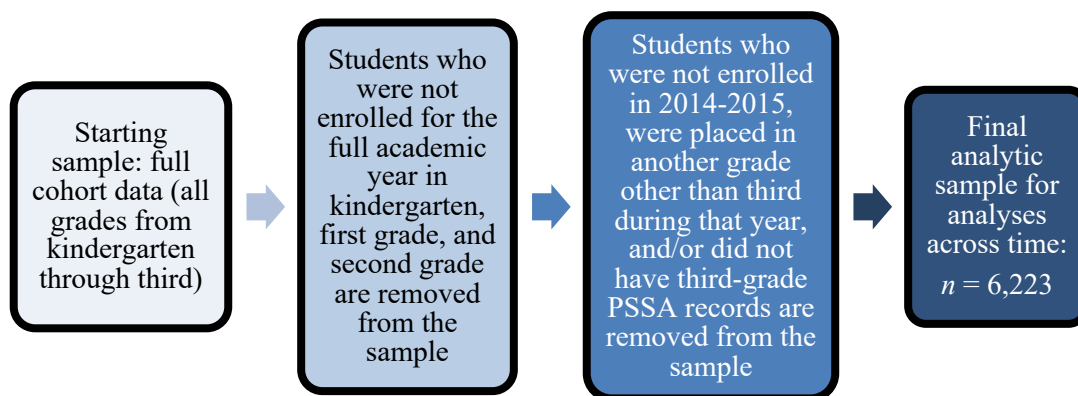
Figure 2. Flow Chart of Sample Creation for the Fourth Analytic Sample

Table 1 presents demographic characteristics for the data across academic years and the final analytic samples for kindergarten, first grade, second grade, and the sample across grades. While there are some minor differences between the original samples and the final analytic samples (e.g., the proportion of Black/African-American students enrolled in kindergarten full-time is 48.72% compared to the kindergarten analytic subsample where the proportion is 47.41%), the demographic characteristics between the original and analytic subsamples are comparable. About half of students in each of the final analytic subsamples are identified as Black/African-American, about 22% are identified as Hispanic/Latino, about 14% are White, 8% are Asian, and 7% are Multi-racial or Other. This distribution is consistent with the overall racial and ethnic distribution for the district. In addition, about 12% of students in the subsamples are English-language learners (ELLs) and about 80% qualify for free or reduced-price lunch (U.S. Department of Agriculture, 2018). Again, these percentages are consistent with district proportions of ELLs and children qualifying for free and reduce-price lunch.

	All Full-Time Enrolled Kindergarten Students (n=10,525)	All Full-Time Enrolled First Grade Students (n=10,234)	All Full-Time Enrolled Second Grade Students (n=9,261)	All Full-Time Enrolled Third Grade Students (n=8,713)	Third Grade Students with Complete PSSA Scores (n=7,803)	Final Analytic Sample for Kindergarten (n=6,800)	Final Analytic Sample for First Grade (n=7,453)	Final Analytic Sample for Second Grade (n=7,254)	Final Analytic Sample for Analyses across Grade (n=6,223)
Gender (male)	51.24%	51.17%	50.76%	51.22%	49.42%	49.21%	49.20%	49.13%	48.80%
Race/Ethnicity									
Black/African American	48.72%	50.99%	49.97%	49.12%	49.43%	47.41%	49.35%	48.66%	46.92%
Hispanic/Latino	19.93%	20.01%	21.19%	21.52%	21.66%	22.01%	21.67%	22.17%	22.42%
White	15.98%	13.81%	13.98%	13.84%	13.97%	15.06%	14.12%	14.35%	15.47%
Asian	6.94%	6.92%	6.76%	7.28%	7.98%	8.35%	7.96%	7.82%	8.10%
Multi-Racial/Other	8.42%	8.28%	8.10%	8.24%	6.47%	7.16%	6.91%	7.00%	7.09%
English-Language Learner	11.35%	11.87%	11.44%	10.85%	11.24%	12.99%	13.38%	12.48%	N/A
Special Education	7.46%	9.98%	12.99%	14.03%	10.50%	4.94%	6.83%	8.64%	N/A
Free or Reduced-Price Lunch Free Lunch through TANF	65.20%	73.46%	68.96%	71.15%	70.56%	65.40%	73.08%	68.25%	N/A
Free Lunch Application	5.17%	4.55%	6.88%	0.11%	0.13%	5.93%	4.87%	7.31%	N/A
Reduced Lunch	2.02%	1.57%	1.55%	0.03%	0.04%	2.15%	1.85%	1.70%	N/A
Not Applicable	27.61%	20.41%	22.61%	28.70%	29.27%	26.53%	20.20%	22.75%	N/A

Table 1. Student Characteristics by Grade

Between 5% and 9% of the analytic subsamples are students needing special education services, dependent on the grade level; this discrepancy of students identified as needing special education across years is consistent with the trend exhibited in the enrollment population (i.e., fewer students are identified as needing special education services in kindergarten and more students are identified over time). Note that ELL status, special education status, and free and reduced-price lunch status could not be determined for the fourth subsample, as these statuses may differ across year (i.e., a student could qualify for free lunch in kindergarten but not in first or second grade). Overall, the analytic samples are similar to the population samples.

Measures

Student Absences

The independent variables of interest relate to student absences from school. The School District of Philadelphia records student absences as "excused", "unexcused", or "due to out-of-school suspension". For the purpose of this study, only excused and unexcused absences will be considered. Absences caused by out-of-school suspensions relate to forced non-attendance due to behavioral reasons and are a fairly uncommon occurrence in the early primary years (less than 7% of all students in Philadelphia in kindergarten, first grade, and second grade). All absence information is officially recorded and entered into an electronic database by front office staff at each school. Front office staff are trained in data entry and coding procedures through training manuals and professional development. Teachers send front office staff their attendance list each day and pass along any communications they have with family members to indicate the

reason for students' absences. Though the district does not conduct data audits of attendance information to ensure data quality, attendance records are kept for all students who were enrolled in the district at any point and are diligently maintained as absences are tied to serious legal ramifications (e.g., referral to truancy court or a social service agency) (School District of Philadelphia, 2018).

Excused Absences. Within the school district of Philadelphia, daily attendance is required by all enrolled students, and the school year typically spans 175 to 180 days discounting district-wide school closures (e.g., closures for inclement weather). According to district policy, absences are generally considered excused if the reason for absence relates to illness or injury, religious holidays, authorized school activities (e.g., field trips), death in the family, or required appearance in court (School District of Philadelphia, 2018). Absences must also be accompanied by a family contact with administrative staff (e.g., front office staff) or the student's teacher to be considered excused. Families may call in or submit a written note or email to the school and must do so within three days of the child's absence (School District of Philadelphia, 2018). Even if the absence has been marked as "unexcused" because the family has not contacted the school, it can be changed to "excused" provided the family submits a written note within three days of the absence. The district does not require additional documentation (e.g., a doctor's note) for most excused absences other than notification from the family. Extended absences from school that occur consecutively (e.g., three consecutive days absent) require additional documentation depending on the situation (e.g., if the child is

sick for three days in a row, a doctor's note is required) (School District of Philadelphia, 2018).

 Unexcused Absences. Absences may be considered unexcused for four reasons: the cause of the student's absence, lack of communication by the family with the school, lack of documentation of the absence, or the student skipping school without the family's knowledge. For instance, if a student were absent from school due to a family vacation, that would be considered an unexcused absence, even if the family notified the school of the absence. In this case, the family may be aware of the student's absence but the absence does not meet the district's requirement for exemption. On the other hand, a student may be absent for a legitimate reason (e.g., illness), but the family does not contact the school within three days to notify them of the absence. Thus, that absence would similarly be marked as unexcused even though the reason for absence was legitimate and the family was aware of the absence. Additionally, if the family does not to provide appropriate documentation for consecutive absences (e.g., a doctor's note), these absences would also be marked as unexcused. The final scenario in which a student absence may be marked as unexcused is due to a student willfully skipping school without the family's or school's permission. This is typically what is thought of as "truant", but research suggests that it rarely applies to children in the early elementary grades (Chang & Romero, 2008; Klerman & Glasscock, 1996). Therefore, all unexcused absence within the context of this study are due to: (1) the reason for absence provided by the family (e.g., a vacation), (2) the lack of family communication with the school (e.g., no call or email from the family to the school), or (3) the lack of documentation provided

by the family for consecutive absences (e.g., no doctor's note for three days of consecutive absence).

Unfortunately, the district does not maintain records about the reason for unexcused absences; thus, it cannot be determined *why* the child's absence was marked as unexcused. The main difference between excused and unexcused absences in this dataset, then, is that excused absences necessitate contact between the family and the school, while unexcused absences may or may not involve family contact between the teacher and the school. Research indicates that many unexcused absences may be due to lack of family engagement in the schooling process, lack of regular contact between the family and the school, and/or the family's negative feelings or associations about their own educational experiences that may lead to avoidance of communication with school personnel (Chang & Romero, 2008; Gottfried, 2009). Thus, previous studies suggest that unexcused absences, especially among young students living in poverty, are primarily due to lack of family communication with the school (Jeynes, 2003; McNeal, 1999).

Academic Achievement

The Pennsylvania System of School Assessment (PSSA). Student achievement was measured using "the most widely-used outcome measure in education—the end-of-the-year state achievement test" (Rodrigues, 2017, p. 29). The PSSA is a standards-based, criterion-referenced assessment. All students in the State of Pennsylvania are required to take the PSSA across a number of grades unless specific exemptions apply (e.g., the child has a severe disability). Third grade is the first academic year in which students are tested

using the PSSA. Students are assessed using the PSSA in the spring of their third-grade year.

The PSSA third-grade assessment consists of English language arts (ELA) and mathematics subtests that are used in accordance with mandated federal reporting under the No Child Left Behind Act (NCLB, U.S. Congress, 2001). Reliability and validity of the PSSA scaled scores has been well established (Data Recognition Corporation, 2015), including internal consistency (r range .92 to .94) and validity evidence from factor analysis and differential item functioning.

In addition, the PSSA uses the most prevalent method for determining performance level cut scores—the Bookmark Method (Data Recognition Corporation, 2015). The Bookmark Method involves the mapping of items onto a proficiency distribution where cut scores are set. The method requires items to be empirically sorted by difficulty from least to most difficult using item response theory. A panel of experts then reviews the prearranged test items and places a "bookmark" between two items, such that students with a certain proficiency level would be able to answer the question before the bookmark but would not be able to answer the question after the bookmark (Karantonis & Sireci, 2006). The bookmarking method typically proceeds in rounds, where all items are bookmarked and then there is discussion among the experts; there are usually three rounds of review with each round designed to foster increasing convergence among panelists (Mitzel, Lewis, Patz, & Green, 2001).

PSSA subtests are divided into four performance levels (from least to most proficient): below basic, basic, proficient, and advanced. Below-basic status reflects

inadequate academic performance and minimal display of the skills required to meet the academic standards for that year. Basic status indicates marginal academic performance and a limited display of the skills necessary to meet the academic standards. Proficient status reflects satisfactory academic performance and represents an adequate display of skills needed to meet the standards. Finally, advanced status indicates superior academic performance and an exemplary display of the skills necessary to meet the on-grade academic standards.

For the 2015 ELA subtest of the PSSA, the cut scores for the performance levels were as follows: scaled scores from 600 to 904 were considered below basic, scores from 905-999 were basic, 1000 to 1142 were considered proficient, and scores from 1143 to 1586 were advanced (Data Recognition Corporation, 2015). Similarly, for the 2015 mathematics subtest, the cut scores for the performance levels were as follows: 600 to 922 were considered below basic, scores from 923-999 are basic, 1000 to 1109 were considered proficient, and scores from 1110 to 1594 were advanced (Data Recognition Corporation, 2015). In 2015, 13.3% of students in Pennsylvania scored at the below-basic level on the third-grade ELA subtest, 24.6% scored at the basic level, 49.1% scored at the proficient level, and 13% scored at the advanced level (Pennsylvania School Board Association, 2015). Similarly, 27.9% of students scored at the below-basic level on the third-grade mathematics subtest in 2015, 23.6% scored at the basic level, 28.5 scored at the proficient level, and 20% scored at the advanced level.

Analytic Method

The current study will utilize the most advanced statistical technique available to assess the classification accuracy of a predictor—the receiver (or relative) operating characteristic (ROC). ROC analyses have been widely utilized across a multitude of fields but are less prevalent in the educational literature. Before discussing how this methodology will be utilized to answer the four research questions, this section will describe how ROC analysis was developed, the rationale for the use of ROC analysis in educational science, and fundamental aspects of ROC analysis necessary to understand its application.

The History of ROC Analysis and the Rationale for Its Use in Education Research

ROC analysis originated out of Signal Detection Theory (or SDT), which posits that within data, there are patterns that convey information (or signals) and patterns that convey randomness (or noise). The goal of SDT is to separate the signal from the noise. As McFall and Treat (1999) assert, "Historians trace the roots of contemporary SDT to…work on hypothesis testing and statistical inference, but the underlying probabilistic concepts can be traced backwards chronologically, if not genealogically, more than 200 years" (p. 226). While the roots of SDT can be traced back many years, modern invocations of SDT, including ROC analysis, emerged in relation to radio signals, and specifically, how radio signals could be reliably discriminated from background noise (Pintea & Moldovan, 2009; Smolkowski & Cummings, 2015; Zou, O'Malley, & Mauri, 2007). McFall and Treat (1999) note, "Engineers originally developed ROC analysis to quantify how well an electronic receiver detects electronic signals in the presence of noise; ROC analysis acquired its name from its application to radar detection problems

during World War II" (p. 229-230). ROC analysis was soon adopted in the biomedical field in the 1960s to promote the diagnostic accuracy of medical tests in discriminating between patients with diseases or diagnoses and those unaffected by them (Pintea & Moldovan, 2009). ROC then became a popular technique among psychologists to help with clinical diagnosis. The disease or condition is thus the "signal" that must be discerned amidst the "noise."

Decades of research confirm that ROC analysis is the best means of evaluating the diagnostic accuracy of a test. As Jordan et al. (2010) note, "ROC is 'the state-of-the-art method' for describing the diagnostic accuracy of a test and is 'recognized widely as the most meaningful approach to quantify the accuracy of diagnostic information and diagnostic decisions' (Metz & Pan, 1999, p. 1)" (ps. 184-185). In fact, meta-analyses from the biomedical field reveal that ROC methodology has been widely employed to assess the classification accuracy of a variety of tests for many serious medical conditions. For example, ROC analyses have been used in thousands of studies to test: the accuracy of positron emission tomography (PET) scans in identifying Alzheimer's Disease (Patwardhan et al., 2004); the diagnostic utility of a protein in identifying acute appendicitis (Hallan, & Åsberg, 1997); the comparative accuracy of several biological indices in determining the risk of heart disease (Lee et al., 2008); the ability of radiologic scans to detect prostate cancer (Engelbrecht et al., 2002); and the diagnostic accuracy of cell specimens in detecting human papillomavirus (HPV) (Ogilvie et al., 2005). Furthermore, Baker (2003) stated in an article published in the *Journal of the National*

Cancer Institute that "ROC curves should be the primary method for evaluating the performance of early detection tests of cancer" (p. 511).

Despite its widespread application in fields such as biomedicine, ROC analysis is little used within certain disciplines—such as education. As Swets et al. (2000) state, "Diagnostic problems abound for individuals, organizations, and society. The stakes are high, often life and death. Such problems are prominent in the fields of health care, public safety, business, environment, justice, education, manufacturing, [etc.]….this incipient discipline has been demonstrated to improve diagnosis in several fields, but is nonetheless virtually unknown and unused in others" (p. 1). Rather than utilizing ROC analysis, statistically significant mean differences are typically the benchmark used to determine whether a test is useful at discriminating between two groups in educational research (Jordan et al., 2010; Wilson et al., 2015). Mean scores on particular outcomes are compared between different groups. Statistically significant differences are touted as evidence of the usefulness of a particular variable or test in discriminating between two groups. This methodology is, however, problematic. Jordan et al. (2010) note:

> Although the mean score differences indicate that *groups* can be discriminated, this conventional validity approach cannot be uncritically extended to conclude that mean group differences are distinctive enough to differentiate among *individuals*….In other words, group mean differences are necessary but not sufficient for making accurate decisions about individuals because they do not take into account the overlap in score distributions between groups (p. 184).

Thus, statistically significant mean differences are insufficient in determining whether the variable or test is an accurate diagnostic classifier of any given individual. Studies have shown that ROC methods are superior, in terms of evaluating the diagnostic accuracy of a test, to group mean differences and other more simplistic methods for determining

accuracy, like the ratio of false positives to true positives (Bossuyt et al., 2003; Mossman, 1994). It is, therefore, imperative, that studies designed to assess the validity of a variable or test in discriminating between groups go beyond the traditional methods of comparing mean differences and utilize a more advanced methodology that directly assesses classification accuracy.

This shift toward a more advanced methodology is particularly critical in the field of education where diagnostic decisions for teachers, school psychologists, administrators, and school districts abound. In most cases, these diagnostic decisions are based on professional judgments rather than on empirical evidence, despite the fact that "studies of the diagnostic decision-making process suggest that judgements grounded on data, statistical models, and even informal prediction models outperform those based on intuition alone" (p. 41, Smolkowski & Cummings, 2015). The goal of educational research in this area should, therefore, be to maximize classification accuracy so that school and district staff can capitalize on the power of data and diagnostic systems, minimize the time required for them to categorize students based on professional judgments, limit biases from factors unrelated to student outcomes, and ensure that students in need receive appropriate services (Smolkowski & Cummings, 2015). This shift toward employing more rigorous methodology also reflects the changing nature of the educational landscape where researchers, practitioners, and policymakers are increasingly urged to take advantage of high-quality science. The Every Student Succeeds Act (2015), for instance, references the need for "evidence based" research, practices, and policies 70 times. Thus, educational science should seek to employ the

most rigorous research methods—such as ROC analyses—to determine diagnostic

accuracy, such that teachers, schools, and policymakers can better discern which students

are at risk and need support services.

The Basic Principles of ROC Analysis

Before delving into how ROC analysis will be applied to the research questions, a

fundamental understanding of the basic concepts of this method is necessary. As noted

before, "Receiver Operating Characteristic (ROC) analysis is a procedure used in

assessing diagnostic properties of tests, namely in assessing the way various measures

generally discriminate between different categories of subjects" (p. 49, Pintea &

Moldovan, 2009). ROC analysis is the preferred methodology when the outcome of

interest is a binary variable (e.g., at-risk students vs. students not at risk) and the predictor

is continuous (e.g., numerical score on a risk assessment inventory) (Gönen, 2007).

ROC analysis is based on the idea that given a predictor measure, there will be

two underlying distributions that correspond to the two groups of interest (i.e., people

with the condition and people without the condition). Provided there is no overlap

between the two distributions, the predictor will be able to perfectly discriminate between

the two groups (Smolkowski & Cummings, 2015; Swets et al., 2000. Figure 3 provides

an illustration of this theoretical concept.

Given a certain cut-off point, the predictor measure illustrated here would be able

to perfectly discriminate between those with the condition and those without the

condition. For example, if students that will eventually read at grade level by third grade

received scores ranging from 51-100 on a kindergarten reading test and students that

Figure 3. Theoretical Distribution of Populations

cannot read at grade level by third grade score between 0 and 49 on the same

kindergarten reading test, then a cut-off score of 50 would perfectly discriminate between

the two groups. The reading test would thus have 100% diagnostic accuracy. In reality,

the distributions amongst the two groups typically overlap in relation to the predictor

variable. Even among applications in the biomedical field, there is usually some degree

of overlap between the two distributions (Gönen, 2007; Youngstrom, 2014; Zweig &

Campbell, 1993). Figure 4 provides an illustration of what these distributions realistically

look like, as compared to the theoretical model.

Because the populations with and without the condition will typically overlap in

their distribution along the predictor measure, the cut-off point will inevitably produce

four possible outcomes: a true negative (TN), where the individual is identified as not

having the condition by the predictor and does not actually have the condition; a true

positive (TP), where the individual is identified as having the condition by the predictor
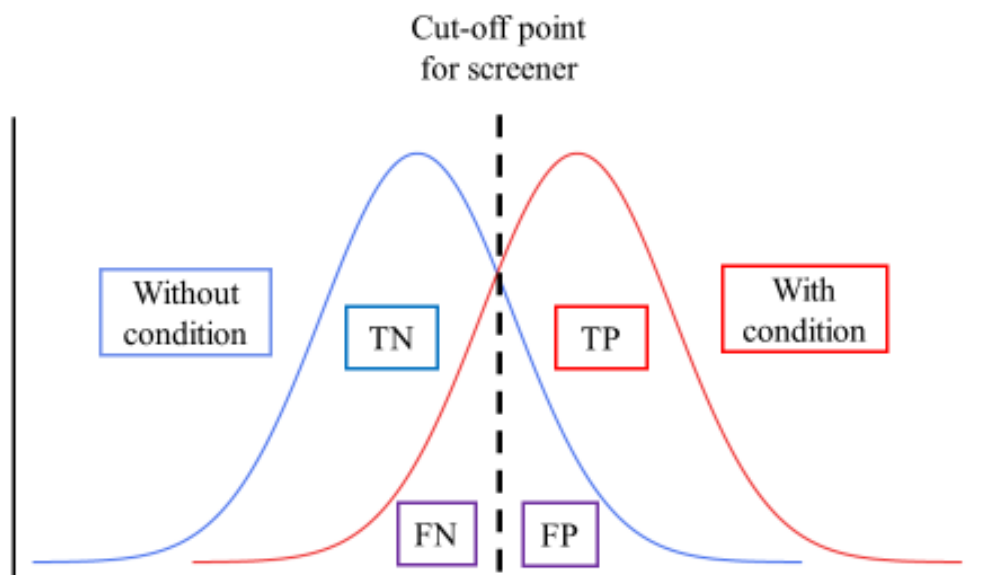
Figure 4. Realistic Distribution of Populations

and actually has the condition; a false negative (FN), where the person is identified as not

having the condition by the predictor but actually does have the condition in reality; and a

false positive (FP), where the person is identified by the predictor as having the condition

but in reality does not have the condition (Pintea & Moldovan, 2009). For example, if all

students who eventually read at grade level by third grade received scores ranging from
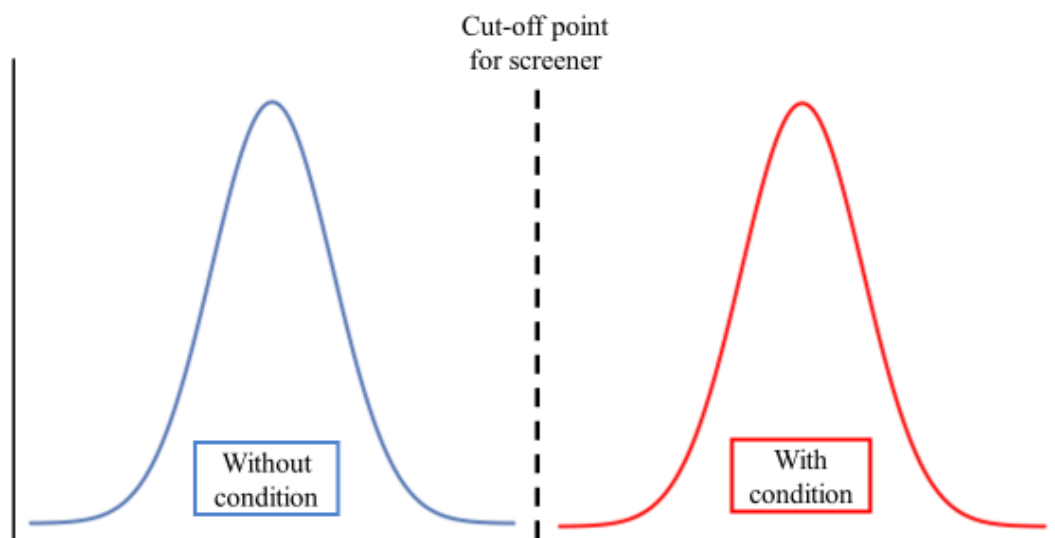
38-100 on a kindergarten reading test and all students who cannot read at grade level by

third grade score between 0 and 62 on the same reading test, a cut-off score of 50 would

not be able to perfectly discriminate between the two groups. This cut-off point would

produce all four outcomes (i.e., true negatives, true positives, false negatives, and false

positives), and the accuracy of the test would thus need to be evaluated in terms of these

outcomes; in addition, the cut-off score would potentially need to be modified to

maximize true positive or true negative results, depending on the context in which the test

is being used.

These four possible outcomes form the basis for ROC analysis. With this

information, the true positive rate of the test (i.e., the sensitivity of the test) and the true

negative rate of the test (i.e., the specificity of the test) can be calculated. Table 2 depicts

the four possible outcomes along with the accompanying calculations of sensitivity and

specificity.

Table 2. Possible Outcomes Determining Sensitivity and Specificity

		Actual Condition	
		Positive	Negative
Test	Positive	True positive (TP)	False Positive (FP)
Result	Negative	False negative (FN)	True negative (TN)
	Total	Number with condition (C+)	Number without condition (C-)

$$\text{Sensitivity} = \frac{TP}{C+} \qquad \text{Specificity} = \frac{TN}{C-}$$

As Pintea and Moldovan (2009) note, "*Sensitivity*, also called the *true positive*

rate (when expressed as a percentage) is defined as the probability that test result will be

positive when the disorder is present. *Specificity*, also called the *true negative rate* (when

expressed as a percentage), represents the probability that a test result will be negative

when the disorder is not present. These two indicators are essential for ROC curves

analysis" (p. 52). Sensitivity and specificity likewise correspond to specific cut points of

the predictor. It is thus possible to achieve 100% sensitivity or 100% specificity with a

given cut score. However, the higher the sensitivity, the lower the specificity and vice

versa (Youngstrom, 2014). Thus, a test that captures all existing true positives would also

capture the most false positives, just as a test that captures all true negatives would also

capture the most false negatives. Given the example of the kindergarten reading test

(where the range of scores for those eventually reading on grade level was 38-100 and the range of scores for those not reading on grade level was 0-62), if the cut score was set at 37, it would pick up all the students who eventually read on grade level by third grade (i.e., all students who scored from 38 to 100). It would also, however, classify a lot of students who will not be able to read on grade level by third grade (i.e., all those who scored in the 38 to 62 range) as being on track. Thus, sensitivity and specificity are inversely related and involve tradeoffs in diagnostic accuracy depending on the cut score chosen (Smolkowski & Cummings, 2015).

One of the major benefits of ROC analysis is that it does not look at a single cut score to evaluate the diagnostic accuracy of a predictor. Rather, ROC analysis involves considering all true positive rates and all false positive rates for all possible cut scores of the diagnostic assessment. ROC analysis plots these rates graphically, which serves as a useful visual representation of the overall accuracy of the predictor. Pintea and Moldovan (2009) describe this visual representation: "ROC graphs are bidimensional representations of the sensitivity (also called the *true positive rate* – on the X axis) and 1-specificity (also called the *false positive rate* – on the Y axis), corresponding to each possible cut-off point (classifying value). In other words, they represent tradeoffs between benefits (true positives) and costs (false positives)" (p. 53). Sensitivity and 1-specificity plotted on a bidimensional graph typically forms a curve, where the closer the curve is to the upper left-hand side of the graph, where sensitivity is maximized and the false positive rate is minimized, the more accurate the predictor is. Conversely, the flatter the curve (i.e., the closer the curve is a diagonal line where y = x), the less accurate the

test. This diagonal line where the true positive rate is equal to the false positive rate represents a test that is no better than random chance. For example, if the kindergarten reading test produced just as many true positives as false positives, then it would perform no better than using a coin flip to categorize students at risk for future reading difficulties. Although it is possible for the ROC curve to dip below the diagonal line (i.e., the test performs *worse* than random chance), instances of this occurring are rare. Thus, the closer the ROC curve is to the diagonal line, the worse its discriminatory accuracy. Figures 5 and 6 show a visual representation of two ROC curves, one with high diagnostic accuracy (the green line in Figure 5) and one with low diagnostic accuracy (the red line in Figure 6).
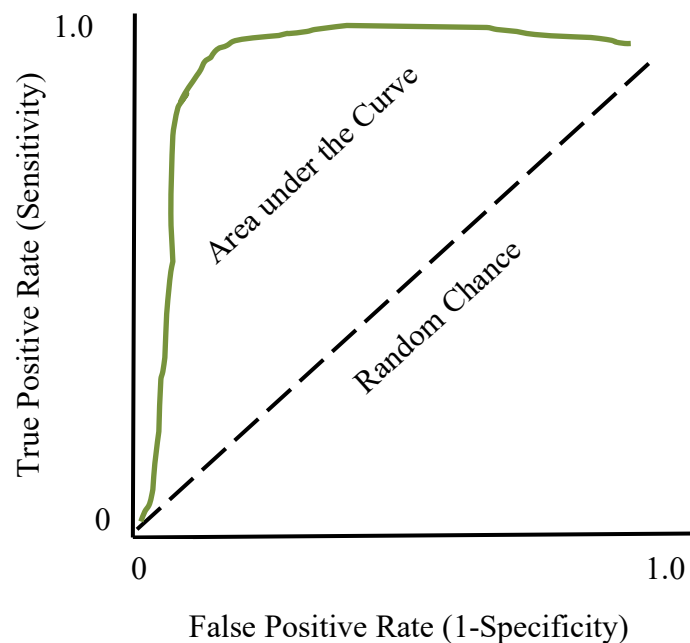
Figure 5. ROC Graph of a Predictor Measure with High Diagnostic Accuracy
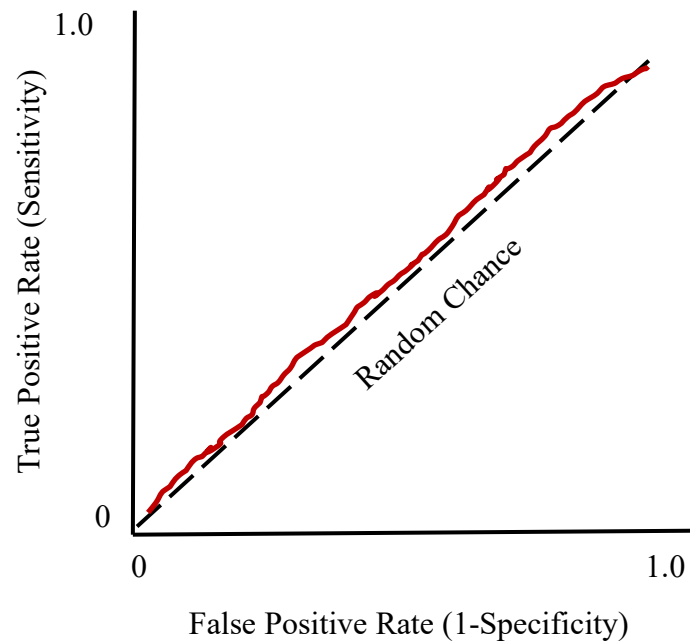
Figure 6. ROC Graph of a Predictor Measure with Low Diagnostic Accuracy

In addition to a visual representation of diagnostic accuracy, ROC analysis also produces a quantitative value of diagnostic accuracy called the area under the curve (AUC or *c* statistic). As Pintea and Moldovan (2009) state, "As concerning the statistical indicators of the ROC curve, the primary statistic derived from the ROC is the *area under the curve* (AUC). The total area under the ROC curve is a measure of the overall performance of the diagnostic test: the larger the area, the better the performance" (p. 54). The area under the curve is calculated using the trapezoidal rule, which involves dividing the area under the curve into a series of strips of equal width, calculating the area of each trapezoidal-shaped strip, and summing the strips (Delong, Delong, & Clarke-Pearson, 1988; Gönen, 2007). The resulting numerical value ranges from 0 to 1, with 0 representing no diagnostic accuracy and 1 representing perfect diagnostic accuracy. While the AUC can range in value from 0 to 1, an AUC of 0.50 would fall along the

diagonal line indicating that the test performs no better than random chance. Thus, AUC

values of less than 0.50 are rarely seen and would indicate that the given test is not useful

in discriminating between two groups. Zou et al. (2007) confirm that "the AUC is an

overall summary of diagnostic accuracy. AUC equals 0.5 when the ROC curve

corresponds to random chance and 1.0 for perfect accuracy" (p. 656). Another useful way

of understanding the AUC is that it represents the probability that a randomly selected

person with the condition would have a higher score on the predictor variable than a

person without the condition. Thus, if the kindergarten reading test produced an AUC of

0.51, someone who ultimately reads on grade level would have only a 51% chance of

having a higher score on the kindergarten reading test than someone who ultimately does

not read on grade level by third grade; this means that the kindergarten reading test would

be only marginally better than random chance and would not be a good classifier for

students at risk for later reading difficulties.

While there are various heuristics for interpreting the size of the AUC in

relationship to the diagnostic accuracy of the test, benchmarks differ by fields. For

instance, in biomedical or engineering applications, an AUC of 0.80 or above would be

considered strong, whereas that would be inappropriate in another field, such as

education (Youngstrom, 2014). In contrast, Rice and Harris (2005), in a widely cited

paper, translated AUC into measures of effect size (Cohen's *d*), where: an AUC of 0.556

corresponds to an effect size of 0.20, which is considered small; an AUC of 0.639

corresponds to an effect size of 0.50, which is considered medium; and an AUC of 0.714

corresponds to an effect size of 0.80, which is considered large. This heuristic is more

appropriate for the current study considering that the average effect size in education for

intensive interventions aimed at improving student achievement is between 0.20 and 0.51

(i.e., the small to medium range of Cohen's *d*) (Hill, Bloom, Black, & Lipsey, 2008).

Because ROC methodology is rarely used in education, there are not well-established

rules of thumb for judging the magnitude of the AUC and thus more flexible benchmarks,

like those proposed by Rice and Harris (2005) are more applicable.

      While a generalized rule of thumb for judging the size of the AUC poses a

challenge in education research, the AUC *can* be used to make relative judgments about

the diagnostic accuracy of two predictor measures on the same group of people. Figure 7

displays the results for two ROC curves and their accompanying AUCs tested on the

same group of people. While Test B has a higher ROC curve and larger AUC value than
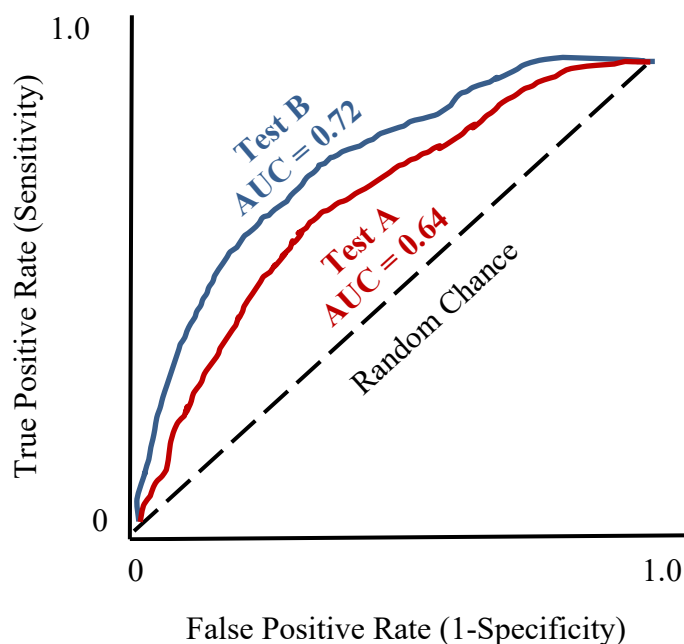
Figure 7. Comparison of Two ROC Curves

Test A, that does not necessarily mean that Test B performs significantly better in terms of classification accuracy. A statistical method is needed for determining whether the difference in these two curves is significant.

Fortunately, there is a widely used statistical test that can be applied to a situation where two ROC curves need to be compared (Gönen, 2007). Demler, Pencina, and D'Agostino (2012) describe this approach for comparing the diagnostic accuracy of two tests: "A widely used test to compare the difference between two AUCs relies on the method developed in a seminal paper by DeLong et al. (henceforth 'the DeLong test'). It provides a confidence interval and standard error of the difference between two (or more) correlated AUCs. This procedure has been frequently applied to test the incremental gain in model discrimination" (p. 2). The DeLong test utilizes a nonparametric approach by using a theory developed for generalized U-statistics; this method estimates a covariance matrix "and the resulting test statistic has asymptotically chi-square distribution" (p. 844, Delong et al., 1988). Utilizing the DeLong test, the null hypothesis would be that there is no discriminatory difference between the two diagnostic measures; the alternative hypothesis would be that there is a discriminatory difference between the two predictors. A chi-square statistic can be used to determine whether there is a difference in the diagnostic accuracy of the two tests (Hajian-Tilaki, 2013). Thus, the relative strength of two ROC curves can be empirically determined, regardless of the heuristic applied to judge the magnitude of each AUC.

The ability to compare two correlated ROC curves speaks to one of the many benefits of this type of analysis. This form of ROC analysis uses an empirical approach,

rather than a parametric approach. In a parametric approach, the data represents a sample of information drawn from a larger population, about whom certain assumptions must be made. In contrast, the type of ROC analysis utilized by the DeLong approach is non-parametric and does not necessitate any assumptions about the distribution of the data. As Zou et al. (2007) note, "An advantage of this method is that no structural assumptions are made about the form of the plot, and the underlying distributions of the outcomes for the two groups do not need to be specified" (p. 655). Thus, there is no need to determine whether the data meet any specific criteria (e.g., a normal distribution) in order to use this methodology.

Another benefit of ROC analysis is that it allows for the determination of a particular cut score based on the context. Because there is no true optimal cut-off point, trade-offs between sensitivity and specificity can be considered as they relate to outcomes. For example, when considering the cut score for the kindergarten reading test, one would need to consider the benefits of true positives and true negatives and the cost of false positives and false negatives. For instance, in one context, a false negative may be more dangerous than a false positive because at-risk students wouldn't receive the additional supports they need to succeed. In another context, false positives may be more problematic, as providing additional resources to students who don't need them is time- and cost-intensive, especially within schools where personnel and funding are limited (Smolkowski & Cummings, 2015). Thus, ROC analyses allow for these considerations in determining which cut score makes the most sense for a given context.

There are two potential drawbacks of ROC analysis that should be considered. The first drawback concerns the criterion variable or diagnosis (e.g., student at risk for academic difficulties) and is typically referred to as "the gold standard problem" (McFall & Treat, 1999). The gold standard problem reflects the issue that the criterion variable may not reflect the true status of the individual. For instance, to determine the accuracy of a test of kindergarten reading ability, an outcome that reflects future reading difficulty is needed. This outcome is likely to be captured by another test of reading ability. This test, no matter how carefully it is designed and administered, will contain measurement error. That is to say, the test cannot perfectly capture reading ability. Thus, the accuracy of the kindergarten screener is being judged based on an outcome measure that does not perfectly reflect true reading ability. While this problem abounds in all fields that utilize ROC analyses, even in the biomedical literature where the gold standard is still considered difficult to obtain (Zou et al., 2007), it is particularly important to consider in educational science, where measurement standards are of variable quality. It is thus important to select well tested, high-quality criterion variables, such as state-wide standardized tests, that have undergone a rigorous development process in an attempt to limit measurement error.

The other potential drawback of ROC analysis—called spectrum—should also be considered in the context of educational research (Hajian-Tilaki, 2013). Spectrum refers to the range of the condition in the people being studied. It is important that the subjects being studied represent a broad range in relation to the severity of the diagnosis. For example, validating the diagnostic accuracy of the kindergarten reading test would

require that the outcomes of the subjects (i.e., their future reading test performance) show breadth. In other words, it would not be useful to have children with only moderately low scores on the outcome test; children with extremely low, low, and moderately low scores on the outcome test would be necessary for the subject pool. The same applies to subjects without the condition; thus, children with moderately high, high, and extremely high scores on the outcome test would be necessary for the subject pool. This is typically more of a concern in the biomedical field, where there can be a lack of heterogeneity in the severity of the condition for a number of reasons (e.g., someone with a mild case may not have been diagnosed yet) but is less of a concern in fields like education when a normed measure is being used as the criterion variable (e.g., a standardized test). While the potential drawbacks of this methodology should be examined within the context of the study, ROC analysis still remains the most sophisticated approach to determining the classification accuracy of a predictor and should be more readily considered for its application to questions of educational science.

Procedure for Conducting ROC Analyses

There are several steps that must be taken to conduct ROC analyses to determine the difference in diagnostic accuracy of two predictor measures. Before analyses are conducted, the predictors of interest should be defined and an appropriate criterion variable related to those predictors should be selected; the criterion variable must be binary or be transformed into a binary variable if necessary (Youngstrom, 2014). Once the predictor and criterion variables have been selected, an appropriate sample should be identified.

The first step of the ROC analysis is to produce the ROC curve and AUC for each relevant predictor and determine whether there is a statistically significant difference between the AUC of the predictor and that of random chance (AUC = 0.50) (Pintea & Moldovan, 2009; Youngstrom, 2014). If the AUC is not significantly different from random chance, it would not be considered useful as a diagnostic variable. If the AUC is significantly different than random chance, it can be considered for use as a diagnostic instrument and should can be compared to other predictor variables that are significantly different from random chance. The AUC can also be compared to a pre-established heuristic, such as the thresholds outlined by Rice and Harris (2005).

The second step involves comparing the AUC for each predictor variable that is significantly different than random chance. The diagnostic performance of these variable can then be assessed through use of the DeLong test. If there is a statistically significant difference between the AUC of the two predictors, the variable with the lower AUC should be removed from consideration in favor of the variable with the higher AUC.

The final step of ROC analysis is to optimize cut-score thresholds for the predictor variables with the highest AUCs (Youngstrom, 2014). As mentioned in the previous section, this optimization depends heavily on the intended use of the predictor measure, the context in which it will be used, and the relative costs and benefits of correct classification vs. misclassification (Zweig & Campbell, 1993). The process for determining optimal cut scores is often time intensive and requires many contextual considerations.

Analytic Approach to Research Questions

The analytic approach to the research questions will follow steps one and two in the procedures outlined above. Step three will not be addressed in this study, as it is outside the focus of the research questions. The approach to the five research questions will be identical. The first three research questions will utilize the outlined approach at a static time point (i.e., the end of kindergarten, first grade, and second grade) to determine whether there is a difference in classification accuracy *between* excused and unexcused absences at each grade level. The fourth and fifth research questions will utilize the procedure above to determine whether the classification accuracy *within* absence type (i.e., both excused absences and unexcused absences) is consistent in magnitude across kindergarten, first grade, and second grade.

The diagnostic accuracy of absences will be tested against two criterion variables for each analytic subsample: below-basic performance on the ELA PSSA and below-basic performance on the mathematics PSSA. The below-basic designation is the lowest performance level on the PSSA and indicates inadequate academic performance and minimal display of the skills required to meet third-grade academic standards. It is, thus, a negative educational outcome that signifies academic difficulty in third grade. Furthermore, standardized test scores, as well as third-grade reading and math skills, have been linked to a host of future negative academic and behavioral outcomes (Annie E. Casey Foundation, 2010; Wieman, 2007), so this negative educational outcome may also signal future negative outcomes. Below-basic performance on ELA and mathematics will thus be considered the "condition" or "disease" in this study as it represents a negative outcome. All students scoring below basic in ELA (a scaled score between 600 and 904)

will be coded as a "1" in the dataset, and all students scoring above 904 will be coded as

a "0". Similarly, all students scoring below basic in math (a scaled score between 600 and

922) will be coded as a "1" in the dataset, and all students scoring above 922 will be

coded as a "0".

For the first three analytic samples (corresponding to the first three research

questions) there will be two predictor variables of interest (i.e., excused absences and

unexcused absences at each grade level) and two criterion variables (i.e., below basic in

ELA and below basic in math third grade). The analyses will test the diagnostic accuracy

of both predictor variables on both criterion variables. Table 3 provides an overview of

the predictors and outcome variables to be tested with each analytic sample for research

questions 1a, 1b, and 1c.

Table 3. Analytic Strategy for Research Questions 1a, 1b, and 1c

	Kindergarten Analytic Sample	**First Grade Analytic Sample**	**Second Grade Analytic Sample**
Predictor Variables	Excused absences in kindergarten	Excused absences in first grade	Excused absences in second grade
	Unexcused absences in kindergarten	Unexcused absences in first grade	Unexcused absences in second grade
Criterion Variables	Below basic in ELA	Below basic in ELA	Below basic in ELA
	Below basic in math	Below basic in math	Below basic in math

Each predictor variable will be tested against each criterion variable. ROC graphs

and AUCs will be produced for these predictor variables and each corresponding criterion

variable across the analytic subsamples. If the AUC for a particular predictor variable is

not statistically significantly different from random chance ($AUC = 0.50$), then it will not

be considered for comparison against the other predictor variable in the next stage of the

analysis. Additionally, the magnitude of the AUC will be judged by the heuristic

suggested by Rice and Harris (2005). AUCs below 0.555 will be considered negligible,

AUCs between 0.556 and 0.638 will be considered small, AUCs between 0.639 and

0.713 will be considered medium, and AUCs of 0.714 and above will be considered

large. AUCs in the small or medium range will be considered desirable, as they are

comparable to the average effect size of intensive educational interventions aimed at

improving student achievement per Hill et al. (2008), whereas AUCs in the large range

are unlikely to be observed given how rare it is to see large effect sizes in educational

research.

The AUC for each of the predictor (and corresponding criterion) variables that has

met the significance tests of the previous step will then be compared using the DeLong

test. It is possible that certain sets of predictors will not be compared using the DeLong

test. For example, if the AUC produced by the ROC curve for excused absences in

kindergarten on math performance is not significantly different from random chance, and

the AUC for unexcused absences in kindergarten is significantly different, no comparison

would be necessary; unexcused absences would be the more accurate classifier in this

case. In these instances, the DeLong test will not be performed. In instances where there

are two AUCs to compare, a chi-square statistic with a p value < .05 will be considered

statistically significant and thus indicate that one of the variables exhibits better

diagnostic accuracy than the other. This step will be repeated across the first three

analytic samples.

Finally, to address the last two research questions, this same set of steps will be applied to the fourth analytic subsample. This requires testing the diagnostic accuracy of each absence type across grade levels (e.g., kindergarten excused absences, first grade excused absences, and second grade excused absences) to determine the consistency of their classification accuracy for each criterion variable. These analyses will thus be performed *within* absence type *across* grade level (as opposed to the previous three analyses which were conducted *across* absence types *within* grade level). Table 4 depicts the analytic strategy for the research questions 2a and 2b.

Table 4. Analytic Strategy for Research Questions 2a and 2b

	Analytic Sample for Analyses across Grade	
Predictor Variables	Excused absences in kindergarten	Unexcused absences in kindergarten
	Excused absences in first grade	Unexcused absences in first grade
	Excused absences in second grade	Unexcused absences in second grade
Criterion Variables	Below basic in ELA	Below basic in ELA
	Below basic in math	Below basic in math

Again, all predictor variables will be tested against each criterion variable. ROC graphs and AUCs will be produced for these predictor variables and each corresponding criterion variable; only AUCs significantly different than random chance will be considered for comparison using the DeLong test. For example, if the AUC for excused absences on ELA performance was 0.65 in kindergarten, 0.67 in first grade, and 0.60 in second grade (all of which were significantly different than random chance), the DeLong test will determine if there are significant differences in classification accuracy across

years (i.e., first grade excused absences have higher diagnostic accuracy than second grade excused absences). Again, a DeLong test that produces a chi-square statistic where $p < 0.05$ indicates that the there is a significant difference in classification accuracy. This analysis will provide a sense of whether diagnostic accuracies are variable across time, an important factor to consider when assessing classification accuracy.

Summary of Methods

The current study will explore whether there is a difference in classification accuracy between excused and unexcused absences in early primary school and whether the classification accuracy of excused and unexcused absences is stable across these early grades. The study will utilize ROC analysis to determine the degree of classification accuracy of each absence type and utilize the DeLong test to determine whether classification accuracy between excused and unexcused absences within each grade level is significantly different. The DeLong test will also be used to determine whether the classification accuracy of excused and unexcused absences is significantly different across grade levels. By evaluating the diagnostic accuracy of absence types using a rigorous methodology, this study seeks to ensure that widespread policy indicators used to identify at-risk students, such as chronic absenteeism, are backed by sound science. This, in turn, will ensure that American schools are better able to identify and serve their most vulnerable children.

CHAPTER 3: RESULTS

This chapter presents the findings of the current study to determine (1) the classification accuracy of absence types in kindergarten, first grade, and second grade on academic achievement status and (2) the stability of this classification accuracy across time within each absence type. The chapter will first present descriptive information about the four analytic subsamples in relation to the predictor and criterion variables. Second, the findings related to the classification accuracy of absence type within each grade (i.e., research questions 1a, 1b, and 1c) will be examined. Finally, this section will present the findings related to the consistency of the magnitude of classification accuracy within each absence type across kindergarten, first grade, and second grade (i.e., research questions 2a and 2b).

Descriptive Statistics for the Four Analytic Subsamples on Key Variables

Table 5 presents the mean and standard deviation of the two predictor variables and two criterion variables for each of the within-grade analytic subsamples. While PSSA scaled scored for both math and ELA remain consistent across the subsamples, there are several notable differences related to the two predictor variables. In kindergarten, the mean number of excused absences is about 1.5 times greater than the mean number of unexcused days (6.93 days vs. 4.53 days). In first grade, however, the mean number of excused days actually decreases by almost a full day, while the mean number of unexcused days rises by almost two full days, such that the mean number of excused and unexcused days in first grade are roughly the same (6.02 vs. 6.22). Finally, in second grade, the mean number of excused absence days drops again by about a half day, and the

mean number of unexcused days increases by about three-fourths of a day. Thus, in

second grade, students have a higher number of unexcused absences than excused

absences by about 1.5 days on average (7.02 days vs. 5.53). Additionally, the spread of

absence days—as measured by the standard deviation—also changes across the

subsamples. The initial standard deviation is roughly the same for excused and unexcused

days in kindergarten (7.43 vs. 7.23); however, the standard deviation decreases across

both years for excused absence days and increases across both years for unexcused days.

For the second-grade sample, the standard deviation of excused days is more constricted

than that of unexcused days (6.39 vs. 9), meaning that there is less variability in excused

absence days as opposed to unexcused absence days. The descriptive information from

Table 5. Descriptive Statistics for Kindergarten, First Grade, and Second Grade Analytic
Samples

	Analytic Sample for Kindergarten ($n = 6,800$)	Final Analytic Sample for First Grade ($n = 7,453$)	Final Analytic Sample for Second Grade ($n = 7,254$)
	Mean (SD)	Mean (SD)	Mean (SD)
Excused Days	6.93 (7.43)	6.02 (6.56)	5.53 (6.39)
Unexcused Days	4.53 (7.23)	6.22 (8.59)	7.02 (9.00)
Math PSSA Scaled Score	929.56 (103.82)	926.29 (102.62)	925.82 (102.49)
ELA PSSA Scaled Score	965.10 (97.74)	962.65 (96.73)	962.18 (96.65)

these study samples is of note considering the widespread notion that young students

generally do not accumulate many unexcused absences (Balfanz, 2016). On the contrary,

the youngest students in Philadelphia average between 4.5 and 7 unexcused absences per

year, with a higher number of average unexcused days than average excused days in first

grade and second grade.

The descriptive information presented in Table 5 is similar to the data presented in Table 6 for the fourth analytic subsample. Rather than within-year descriptive information, Table 6 shows the mean and standard deviation of all the predictor and criterion variables at each timepoint for the longitudinal sample of students. While the mean number of unexcused absence days for each grade is slightly different in this subsample compared to the other three subsamples (e.g., 6.22 unexcused days in the first-grade subsample vs. 5.5 unexcused days in the across-grade sample), the number of excused absence days for each grade is comparable. Additionally, the data for this subsample show similar differences between excused and unexcused absences compared

Table 6. Descriptive Statistics for Analytic Sample for Analyses across Grade

	Final Analytic Sample for Analyses across Grade (n = 6,223)
	Mean (SD)
Kindergarten Excused Days	6.87 (7.29)
Kindergarten Unexcused Days	4.34 (6.75)
First Grade Excused Days	6.01 (6.51)
First Grade Unexcused Days	5.50 (7.49)
Second Grade Excused Days	5.48 (6.25)
Second Grade Unexcused Days	6.46 (8.46)
Math PSSA Scaled Score	929.63 (103.37)
ELA PSSA Scaled Score	965.62 (97.10)

to the other three subsamples. As with the kindergarten subsample, kindergarteners average more excused days than unexcused days in the across-time sample (6.87 days vs. 4.34 days). This finding changes across years, however, with excused days decreasing by

about a full day and unexcused days increasing by a full day in first grade (6.01 days vs. 5.5 days), and excused days dropping again by about a half day and unexcused days increasing by almost a full day in second grade (5.48 days vs. 6.46 days). The standard deviation for excused days in this subsample also decreases across time, while the variability for unexcused days increases across grades. Thus, for all subsamples, it appears that the average number of excused absence days decreases across grades and unexcused absences increases across grades.

Table 7 presents descriptive information about the criterion variables for each analytic subsample. Instead of a scaled score, this information shows the percentages of students scoring within each performance level of the PSSA (i.e., advanced, proficient, basic, and below basic). This information is particularly relevant as the ROC analyses

Table 7. Percentages of Students Scoring in Each PSSA Level for All Analytic Samples

	Analytic Sample for Kindergarten (*n* = 6,800)		Final Analytic Sample for First Grade (*n* = 7,453)		Final Analytic Sample for Second Grade (*n* = 7,254)		Final Analytic Sample for Analyses across Grades (*n* = 6,223)	
PSSA Level	Math	ELA	Math	ELA	Math	ELA	Math	ELA
Advanced	6.25%	4.04%	5.90%	3.70%	5.86%	3.68%	6.15%	4.00%
Proficient	14.62%	32.04%	14.02%	31.44%	13.80%	31.27%	14.61%	32.19%
Basic	22.40%	34.62%	21.95%	34.91%	22.07%	34.92%	22.64%	34.90%
Below Basic	56.74%	29.29%	58.12%	29.95%	58.27%	30.14%	56.60%	28.91%

will be conducted using below-basic performance for each subtest as a binary indicator. Just as the mean scaled scores for both math and ELA are similar across subsamples, the percentages of students scoring within each performance level is comparable across each subsample: about 6% scored in the advanced range, 14% proficient, 22% basic, and 57%

below basic in math, and about 4% scored in the advanced range, 32% proficient, 35% basic, and 30% below basic in ELA. While these percentages are consistent across subsamples, the difference between these percentages and the percentage of students scoring at each performance level across the state is notable. For the test administration year represented in the data (i.e., 2015), the percentage of third-grade students in each performance range across Pennsylvania was 20% advanced, 29% proficient, 24% basic, and 28% below basic in math and 13% advanced, 49% proficient, 25% basic, and 13% below basic in ELA (Pennsylvania School Board Association, 2015). In the School District of Philadelphia, a much higher percentage of students score below basic in both math and ELA as compared to the rest of the state (i.e., 57% vs. 28% below basic in math and 30% vs. 13% below basic in ELA). This finding comports with national research that shows that large urban school districts tend to have lower standardized test score performance than other districts (Logan, Minca, & Adar, 2012).

The Differential Diagnostic Accuracy between Absence Types within Grades

Research Question 1a: The Diagnostic Accuracy of Excused vs. Unexcused Absences and PSSA Outcomes in Kindergarten

Both excused and unexcused absences in kindergarten were tested for their accuracy in classifying students as being in the below-basic performance level for both the math and ELA subtests of the PSSA. For each pair of predictor and criterion variables, ROC curves and AUC statistics were produced and are reported below. Figure 8 presents the ROC curves for kindergarten excused and unexcused absence days classifying students as below basic on third-grade PSSA math. The ROC curve for

excused absence days shows a fairly flat line that falls closely along the line indicating

random chance (y = x). In contrast, the ROC curve for kindergarten unexcused absence
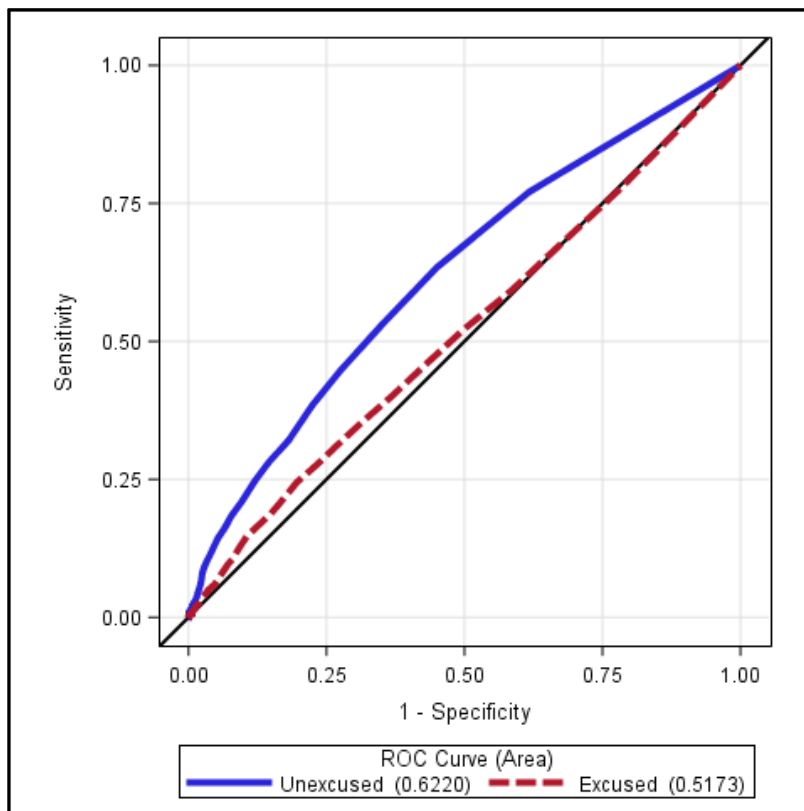
Figure 8. ROC Curves Comparing Kindergarten Excused and Unexcused Absence Days
and PSSA Math Below Basic

days shows an arc toward the upper left corner of the graph (i.e., away from the line of

random chance), indicating better classification accuracy. The ROC curves comparing

kindergarten excused and unexcused absence days and below-basic status in ELA (Figure

9) show similar differences in classification accuracy. The ROC curve for excused

absence days lays roughly flat against the line of random chance and appears to be

slightly flatter than the curve for excused absence days and math. Similarly, the curve for
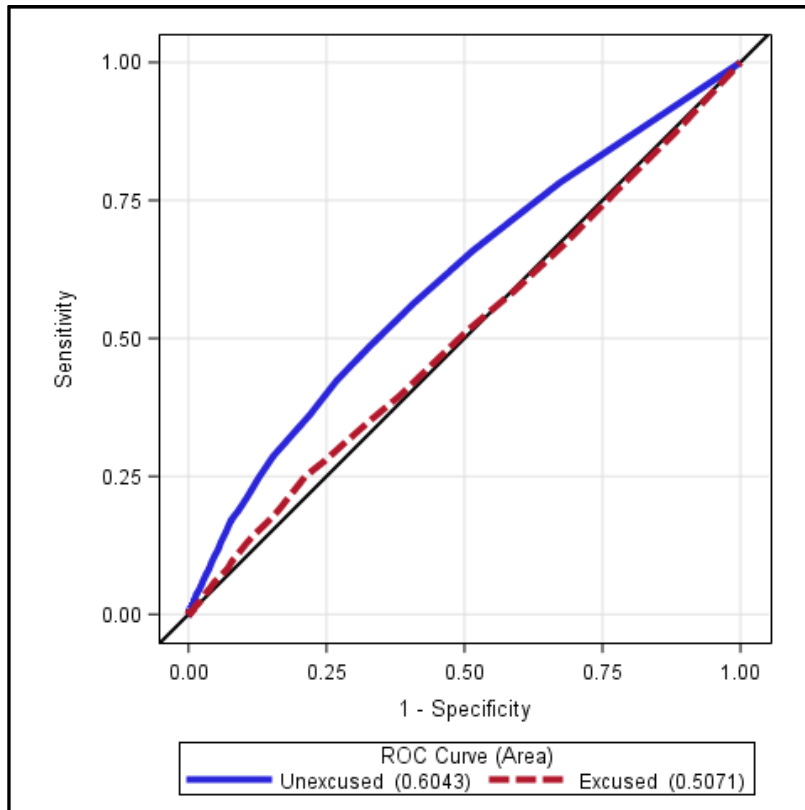
Figure 9. ROC Curves Comparing Kindergarten Excused and Unexcused Absence Days
and PSSA ELA Below Basic

unexcused absence days and ELA shows an arc toward the upper left corner of the graph,

while still appearing slightly flatter than the curve for unexcused absence days and Math.

From these graphs, it appears that unexcused absences in kindergarten have better

diagnostic accuracy than excused absences for both Math and ELA.

The AUC statistics accompanying these graphs are presented in Table 8. As the

table shows, the AUCs for kindergarten excused days for both math and ELA are close to

random chance (0.52 and 0.51, respectively); however, the AUC for excused days and

math is marginally statistically significant (at the .05 level). The AUC for ELA is not

statistically significant. Both of these AUCs are considered negligible in terms of

Table 8. AUC Statistics for Kindergarten Excused and Unexcused Days and PSSA Outcomes

	AUC	SE	95% Confidence Interval		p	Cohen's d
			Lower	Upper		
PSSA Math Below Basic						
Kindergarten Excused Days	0.52*	0.01	0.50	0.53	0.014	0.07
Kindergarten Unexcused Days	0.62****	0.01	0.61	0.64	<0.0001	0.44
PSSA ELA Below Basic						
Kindergarten Excused Days	0.51	0.01	0.49	0.52	0.36	0.04
Kindergarten Unexcused Days	0.60****	0.01	0.59	0.62	<0.0001	0.37

Note. *p < .05. **p < .01. ***p < .001; ****p < .0001.

magnitude based on the previously cited heuristic (Rice & Harris, 2005) and correspond to a Cohen's d effect size of 0.07 and 0.04, respectively. The AUCs for unexcused absences, on the other hand, are statistically significant at the .0001 level for both math and ELA (0.62 and 0.60, respectively). Both of these AUCs are the .0001 level for both math and ELA (0.62 and 0.60, respectively). Both of these AUCs are considered small, corresponding to effect sizes of 0.44 and 0.37, respectively, but are stronger in terms of magnitude than the AUCs for excused absences.

Because the AUC for kindergarten excused absence days and PSSA math was marginally statistically significant, the DeLong test was performed to determine whether this difference in diagnostic accuracy between absence types is statistically meaningful. Table 9 presents the findings of the DeLong test for kindergarten excused and unexcused absence days and math.

As Table 9 indicates, the results of the DeLong test are statistically significant at

Table 9. DeLong Test for AUCs of Excused and Unexcused Days and PSSA Math

	Difference in AUC	SE	95% Confidence Interval		p
			Lower	Upper	
PSSA Math Below Basic					
Kindergarten Excused vs. Unexcused Days	0.11****	0.01	0.09	0.12	<0.0001

*Note. *p < .05. **p < .01. ***p < .001; ****p < .0001.*

the .0001 level. This means that unexcused absence days have significantly higher

classification accuracy for PSSA math than excused absence days. The DeLong test was

not performed for ELA, as excused days were not significantly different than random

chance. Thus, for both math and ELA, unexcused absences in kindergarten show better

diagnostic accuracy than excused absences.

Research Question 1b: The Diagnostic Accuracy of Excused vs. Unexcused Absences and PSSA Outcomes in First Grade

Similar to the previous research question, excused and unexcused absence days in

first grade were tested for their accuracy in classifying students as being in the below-

basic performance level for PSSA math and ELA. ROC curves and AUC statistics for

each pair of predictor and criterion variables are presented below. Figure 10 displays the

ROC curves for excused and unexcused absence days and PSSA math. Similar to the

ROC curves generated for the kindergarten subsample, the ROC curve for excused days

is roughly a flat line, falling almost directly on the line of random chance; the curve for

unexcused absence days, in contrast, presents more of an arc, indicating better
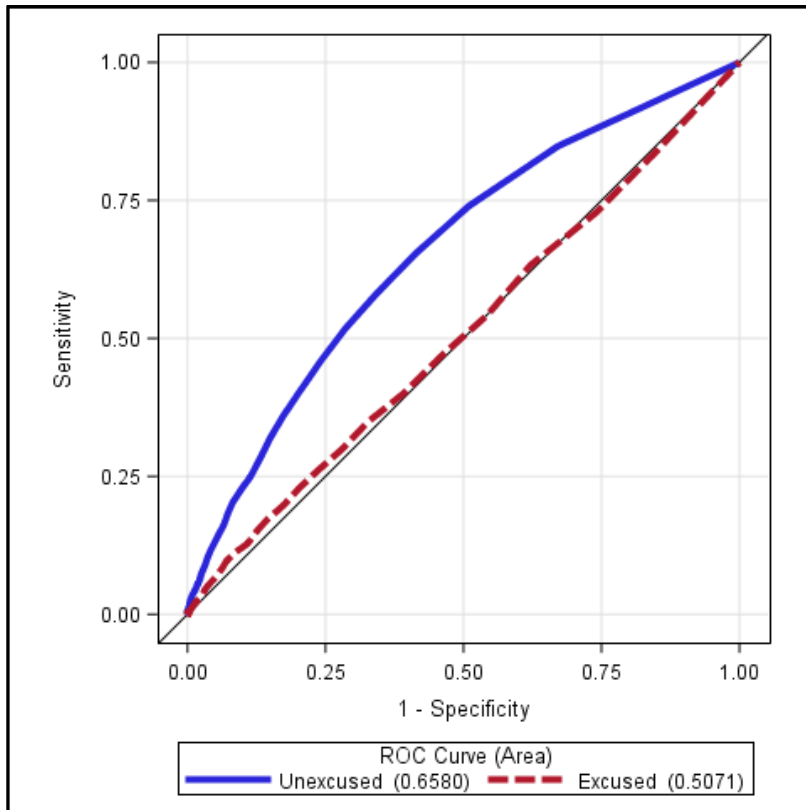
Figure 10. ROC Curves Comparing First Grade Excused and Unexcused Absence Days and PSSA Math Below Basic

classification accuracy. Figure 11 presents the ROC curves comparing first grade excused

and unexcused absence days and below-basic status in ELA. These curves show similar

differences in classification accuracy, while being slightly flatter than the curves for

math. This also mirrors the results presented for the kindergarten subsample. Again, the

curve representing excused absence days appears close to the line of random chance,

even dipping below the line as sensitivity and false positive rate increase, while the line

for unexcused absence days shows more of an arc and signals better diagnostic accuracy.
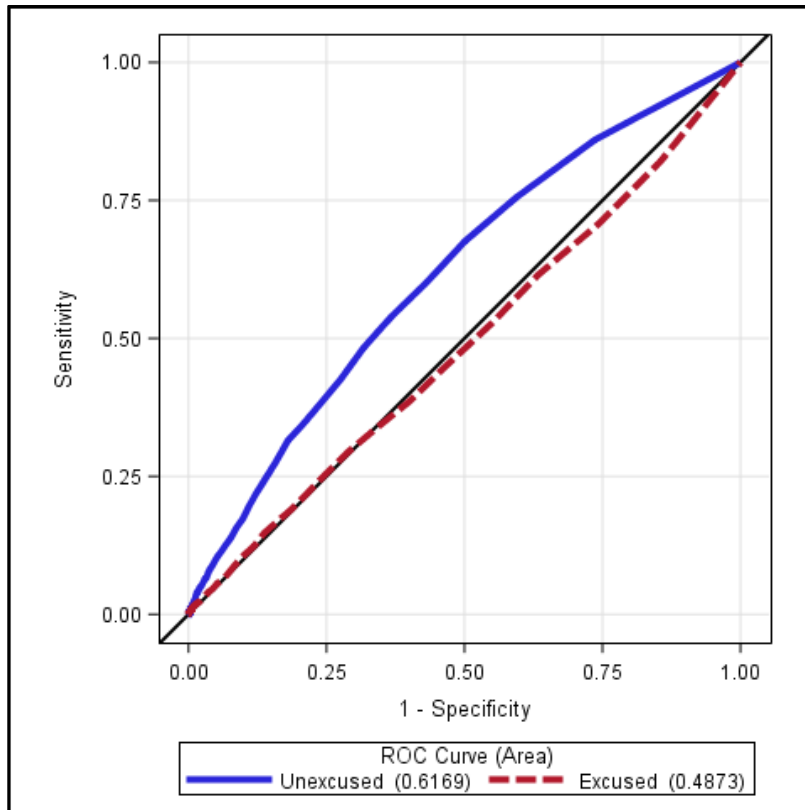
Figure 11. ROC Curves Comparing First Grade Excused and Unexcused Absence Days and PSSA ELA Below Basic

The AUC statistics for these graphs are presented in Table 10. For both math and ELA, the AUCs for excused absences (0.51 and 0.49, respectively) are not statistically significant, meaning they are no different than random chance, while the AUCs for unexcused absences (0.66 and 0.62, respectively) are statistically significant at the .0001 level. In addition, both of the AUCs for excused absences would be considered negligible in terms of magnitude with effect sizes ranging from 0 to 0.04, while the AUC for unexcused absences and math would be considered medium (effect size of 0.59) and the AUC for ELA would be considered small (effect size of 0.44). The DeLong test was not performed, as excused absence days were not significantly different than random chance

Table 10. AUC Statistics for First Grade Excused and Unexcused Days and PSSA Outcomes

	AUC	SE	95% Confidence Interval		*p*	Cohen's *d*
			Lower	Upper		
PSSA Math Below Basic						
First Grade Excused Days	0.51	0.01	0.49	0.52	0.29	0.04
First Grade Unexcused Days	0.66****	0.01	0.65	0.67	<0.0001	0.59
PSSA ELA Below Basic						
First Grade Excused Days	0.49	0.01	0.47	0.50	0.09	0
First Grade Unexcused Days	0.62****	0.01	0.60	0.63	<0.0001	0.44

*Note. *p* < .05. **p* < .01. ***p* < .001; ****p* < .0001.*

for either criterion variable. The ROC graphs, statistical significance of the AUCs, and magnitude of the AUCs all indicate that unexcused absence days in first grade have better diagnostic accuracy than excused absence days for below- basic status in math and ELA. These findings are consistent with those presented for the kindergarten subsample.

Research Question 1c: The Diagnostic Accuracy of Excused vs. Unexcused Absences and PSSA Outcomes in Second Grade

The third research question investigated the difference in diagnostic accuracy of absence type for the final within-grade analytic subsample. As with the kindergarten and first-grade subsamples, excused and unexcused absence days in second grade were tested for their accuracy in classifying students as being in the below-basic performance level for third-grade PSSA math and ELA. ROC curves and AUC statistics for each pair of predictor and outcome variables are presented below. Figure 12 displays the ROC curves for excused and unexcused absence days and PSSA math. Again, these ROC curves
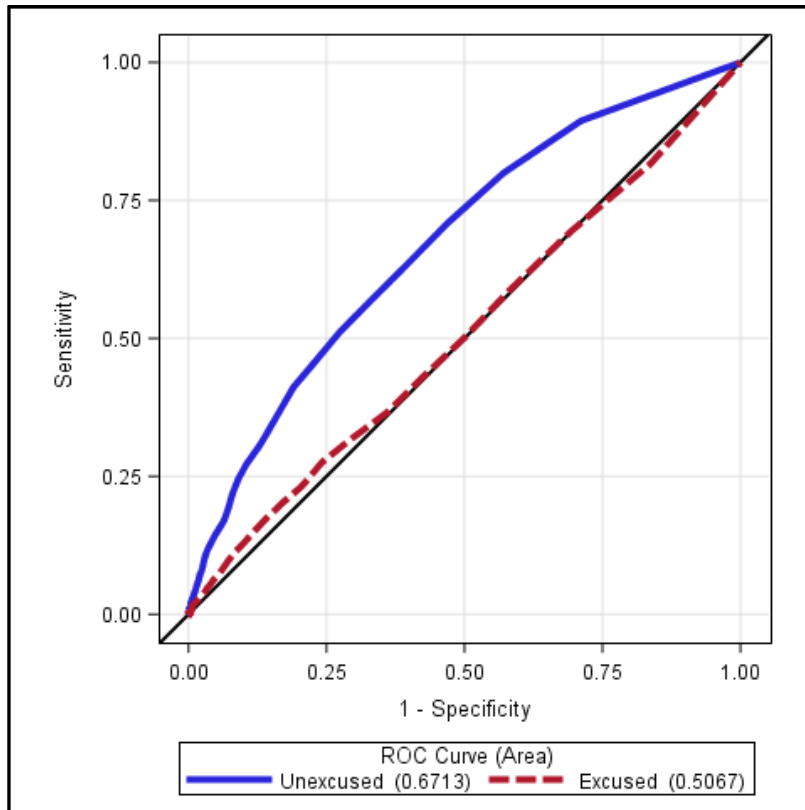
Figure 12. ROC Curves Comparing Second Grade Excused and Unexcused Absence Days and PSSA Math Below Basic

appear similar to those generated for the kindergarten and first-grade subsamples. The

ROC curve for excused days appears flat and falls almost directly on the line of random

chance, while the curve for unexcused absence days arcs toward the upper left corner of

the graph, indicating better classification accuracy. Figure 13 displays the ROC curves

comparing second grade excused and unexcused absence days and below-basic status in

ELA. These curves also show similar differences in classification accuracy, while again

being slightly flatter than the curves for math. The curve for excused absence days again

falls along the line of random chance, and the curve for unexcused absence days arcs

upward to the left. Again, these curves look similar to the curves produced for the
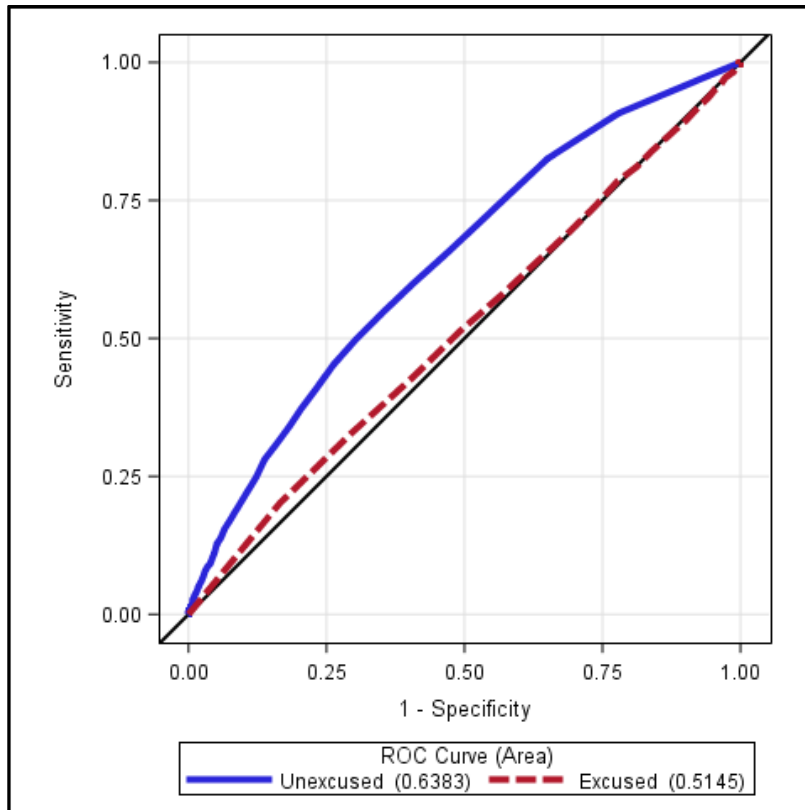
Figure 13. ROC Curves Comparing Second Grade Excused and Unexcused Absence Days and PSSA ELA Below Basic

kindergarten and first-grade subsamples, indicating that there is consistency in the

differential classification accuracy of absence types across the three analytic samples,

with unexcused absences being more accurate classifiers of achievement status than

excused absences.

The AUC statistics that accompany these graphs are presented in Table 11. For

the math and ELA subtests, the AUCs for excused absences (0.51 and 0.52, respectively)

are not statistically significant, meaning they are no different than random chance, while

the AUCs for unexcused absences (0.66 and 0.62, respectively) are statistically

significant at the .0001 level. In addition, both of the AUCs for excused absences would

Table 11. AUC Statistics for Second Grade Excused and Unexcused Days and PSSA Outcomes

	AUC	SE	95% Confidence Interval		*p*	Cohen's *d*
			Lower	Upper		
PSSA Math Below Basic						
Second Grade Excused Days	0.51	0.01	0.49	0.52	0.32	0.04
Second Grade Unexcused Days	0.67****	0.01	0.66	0.68	<0.0001	0.62
PSSA ELA Below Basic						
Second Grade Excused Days	0.52	0.01	0.50	0.53	0.05	0.07
Second Grade Unexcused Days	0.64****	0.01	0.63	0.65	<0.0001	0.51

Note. *p < .05. **p < .01. ***p < .001; ****p < .0001.

be considered negligible in terms of magnitude with effect sizes of 0.04 and .07, respectively; the AUCs for unexcused absences for both math and ELA, in contrast, are medium in magnitude with effect sizes of 0.62 and 0.51, respectively. Similar to the first-grade subsample, the DeLong test was not performed, as excused absence days were not significantly different than random chance for either criterion variable. The ROC graphs, statistical significance of the AUCs, and magnitude of the AUCs indicate that unexcused absence days in second grade have better diagnostic accuracy than excused absence days for below-basic status in both math and ELA. These findings comport with those presented for the kindergarten and first-grade subsamples.

The Stability of Diagnostic Accuracy within Absence Types across Grades

Research Question 2a: Consistency in the Degree of Diagnostic Accuracy of Excused Absences across Grades

The findings for the fourth research question address whether the classification accuracy of excused absences is consistent in magnitude across grade levels. ROC curves and AUC statistics were produced for excused absences related to below-basic

performance in math and ELA across each grade for the fourth analytic subsample.

Figure 14 shows the ROC curves for kindergarten, first grade, and second grade excused
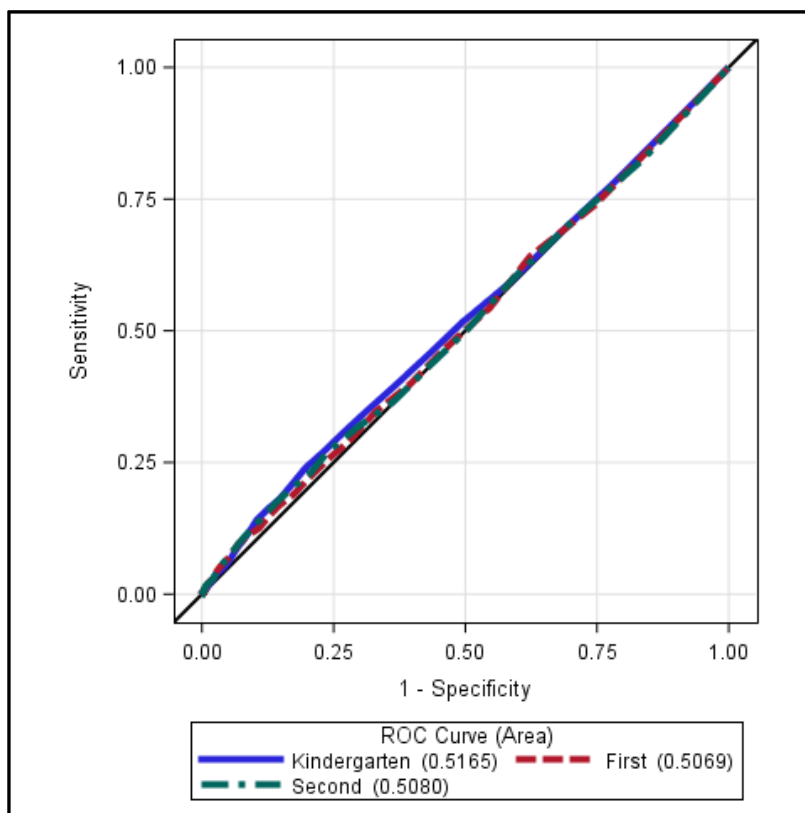
Figure 14. ROC Curves Comparing Kindergarten, First Grade, and Second Grade Excused Absence Days and PSSA Math Below Basic

absence days and below-basic status in math. As the graph indicates, the ROC curve for

each year appears flat and follows the line of random chance. As the lines significantly

overlap, it is difficult to determine any distinctions between them visually, however they

all indicate low diagnostic accuracy. Similarly, Figure 15 presents the ROC curves for

kindergarten, first grade, and second grade excused absence days and below-basic status
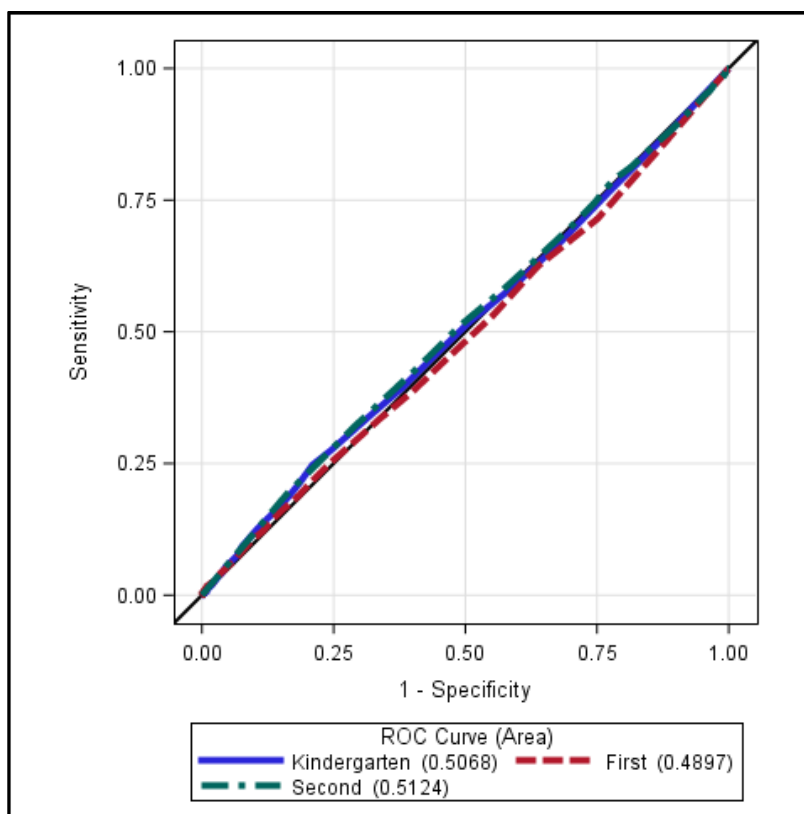
Figure 15. ROC Curves Comparing Kindergarten, First Grade, and Second Grade Excused Absence Days and PSSA ELA Below Basic

in ELA for the fourth analytic subsample. Again, the curves appear flat and clumped together around the line indicating random chance. The curve for first grade excused days is more visually distinguishable from the other two curves, falling slightly below the level of random chance as sensitivity and false positive rate increase. As with the previous graph, these curves indicate poor diagnostic accuracy.

The AUC statistics that accompany these graphs are presented in Table 12. The AUCs for excused absences across all years for both math and ELA below-basic status are stable across grades, with a range of AUC estimates from 0.49 to 0.52. Of these results, only the AUC for kindergarten excused days and math was marginally significant

Table 12. AUC Statistics for Excused Days across Grades and PSSA Outcomes

	AUC	SE	95% Confidence Interval Lower	Upper	p	Cohen's d
PSSA Math Below Basic						
Kindergarten Excused Days	0.52*	0.01	0.50	0.53	0.02	0.07
First Grade Excused Days	0.51	0.01	0.49	0.52	0.35	.04
Second Grade Excused Days	0.51	0.01	0.49	0.52	0.28	.04
PSSA ELA Below Basic						
Kindergarten Excused Days	0.51	0.01	0.49	0.52	0.41	0.04
First Grade Excused Days	0.49	0.01	0.47	0.51	0.21	0
Second Grade Excused Days	0.51	0.01	0.50	0.53	0.13	0.04

Note. *$p < .05$. **$p < .01$. ***$p < .001$; ****$p < .0001$.

at the .05 level, indicating that it was statistically different from random chance. At 0.52, however, the magnitude of this AUC would still be considered negligible in terms of classification accuracy according to Rice and Harris (2005). None of the other AUCs were significantly different from random chance, and all are negligible in magnitude based on the aforementioned heuristic with Cohen's d effect sizes ranging from 0 to 0.07. The DeLong test was not performed, as only kindergarten absence days for math reached statistical significance. These results indicate that the classification accuracy of excused absences remains relatively stable across grades, with kindergarten excused absences being slightly more accurate than first and second grade excused absences for math but still resulting in weak diagnostic accuracy. Thus, excused absences have no meaningful classification accuracy for below-basic status across the early elementary grades.

Research Question 2b: Consistency in the Degree of Diagnostic Accuracy of Unexcused Absences across Grades

The findings for the final research question address whether the classification accuracy of unexcused absences is consistent in magnitude across grades. ROC curves and AUC statistics were produced for excused absences related to below-basic status in math and ELA across each grade for the fourth subsample. Figure 16 presents ROC curves for kindergarten, first grade, and second grade unexcused absence days and below-basic status in math. In contrast with the previous graphs for excused absences,

Figure 16. ROC Curves Comparing Kindergarten, First Grade, and Second Grade Unexcused Absence Days and PSSA Math Below Basic

these curves arc above the line of random chance toward the upper left corner of the graph. Additionally, these lines appear visually distinct, with kindergarten unexcused

days closest to the line of random chance, first grade above kindergarten, and finally second grade farthest from the line of random chance. From this graph, it appears that the diagnostic accuracy of unexcused absences increases across grades.

Finally, Figure 17 presents the ROC curves for kindergarten, first grade, and second grade unexcused absence days and below-basic status in ELA. Consistent with the findings for math, the ROC curves for ELA show a similar visual pattern across grades. Arcing away from the line of random chance, kindergarten unexcused days appear closest

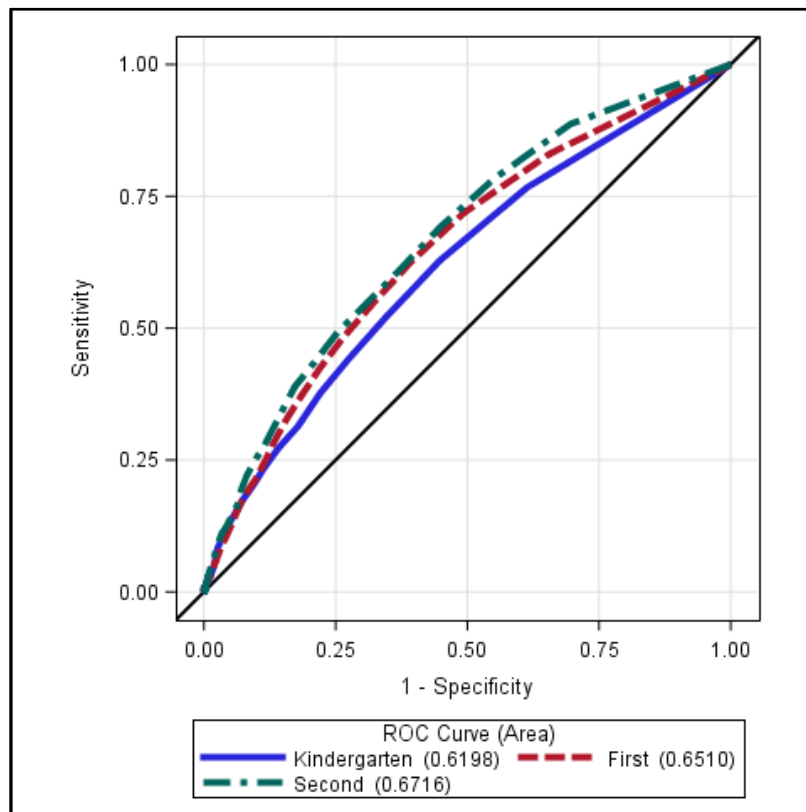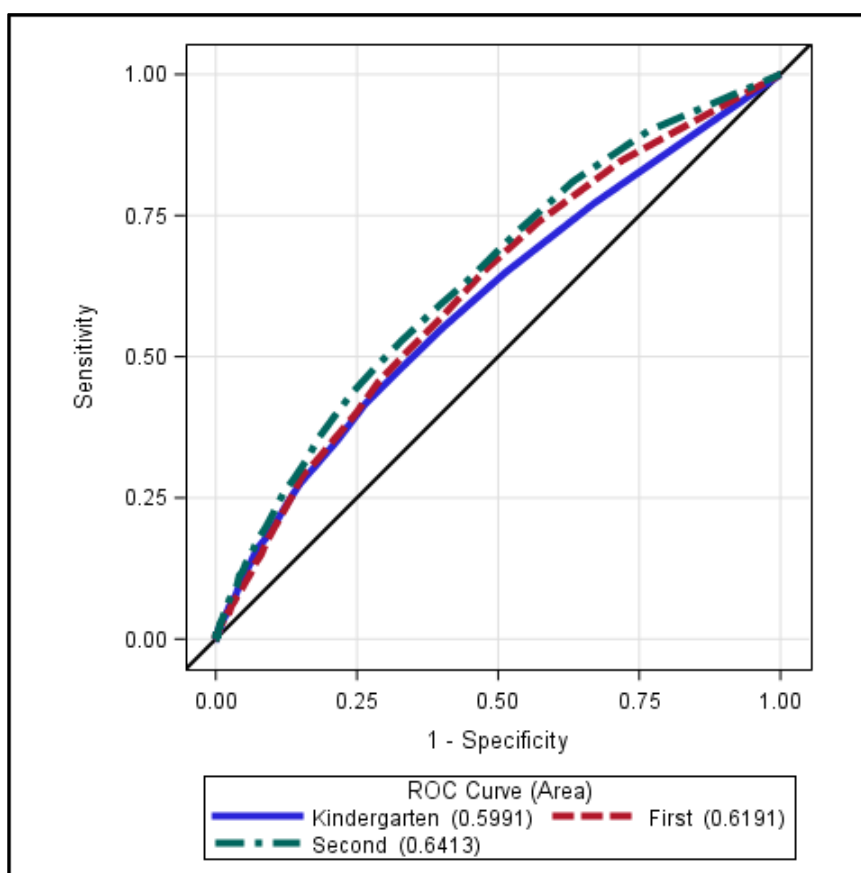**Figure 17.** ROC Curves Comparing Kindergarten, First Grade, and Second Grade Unexcused Absence Days and PSSA ELA Below Basic

to the line, followed by first grade, and finally second-grade unexcused days are farthest from the line of random chance. Again, this appears to indicate that the classification of

accuracy of unexcused absence days increases across grades. While these curves are

slightly flatter than the curves for math, they still indicate that the diagnostic accuracy of

unexcused absence days increases across grades for ELA.

Table 13 presents the AUC statistics for unexcused absences across all years for

both math and ELA below-basic status. All AUCs are statistically significant at the .0001

level. For math, the AUC for kindergarten is small in magnitude but becomes medium in

Table 13. AUC Statistics for Unexcused Absence Days across Grades and PSSA
Outcomes

| | AUC | SE | 95% Confidence Interval | | p | Cohen's d |
			Lower	Upper		
PSSA Math Below Basic						
Kindergarten Unexcused Days	0.62****	0.01	0.61	0.63	<0.0001	0.44
First Grade Unexcused Days	0.65****	0.01	0.64	0.67	<0.0001	0.55
Second Grade Unexcused Days	0.67****	0.01	0.66	0.69	<0.0001	0.62
PSSA ELA Below Basic						
Kindergarten Unexcused Days	0.60****	0.01	0.58	0.62	<0.0001	0.37
First Grade Unexcused Days	0.62****	0.01	0.60	0.63	<0.0001	0.44
Second Grade Unexcused Days	0.64****	0.01	0.63	0.66	<0.0001	0.51

Note. *p < .05. **p < .01. ***p < .001; ****p < .0001.

magnitude for first and second grade with effect sizes ranging from 0.37 to 0.62.

Additionally, the AUCs for kindergarten and first grade unexcused days and ELA are

small but become medium in magnitude by the second grade. These findings are

consistent with the ROC graphs, showing better diagnostic accuracy on both criterion

variables across time.

Because all results were statistically significant, DeLong tests comparing the

AUC of each grade level were necessary to determine if the differences in AUCs were

statistically meaningful. Table 14 shows the results of the DeLong tests for unexcused

Table 14. DeLong Test for AUCs of Unexcused Days across Grades and PSSA Outcomes

	Difference in AUC	SE	95% Confidence Interval		p
			Lower	Upper	
PSSA Math Below Basic					
Kindergarten Unexcused vs. First Unexcused Days	0.03****	0.01	0.04	0.07	<0.0001
Kindergarten Unexcused vs. Second Unexcused Days	0.05****	0.01	0.01	0.02	<0.0001
First Unexcused vs. Second Unexcused Days	0.02***	0.01	0.01	0.03	0.0006
PSSA ELA Below Basic					
Kindergarten Unexcused vs. First Unexcused Days	0.02**	0.01	0.01	0.04	0.0085
Kindergarten Unexcused vs. Second Unexcused Days	0.04****	0.01	0.03	0.06	<0.0001
First Unexcused vs. Second Unexcused Days	0.02****	0.01	0.01	0.04	0.0007

Note. $*p < .05.$ $**p < .01.$ $***p < .001;$ $****p < .0001.$

absences across grades for both math and ELA. As the table indicates, all tests were

significant at the .01 level. These results suggest that the improvement in diagnostic

accuracy across grade levels is statistically meaningful; thus, first grade unexcused

absences are a better classifier of achievement risk status than kindergarten unexcused

absences, and second grade unexcused absences are a better classifier of achievement risk

status than first grade unexcused absences, with the largest difference estimates between the AUCs for kindergarten and second grade (0.05 for math and 0.04 for ELA). These results indicate that, unlike the diagnostic accuracy of excused absence days, unexcused absence days become more accurate classifiers of below-basic status in math and ELA as children progress through school. The findings for the fourth and fifth research questions are thus consistent with the previous findings—namely, that unexcused absence days have diagnostic value for achievement risk status and that this diagnostic value increases across time, while excused absence days do not provide any diagnostic utility in classifying students as being at risk for academic achievement problems and this low accuracy is stable over time.

CHAPTER 4: DISCUSSION

With the passing of the Every Study Succeeds Act (2015), chronic absenteeism has emerged as a national policy indicator to address persistent attendance and achievement gaps in the U.S. Increasingly, states and school districts are monitoring and reporting chronic absenteeism among their students. In addition, chronic absenteeism is being implemented in schools across the nation as an early warning indicator to identify students at risk for later academic achievement problems. Fundamentally, chronic absenteeism is based on the core theoretical assumption that both excused and unexcused absences are diagnostically equivalent in their ability to identify at-risk students. Despite the prevalent use of this early warning indicator, there have been *no empirical tests of this key theoretical assumption.* Because diagnostic accuracy in predicting future risk status is the primary aim of early warning indicators like chronic absenteeism, the lack of any rigorous empirical evidence to test this fundamental assumption is troubling. The absence of research in this area has even more problematic implications for students in early elementary school, as the preventative focus of early warning indicators is meant to identify and provide supports to the youngest students in the public education system.

The present study is thus the first to empirically test the relative diagnostic accuracy of excused and unexcused in determining future academic risk status for students within and across the early elementary grades. To achieve this aim, the current study consisted of two primary research foci: (1) to determine whether there was differential diagnostic accuracy between excused and unexcused absences for below-basic performance in third-grade math and ELA within kindergarten, first grade, and

second grade; and (2) to determine whether the diagnostic accuracy of each absence type remained consistent in magnitude for below-basic status in third-grade math and ELA longitudinally across kindergarten, first grade, and second grade. These research objectives address whether differential classification accuracy between absence types exists at any time in the early elementary grades and whether the classification accuracy of absence types changes across these early grades. Both of these considerations are critical for determining whether excused and unexcused absences can be used to classify students within and across the early elementary grades as being at risk for negative academic outcomes in the future and, thus, whether chronic absenteeism is being appropriately operationalized and can be effectively used as an early warning indicator, particularly for the youngest students in public school.

The following sections of this chapter discuss the significance of the analytic findings of this study in the context of the existing research literature. Limitations of the current study, which qualify its contribution to the literature, will then be presented and discussed. The final section will review the implications of this study on future research and will consider its importance in relation to critical educational policies and practices.

Is There Differential Classification Accuracy between Excused and Unexcused Absences within the Early Primary Grades?

To address the first research objective, this study examined the differential classification accuracy of absence types on academic achievement risk status within the early primary grades. Findings from the first three research questions demonstrated that *there is differential classification accuracy between excused and unexcused absences* in kindergarten, first grade, and second grade. At all three grade levels, *excused absences*

provided no discernible diagnostic accuracy for classifying students as being in the below-basic category for both math and ELA. *Unexcused absences, on the other hand, provided statistically significant levels of diagnostic accuracy within all three grade levels.* The diagnostic accuracy for unexcused absences ranged from small to medium in terms of the magnitude of the effect size (Cohen's *d* ranging from 0.37 to 0.62). These small-to-medium effect sizes are particularly noteworthy, as the average effect size of an intensive intervention directly aimed at improving academic achievement has been found to be in the small-to-medium range (0.20 to 0.51) (Hill, Bloom, Black, & Lipsey, 2008). The magnitude of the diagnostic accuracy of unexcused absences is thus equivalent to the average strength of the effect for an intensive educational intervention. In contrast, the effect sizes of the diagnostic accuracy for excused absences were negligible, ranging from 0 to 0.07. These findings thus indicate that *unexcused absences in the early primary years provide diagnostic utility for determining future academic risk status, while excused absences in the early primary grades have no diagnostic utility in determining future academic risk status.*

The findings of this study are significant in that they *provide empirical evidence to challenge the theory upon which chronic absenteeism is based*. Chronic absenteeism accounts for both excused and unexcused absences based on the notion that any absence represents lost time in school, which is detrimental to student learning and thus results in lower academic achievement (Balfanz and Byrnes, 2012; Jordan et al., 2018) This theory contends that the absence itself is distinctly more predictive of academic problems than the cause of the absence (Coelho et al., 2015). By this theory, research comparing the

association between absence types and academic outcomes should find no difference between excused and unexcused absences as they relate to achievement. The findings from the first research objective indicate that this assumption is incorrect. It appears that absence type is an important piece of information above and the mere absence itself. In fact, the nonexistent diagnostic utility of excused absences found in this study indicates that excused absences have little connection to academic outcomes at all. This undermines the theory upon which chronic absenteeism is based and suggests that its effectiveness as an early warning indicator could be improved through the exclusive use of unexcused absences.

This critical finding—that unexcused absences provide better diagnostic accuracy than excused absences in classifying students at risk for future achievement problems—comports with the two studies that tested the differential association between absence type and academic achievement. Gottfried's study (2009) of 90,000 second-, third-, and fourth-grade students showed that there was a differential association between excused and unexcused absences and achievement outcomes. In fact, there was a small, but significant, *positive* effect of the proportion of excused absences to total absences on standardized math and reading achievement; in contrast, there was a small, but significant, negative effect of the proportion of unexcused absences to total absences on math and reading achievement. Gershenson et al. (2017) found similar results in their study of 650,000 third-, fourth-, and fifth-grade students in North Carolina. While both excused and unexcused absences were associated with a negative effect on standardized achievement performance, the negative effect of unexcused absences was two to three

times larger in reading and math than it was for excused absences. The findings from both of these studies indicate that there is a differential association between absence type and achievement outcomes, with unexcused absences relating more strongly to negative outcomes than excused absences. This is consistent with the findings of the current study, which showed that excused absences had no diagnostic utility in predicting academic risk status, while unexcused absences were able to classify students as being at risk for future achievement problems.

Additionally, the current study's findings extend the existing research literature in two important ways. First, this study is the only research to provide empirical evidence of differential classification accuracy in absence types as they relate to achievement. The previous two studies provided evidence of a differential *associational relationship* between absence types and achievement but *do not provide substantiation for the use of either absence type as a diagnostic classifier (i.e., an early warning indicator)*. Because early warning indicators work by classifying students into binary categories (i.e., academically at risk vs. not at risk), all variables being used as classifiers should undergo tests for diagnostic accuracy to determine their utility in differential classification. Utilizing the premier methodology from fields such as biomedicine, psychology, and engineering, this study was the first to employ ROC analysis to determine the degree of differential classification accuracy between absence types and achievement risk status. The research presented here thus represents the only empirical test of the use of absence types as a potential classifier for students at risk for future achievement problems. As

such, these findings have more direct implications on the use of chronic absenteeism as an early warning indicator than previous studies.

The second extension this study provided to the existing literature relates to the degree to which it focuses on the youngest students in primary school. Because early warning indicators provide the benefit of early detection of future problems, it is essential that these indicators are tested for their utility with the youngest students in the public education system. The previous two studies focused on the differential relationship of absence types to achievement outcomes in upper elementary school but did not explore whether this differential association existed in the early elementary grades. Though Gershenson et al. (2017) did perform a complementary study within their paper that assessed the differential association of absence types in the early elementary grades, issues with sample bias and lack of statistical power preclude the inclusion of the findings from this discussion. Thus, this study provides the first assessment of whether there is differential classification accuracy between excused and unexcused absences and achievement outcomes in the earliest primary grades—a consideration that is made even more important by the primary aim of early warning indicators to detect students at risk for future educational issues at the youngest possible age.

The findings from this first research objective are significant in how they relate to the current understanding of excused and unexcused absences in the theoretical and empirical research literature. First, these findings conflict with the prevailing assumption of chronic absenteeism and contradict a theoretical model suggesting that all time in school is important because it exposes children to critical developmental processes

(Bronfenbrenner & Morris, 1996). While this contradiction is seemingly counterintuitive, it suggests that excused and unexcused absences differ in what they actually represent, rather than serving as a meaningless marker of the same event (i.e., that a student has missed a day of school). Rather, *absence type may be an important signal for other critical factors* affecting academic outcomes; that is, absence types may connote important qualitative information about the student and the student's family and community context. Take, for example, two students that have each accumulated 20 absence days in kindergarten; Student A has 20 excused absences and Student B has 20 unexcused absences. Student A missed 20 days due to a particularly bad cold and flu season, but his parents, who are highly engaged with the school, kept in touch with his teacher and the front office to let them know about the situation and provided doctor's notes as required. Student B, on the other hand, missed 20 days due to a variety of reasons—a lack of reliable transportation to school; the need to care for younger siblings on days when childcare was unavailable; parent work schedules. Furthermore, neither parent was particularly engaged with the school and did not know the procedure for calling in to report absences. While Student A and Student B have the same number of total absences, the *meaning* of their absences is different and connotes important information about each of them. For example, Student A may have the familial support and accompanying financial resources to easily make up for this lost time in school (e.g., educational resources, like books, in the home), while Student B may not have the analogous support and resources. The theoretical assumption that all absences are created

equal may thus be more nuanced than the operationalization of chronic absenteeism would indicate.

Additionally, the idea that absence types may serve as a proxy for family engagement with the educational process and/or structural issues related to the family is well supported by the empirical research literature. Both qualitative and quantitative research affirms the notion that students experiencing high numbers of unexcused absences are more likely to come from families that have low involvement in their children's education or from families facing serious challenges (Jeynes, 2003; Klerman & Glasscock, 1996; Lehr et al., 2004; McNeal, 1999; Romero & Lee, 2008; Sheldon, 2007; Teasley, 2004). While some of this evidence comes from the truancy literature, which involves older students, it is still relevant to the current findings. For instance, Maynard et al. (2012) studied a nationally representative dataset of students from middle to high school and found that those with the fewest number of unexcused absences had the highest reported parental involvement in education, while those accumulating the largest number of unexcused absences came from families with the lowest reported parental involvement in their education. Furthermore, Teasley (2004) found that students experiencing a large number of unexcused absences were more likely to come from families experiencing one or more risk factors: living in poverty, coming from a single-parent household, crowded living conditions, irregular parent work schedules, and housing instability. As the number of risks accumulate, the likelihood of the student experiencing more unexcused absences increases. These findings are significant in that they support the contention that unexcused absences may be fundamentally *different* than

excused absences and serve as a signal for other critical factors that can influence achievement. In trying to find the most precise early warning indicator, it is essential that the difference in absence types, and all that is connoted therein, not be masked by grouping all absences together. Rather, the differentiation between these two absence designations should be leveraged as a way to most effectively identify young students at risk.

Is There Consistent Classification Accuracy within Absence Types across the Early Primary Grades?

The second research objective sought to examine whether the classification accuracy of each absence type remained consistent in magnitude across the early grades in the longitudinal sample or if it increased or decreased over time. Findings indicated that *there is a differential longitudinal pattern in the classification accuracy of excused and unexcused absences*. Excused absences showed no classification accuracy in kindergarten, first grade, and second grade for both math and ELA with negligible effect sizes ranging from 0 to 0.07. E*xcused absences thus show a stable pattern of providing no diagnostic accuracy across the early elementary school grades.* In contrast, unexcused absences showed variability across time in the magnitude of classification accuracy for math and ELA. Across kindergarten, first grade, and second grade, unexcused absences provided diagnostic accuracy that increased over time for both criterion variables. Effect sizes ranged from 0.37 to 0.64, with the smaller effect sizes seen in kindergarten and the larger effect sizes seen in second grade. Again, the magnitude of these effect sizes is significant, as the average effect size of an intensive educational intervention ranges from 0.20 to 0.51 (Hill, Bloom, Black, & Lipsey, 2008). Furthermore, the differences in

magnitude of the classification accuracy of unexcused absences were significant for all year-to-year comparisons (e.g., diagnostic accuracy was significantly higher in first grade than in kindergarten and significantly higher in second grade than in first grade). The findings thus showed that *the classificatory accuracy of unexcused absences actually increases over time*, with unexcused absences providing significantly higher diagnostic accuracy as children advance from kindergarten to second grade. These findings suggest that *excused absences have no diagnostic utility across the early elementary grades, while unexcused absences provide increasingly better classification accuracy as children progress through early elementary school.*

Because this is the first study to explore the strength of the classification accuracy of absence types across time, there is no empirical research available that serves as a direct source of comparison. There is, however, a broader research base that provides some insight into the finding that unexcused absences become a stronger classifier of academically at-risk students over time. While there is substantial research evidence that absence patterns emerge early in the educational lives of children and persist over time (Bauer et al., 2018; Ehrlich et al., 2014; Hickman & Heinrich, 2011; Mac Iver, 2010; Neild & Balfanz, 2002), there is additional research suggesting that unexcused absences among the most at-risk students (e.g., those who eventually have persistently failing grades or drop out of school altogether) increase in frequency as children progress through school (Heilbrunn, 2007; Maynard et al., 2013; Schoeneberger, 2012). The notion that the number of unexcused absences may intensify as students advance in school and is increasingly predictive of dire educational outcomes supports the findings

of this study. Unfortunately, there are no comparative studies of longitudinal patterns of excused absences, so it is difficult to determine whether excused absences also seem to increasingly predict educational problems or whether they are stable across time in their lack of predictive utility, as this study found.

Because no studies have investigated the predictive efficacy of different absence types across time, this study represents a significant contribution to the knowledge base. In addition, these findings are significant as they relate to the broader truancy literature. Specifically, the truancy literature has well established that students experiencing high numbers of unexcused absences must be identified and provided with supports as early as possible in order to change their educational trajectories. Efforts to remediate truant students become less successful over time, as attendance patterns and their correlates (e.g., educational disengagement, failing grades, etc.) become more entrenched and students fall further behind their peers (Blazer, 2011; Hickman & Heinrich, 2011; Mac Iver, 2010). By the time students are identified as habitually truant in late-middle and early-high school, it is often too late to provide remediation to change educational pathways. It is, thus, imperative that these students be identified as early as possible and given the resources they need to change unexcused absence patterns and ultimately alter their educational trajectories. This speaks to the primary purpose of early warning indicators—to identify and provide support to those students most at-risk for educational failure *as early as possible*. The findings of this study show that unexcused absences are a useful early warning indicator in kindergarten and become a *progressively stronger* classifier of risk status over the course of the early elementary years. As such, unexcused

absences can be used to identify at-risk students in the early primary grades and provide

them with the supports they need to change their educational trajectories before

behavioral patterns become entrenched and corrective measures become ineffectual.

Limitations of the Current Study

The current study provides a contribution to an understudied area of the school

absence literature and has significant implications for both research and policy, which

will be discussed later in this chapter. The findings presented here represent the first

empirical investigation into the diagnostic accuracy of absence types in the early primary

grades in determining future academic risk status. This study thus signifies an initial step

to raise awareness about the lack of empirical testing of a core underlying assumption of

chronic absenteeism. Due to the exploratory nature of this work, there are two important

limitations of this study that require further discussion. The first limitation relates to the

need to test the generalizability of these findings to other educational contexts. The

second limitation has to do with the aforementioned "gold standard" concern inherent in

ROC analysis and the need for replication of these findings with other criterion variables.

The most pertinent limitation of this study is its lack of corroboration with other

research. As the first study to address the classification accuracy of absence types, it is

difficult to know whether the findings presented here are context specific or are more

broadly applicable. Because chronic absenteeism is being used across the nation as a

policy indicator, it is essential that the methodology presented in this study be replicated

across other contexts. It would be inappropriate to assume that the findings from this

research are nationally generalizable, as data from only one school district were used in

this research. On the other hand, the group of students represented in this study come from an urban school district in one of the largest cities in the country. This makes the findings pertinent to other large urban school districts, such as those categorized as "large cities" by the National Assessment of Educational Progress (NAEP District Profiles, n.d.). The applicability to large urban school districts is also highly policy-relevant, as these districts tend to face persistent challenges related to student attendance and achievement and are thus areas of focus for educational policy reform efforts (Chang & Romero, 2008; Jacob & Lovett, 2017; Lleras, 2008; Sandy & Duncan, 2010; U.S. Department of Education, Office for Civil Rights, 2016). The scale of this study, the strength of its findings, and its applicability to other large urban school districts thus necessitate replication to address this first limitation.

The second limitation relates to an issue that is prevalent in the ROC methodology, even within fields such as biomedicine and psychology: the gold standard problem (Zou et al., 2007). ROC methodology relies solely on the use of a criterion variable to determine the degree of diagnostic accuracy of a predictor variable. This inherently means that the assessment of the predictor variable is only useful if the criterion variable reflects the true status of the condition or classificatory category that it represents (McFall & Treat, 1999). The gold standard problem is ameliorated in the context of this study because the most widely used measure of academic performance— the end-of-the-year state assessment of achievement—serves as the criterion variable, and this assessment undergoes a thorough and well-documented process of development and testing (Data Recognition Corporation, 2015; (Karantonis & Sireci, 2006; Mitzel et al.,

2001; Rodrigues, 2017). Utilizing additional criterion variables (e.g., other rigorously developed and tested measures of academic achievement) to assess the diagnostic accuracy of absence types would further ameliorate the issue of the gold-standard problem and should be a consideration for future research.

Implications for Research

As the first study to investigate the differential classification accuracy of absence types on academic achievement risk status within and across the early primary grades, there are numerous opportunities for future research to continue and advance this important area of inquiry. Because this research is nascent, it is important to consider how the findings presented here can be further verified empirically and how they can be cultivated beyond the current study to push the field into important new areas. A research agenda to further investigate whether absence types can be effectively used in the creation of early warning indicators can thus be conceptualized as having two primary categories—opportunities for replication and opportunities for expansion.

There are several important ways in which this research should be replicated. One self-evident opportunity for replication would be to recreate the study with other cohorts of students in the early primary grades in Philadelphia. The existing study used student data from the 2011-2012 school year through the 2014-2015 school year; conducting the same study with more contemporaneous data would lend additional weight to the current findings within the context of the School District of Philadelphia. Additionally, this study could be replicated with the same population of students using a different criterion measure for standardized academic achievement (e.g., another well-developed and

rigorously tested assessment of academic achievement). Conducting the same study using a similar criterion variable would further ameliorate concerns related to the gold-standard problem discussed in the previous section (McFall & Treat, 1999; Zou et al., 2007).

Beyond the context of the School District of Philadelphia, there are two other avenues for replication that should be considered to test the generalizability of these findings. As one of the largest cities in the U.S. with one of the highest rates of early childhood poverty in the country, Philadelphia represents a unique context in which to study the issue of absences. Educational research has consistently found that attendance and achievement challenges are particularly acute within urban settings and among low-income children (Chang & Romero, 2008; Gonzales et al., 2002; Gottfried, 2015; Lleras, 2008; Ready, 2010; Sandy & Duncan, 2010; Spencer, 2009; Teasley, 2004). Thus, studying the discriminatory accuracy of student absences for academic risk status in other urban settings with high concentrations of children living in poverty is a logical context for this initial inquiry. According to the National Assessment of Educational Progress (NAEP), a national comparison of student achievement among school districts in large cities shows that places like Detroit, Baltimore, and Milwaukee, which have academic achievement outcomes similar to Philadelphia, would be especially pertinent contexts for replication (NAEP District Profiles, n.d.). Additionally, places like New York City and Chicago, which are also classified as large cities and where school attendance challenges are well documented, would be key areas to test the classification accuracy of absence types (Allensworth & Easton, 2007; Allensworth et al., 2014; Balfanz & Byrnes, 2013; Ehrlich et al., 2014; Ou & Reynolds, 2008). Applying the rationale and methodology of

this study to the school districts of cities such as Detroit, Baltimore, Milwaukee, New York City, and Chicago would constitute an essential replication of this work and would provide further insight into whether absence types have differential classification accuracy in the early grades in other large urban settings.

A final consideration for replication relates to non-urban settings (i.e., suburban and rural areas). Though the findings of this study may not be generalizable to non-urban settings, the question of whether absence types provide differential diagnostic accuracy is still essential in places where chronic absenteeism is being used as an early warning indicator or is mandated for use in compliance with state ESSA requirements. Suburban and rural areas using chronic absenteeism as an early warning indicator have no empirical evidence that it is a useful diagnostic tool or whether it could be changed to be made more useful (for instance, by disregarding excused absences in favor of unexcused absences). Until the diagnostic efficacy of both absence types has been established, the importance of reproducing the current study in non-urban settings is also critical. Thus, replication of this study in Philadelphia, other comparable large cities, and in non-urban settings that are using chronic absenteeism as an early warning indicator is a critical next step in the field.

Other opportunities for future research involve expansion beyond the scope of the current study. First, the current study could be expanded upon by using criterion variables outside of academic achievement to determine classification accuracy. While this study only investigated the relationship of absence types to standardized achievement outcomes, research has shown that absences relate to a host of negative outcomes,

including suspensions, failing grades, expulsions, social-emotional issues, dropout, substance abuse, teen pregnancy, and incarceration (Baker et al., 2001; Dryfoos, 1990; Garry, 1996; Gottfried, 2014a; Huizinga, Loeber, Thornberry, & Cothern, 2000; Jarjoura, 1993; Kearney, 2008; Kearney & Graczyk, 2014; Maynard et al., 2012; Mogulescu & Segal, 2002; Mueller et al., 2006; Rumberger & Thomas, 2000; Sheldon & Epstein, 2004; Smink & Heilbrunn, 2006; Spencer, 2009). Studies that investigate the diagnostic accuracy of absence types as they relate to these other important outcomes would be tremendously beneficial in determining the utility of absences as early warning indicators. For instance, because attendance is so strongly related to high school dropout, the diagnostic accuracy of absences in the early primary grades could be considerable for this outcome; if unexcused absences in kindergarten, first grade, and second grade were shown to have high diagnostic accuracy for dropout, they could be incredibly useful as early warning indicators for this detrimental outcome. Because the effects of student absences are not limited to academic considerations, investigating how they relate to a variety of outcomes is essential for future research.

Another area of expansion for this work relates to a more distal but nonetheless crucial avenue for research. While the current study provides evidence to suggest that there are differential meanings and factors associated with excused vs. unexcused absences, there may be similar differential meanings and nuance among types of unexcused absences. Because the literature confirms that unexcused absences are caused by a variety of factors, treating them as a unidimensional construct may be as ill-advised as not discriminating between excused and unexcused absences (Chang & Romero, 2008;

Jeynes, 2003; Klerman & Glasscock, 1996; Lehr et al., 2004; Maynard et al., 2012; McNeal, 1999; Romero & Lee, 2008; Sheldon, 2007; Teasley, 2004). Establishing a taxonomy of unexcused absences and then determining whether these categories of unexcused absences differentially relate to outcomes would be an enormous contribution to the field. For instance, suppose unexcused absence days caused by transportation issues have little classification accuracy, while those related to family disengagement with the school have extremely high classification accuracy. Distinguishing between these two types of unexcused absences would be essential for developing the most effective early warning indicator. Because most schools and school districts do not record information about student absences beyond a binary designation of excused and unexcused, this research would be difficult to conduct on a large scale; it would be better suited for a smaller mixed-methods inquiry that could gather qualitative information about the nature of children's unexcused absences, use this information to create distinguishable categories of unexcused absences, and then perform ROC analysis to determine whether the various categories of unexcused absences relate differentially to particular outcomes. Though this represents a more distal expansion of the current study, it is crucial that the nature of unexcused absences be further explored to determine if additional nuance to unexcused absence data could render it more useful in determining which students are at-risk for later educational issues.

Finally, and perhaps most critically, a research agenda within this field must test the second core assumption upon which chronic absenteeism is based. This study addressed the first underlying assumption of chronic absenteeism—that excused and

unexcused absences are of equal predictive value in determining a student's future risk status; the second underlying assumption of chronic absenteeism is that the 18-day cut-off point to determine chronic absentee status is a meaningful threshold. There is, however, no empirical basis for this threshold (Gershenson et al., 2017; Sutphen et al., 2010). To create a scientifically sound early warning indicator, it is thus necessary to first evaluate the diagnostic accuracy of the predictor making up that indicator and then to identify the appropriate cut score threshold for variables that have demonstrated diagnostic accuracy. As this study accomplished the first aim (i.e., determining the diagnostic utility of the variables that constitute chronic absenteeism), future research should investigate the latter issue. For instance, in the context of the current study, additional analysis may reveal that a threshold of 12 unexcused days is the most useful cut point for identifying students at risk. This would mean that the most effective early warning indicator related to student absences in Philadelphia is 12 unexcused absence days rather than 18 total absence days. As discussed in the methods chapter, ROC analysis allows for an empirically sound way to calculate meaningful cut points for binary indicators (Youngstrom, 2014) and should be used by future research efforts to determine a valid threshold for chronic absenteeism. In order to create the most scientifically sound system of early warning indicators, it is thus necessary to further the current study by determining an optimal cut-score threshold for the variables that constitute chronic absenteeism using an empirically rigorous process.

Implications for Policy and Practice

In response to federal policy mandates, about 75% percent of states have adopted chronic absenteeism as a policy indicator of school quality and student success under ESSA (Bauer et al., 2018; Jordan & Miller, 2017, Jordan et al., 2018). Moreover, chronic absenteeism is being used in practice as an early warning indicator around the country, as schools leverage it as a way to determine student risk status and allocate resources accordingly (Allensworth & Easton, 2007; Allensworth et al., 2014; Balfanz et al., 2007; Bruce, Bridgeland, Fox, & Balfanz, 2011; Heppen & Therriault, 2008; Neild & Balfanz, 2006; West, 2013). As such, it is essential for both policy and practice that chronic absenteeism is based on high-quality science substantiating its predictive utility and efficacy for determining which students are at risk for future educational problems. Unfortunately, prior to this study, no such empirical evidence existed. As the first study to attempt to validate the use of this indicator as a predictor of risk status and to utilize an advanced methodological approach to do so, the research presented here has important implications for both policy and practice.

As the findings of this study clearly indicate, the assumption upon which chronic absenteeism is based—that both excused and unexcused absence days should be used when determining student risk status—is questionable, and the use of chronic absenteeism as an early warning indicator deserves serious attention. The research presented here demonstrates that excused absence days have no diagnostic efficacy in determining academic risk status in the early elementary grades; on the other hand, unexcused absences are useful classifiers of academic risk status as early as kindergarten

and appear to grow in their utility across the early primary years. This implies that, as an early warning indicator*, chronic absenteeism should only be based upon unexcused absences, which directly challenges the way it is currently being operationalized*.

Because chronic absenteeism, in its current operationalization, is being used nationwide, the educational risk status of thousands of children is potentially being misidentified; students that are not at risk for educational problems may be identified as at risk and may receive limited resources that provide unnecessary remediation and support, while students that are at risk for educational problems may be identified as on track for positive outcomes and miss out on the critical supports that could help them become successful students (Gottfried, 2009). Chronic absenteeism may, therefore, be producing undue numbers of "false positives" and "false negatives" in terms of academic risk status, which represents a threat to both the well-being of vulnerable students and the prudent allocation of scarce resources. This is problematic for both policymakers and practitioners, as it undermines the entire goal of using early warning indicators to identify at-risk students and provide them with support. While the findings from this study may not be generalizable to all contexts and require further replication before definitive decisions about chronic absenteeism are made, they should, at minimum, require policymakers and practitioners to be more cautious in their use of chronic absenteeism. It also requires policymakers to prioritize studies that test the underlying assumptions of chronic absenteeism, especially among the youngest students in the public education system. Particularly within large, urban school districts, the use of chronic absenteeism as

an early warning indicator should be carefully considered until additional empirical

evidence confirms or challenges the results presented here.

Caution in using chronic absenteeism as an early warning indicator is especially

important in contexts like Philadelphia and other large urban cities, where the findings of

the current study may be most applicable. It is even more essential for policymakers and

practitioners in large urban cities to rethink the use of this indicator if these cities are

nested in states that have already adopted chronic absenteeism and are mandating its use.

Pennsylvania, for instance, has adopted chronic absenteeism as its measure of school

quality and student success under ESSA (Jordan & Miller, 2017). It is also being used in

practice as an early warning indicator in Philadelphia schools to identify students at risk

(Balfanz et al., 2007; Neild & Balfanz, 2006). This means that the School District of

Philadelphia is required to calculate and report incidences of chronic absenteeism and

that some schools are actually using chronic absenteeism to identify at-risk students and

provide them with additional resources; in light of the current study, the use of chronic

absenteeism in Philadelphia is particularly troubling. It is thus critical for policymakers

and practitioners in Philadelphia and analogous contexts to reconsider the use of chronic

absenteeism as it is currently being defined and invest in further research to determine

whether the indicator should be changed in order to make it as efficacious as possible.

Careful consideration of the adoption and implementation of chronic absenteeism

as an early warning indicator is further necessitated by the lack of empirical

substantiation of the second major assumption upon which it is based—the 18-day (or

10%) threshold that serves as a cut point for determining whether a student can be

categorized as chronically absent (Gershenson et al., 2017). While this study did not address the issue of establishing an empirically derived cut score, it is an essential next step in the research agenda around chronic absenteeism and will have important implications for policy and practice. The fact that the 18-day threshold for chronic absenteeism has been so widely accepted without any research evidence to validate its use as an early warning indicator for specific populations of students should be a major source of concern. Because this threshold is arbitrary and is being used as a cut point to determine which students are or are not at risk, it is again a potential source for the misidentification of thousands of children. Policymakers and practitioners using this 18-day cutoff to determine which students are at risk for educational problems should be cautious in using this threshold as a "magic number" at which student risk status is definitive. Again, policymakers must prioritize research that seeks to test this additional underlying assumption of chronic absenteeism and bring sound science to bear on the creation of a cut score for this indicator. Additionally, applying a scientifically sound process for determining cut scores would be informative in generating the appropriate thresholds for different *levels* of risk status. As differentiating students into tiers of risk (e.g., low, medium, high) has become more popular in educational practice in recent years, it is essential that these tiers are constructed with scientific rigor (Attendance Works, 2018; Kearney & Graczyk, 2014). Determining the appropriate threshold of absence days for chronic absenteeism, especially within the early elementary grades, should be an essential consideration for policymakers and practitioners to ensure the accurate early identification of at-risk students.

Another important implication for policy and practice elicited by this study relates to efforts to support family engagement with the school as a primary mechanism to ameliorate educational issues, especially for young students. As the absence research shows, consistent school attendance for young children is highly influenced by the family context; furthermore, unexcused absences often serve as a signal for parental disengagement with the schooling process (Gottfried, 2009; Jeynes, 2003; Klerman & Glasscock, 1996; Lehr et al., 2004; Maynard et al., 2012; McNeal, 1999; Romero & Lee, 2008; Sheldon, 2007; Teasley, 2004). The findings of this study support this previous research and indicate the importance of supporting and engaging families as part of efforts to reduce student absences and prevent negative educational outcomes. Many of the most promising interventions designed to reduce student absences involve connecting families to critical resources that prevent structural issues (such as housing instability or lack of reliable transportation) from disrupting student attendance; these interventions also promote open communication and building bonds between families and schools to foster familial engagement with the education process, which in turn reduces the incidence of unexcused absences (Balfanz, 2016; Chang & Romero, 2008; Faria et al., 2017; Jordan et al., 2018; Kearney & Graczyk, 2014; Lehr et al., 2004; Reimer & Dimock, 2005; Sheldon, 2007; Sutphen et al., 2010; Teasley, 2004). In demonstrating that unexcused absences have a differential relationship to student outcomes, the research presented here supports the idea that working with families to connect them to supports and engage them in the education of young children is a crucial avenue by which to ameliorate student absences and the negative educational effects to which they are linked.

Policymakers and school personnel should take efforts to engage families into consideration when determining which interventions are worthy of investment and which practices to utilize in schools.

Finally, this study speaks to the need for policymakers and practitioners to ensure that attendance data are high quality. Because attendance information is collected as soon as students enter school, is universally recorded across all schools, and is recorded throughout a student's educational life, it represents an essential source of information. As this study shows, much can be gleaned from the use of absences as early indicators of future educational problems; this requires, however, that absence data be valid and reliable. Policymakers and practitioners should consider the mechanisms by which absence data are recorded. Schools and school districts must establish clear protocols for defining excused and unexcused absences and should provide training and assistance for the staff who are responsible for making discernments between these two absence types and recording them. Furthermore, policymakers and practitioners should consider the use of data audits for absence information; while routine data audits may be impractical, conducting sporadic checks on data quality might reveal issues in the process of recording attendance information and provide opportunities for improvement. Because attendance data are so readily available and so potentially valuable in predicting student outcomes, policymakers and practitioners must work to ensure that these data are of the highest possible quality for use.

Summary and Conclusion

Chronic absenteeism has been widely adopted as an early warning indicator for poor academic achievement and a national measure of school quality and student success. As such, the need to ensure the scientific integrity of this policy indicator is vital. Major theoretical assumptions underlying chronic absenteeism, however, have never been empirically tested. The current study represents the first effort to scientifically test one of the most basic assumptions upon which chronic absenteeism is based—both excused and unexcused absences have comparable diagnostic accuracy in the early identification of academically at-risk students. Using the state-of-the-art receiver operating characteristic methodology for determining diagnostic accuracy, this study presented evidence that *only unexcused abse*nces provided diagnostic accuracy for academic risk status in math and ELA for an entire cohort of young students in the School District of Philadelphia. This diagnostic accuracy was evident as early as kindergarten and increased across the early elementary years. Excused absences provided *no* diagnostic utility in differentiating between students at risk for academic problems and students on track for academic success. Overall, these findings indicate that chronic absenteeism could be made a more effective early warning indicator for students in large urban school districts, like Philadelphia, by taking absence types into account.

The findings presented here raise serious questions about how chronic absenteeism is currently being operationalized and call for researchers and policymakers to prioritize rigorous studies that test the generalizability of these results to other school contexts. In addition, this study necessitates further research that carefully examines the

other untested assumption upon which chronic absenteeism is based—the use of the 10%

cut point as a meaningful threshold for differentiating risk status. It is essential that

researchers empirically validate these two critical underlying assumptions of chronic

absenteeism to ensure that policymakers and practitioners can continue their widespread

use of this early warning indicator. Failure to substantiate the evidence base for chronic

absenteeism represents a potential threat to the educational well-being of children in

school systems where it is currently being implemented, as it may misclassify the risk

status of thousands of the nation's youngest students. Ensuring that chronic absenteeism

is scientifically sound for all public-school systems in the U.S. will generate the most

effective early identification of students at risk for future educational problems. This

effective early identification allows for the appropriate allocation of limited but vital

resources for these at-risk students at the earliest possible age. Providing these students

with the supports they need in the most formative years of their educational trajectories

helps ensure that they have what they need to be successful in school and in life.

BIBLIOGRAPHY

Allen, G. (2003). The impact of elementary school nurses on student attendance. *The Journal of School Nursing*, 19, 225-231.

Allensworth, E. M., & Easton, J. Q. (2007). *What matters for staying on-track and graduating in Chicago public high schools: A close look at course grades, failures, and attendance in the freshman year. Research report*. Chicago, IL: Consortium on Chicago School Research.

Allensworth, E. M., Gwynne, J. A., Moore, P., & de la Torre, M. (2014). *Looking forward to high school and college: middle grade indicators of readiness in Chicago public schools*. Chicago, IL: University of Chicago Consortium on School Research.

Annie E. Casey Foundation (2010). *Early warning! Why reading by the end of third grade matters: A KIDS COUNT special report*. Baltimore, MD: Annie E. Casey Foundation.

Applied Survey Research. (2011*). Attendance in early elementary grades associations with student characteristics, school readiness, and third grade outcomes*. San Jose, CA: Applied Survey Research.

Attendance Works (2018). *Three tiered system of intervention*. Retrieved from https://www.attendanceworks.org/resources/toolkits/

Baker, S. G. (2003). The central role of receiver operating characteristic (ROC) curves in evaluating tests for the early detection of cancer. *Journal of the National Cancer Institute*, *95*(7), 511-515.

Baker, M. L., Sigmon, J. N., & Nugent, M. E. (2001). *Truancy reduction: Keeping students in school. Juvenile justice bulletin*. Washington, D.C.: U.S. Department of Justice.

Balfanz, R. (2016). Missing school matters. *Phi Delta Kappan*, *98*(2), 8-13.

Balfanz, R., & Byrnes, V. (2013). *Meeting the challenge of combating chronic absenteeism: Impact of the NYC Mayor's Interagency Task Force on chronic absenteeism and school attendance and its implications for other cities*. Baltimore, MD: Everyone Graduates Center.

Balfanz, R., Herzog, L., & Mac Iver, D. J. (2007). Preventing student disengagement and keeping students on the graduation path in urban middle-grades schools: Early identification and effective interventions. *Educational Psychologist*, *42*(4), 223-235.

Bauer, L., Jordon, P.W., Chang, H.N., & Balfanz, R. (2018). *Taking attendance seriously in the new civil rights data collection.* Washington, DC: Brookings.

Bauer, L., Liu, P., Schanzenbach, D. W., & Shambaugh, J. (2018). *Reducing chronic absenteeism under the every student succeeds act*. Washington, DC: Brookings.

Berliner, D. C. (2009). *Poverty and potential: Out-of-school factors and school success*. Washington DC: Education Policy Research Unit.

Blankenship, J. M. (2015). *The impact of excused versus unexcused absences on academic student achievement among Hispanic economically disadvantaged elementary students*. (Doctoral dissertation, Lamar University). Retrieved from ProQuest Dissertations and Theses database (Proquest no. 10110324).

Blazer, Christie (2011). *Chronic absenteeism in the elementary grades*. Miami, FL: Information Capsule Research Services.

Bossuyt, P. M., Reitsma, J. B., Bruns, D. E., Gatsonis, C. A., Glasziou, P. P., Irwig, L. M., ... & Lijmer, J. G. (2003). The STARD statement for reporting studies of diagnostic accuracy: explanation and elaboration. *Clinical chemistry*, *49*(1), 7-18.

Brazelton, T. B., & Greenspan, S. I. (2009). *The irreducible needs of children: What every child must have to grow, learn, and flourish*. Da Capo Lifelong Books.

Bronfenbrenner, U. (2005). The bioecological theory of human development. In U. Bronfenbrenner (Ed.), *Making human beings human: Bioecological perspectives on human development*. Thousand Oaks, CA: Sage.

Bronfenbrenner, U., & Morris, P. A. (1998). The ecology of developmental processes. In W. Damon & R. M. Lerner (Eds.), *Handbook of child psychology: Theoretical models of human development* (pp. 993-1028). Hoboken, NJ: John Wiley & Sons Inc.

Bruce, M., Bridgeland, J. M., Fox, J. H., & Balfanz, R. (2011). On Track for Success: The Use of Early Warning Indicator and Intervention Systems to Build a Grad Nation. *Civic Enterprises*.

Caldas, S. J. (1993). Reexamination of input and process factor effects on public school achievement. *The Journal of Educational Research*, *86*(4), 206-214.

Chang, H. N., & Davis, R. (2015). *Mapping the early attendance gap: Charting a course for school success*. San Francisco, CA: Attendance Works.

Chang, H. N., & Romero, M. (2008). *Present, engaged, and accounted for: The critical importance of addressing chronic absence in the early grades*. New York City, NY: National Center for Children in Poverty.

Chen, C. C., Culhane, D. P., Metraux, S., Park, J. M., & Venable, J. C. (2016). The heterogeneity of truancy among urban middle school students: A latent class growth analysis. *Journal of Child and Family Studies*, *25*(4), 1066-1075.

Chen, P., & Rice, C. (2016). *Showing up matters: Newark chronic absenteeism in the early years*. Newark, NJ: Advocates for Children of New Jersey.

Christie, K. (2005). *Target attendance and graduation rates and how rates are calculated*. Washington, D.C.: Education Commission of the States.

Christie, K. (2006). Counting the truants. *Phi Delta Kappan*, *87*(7), 485.

Clogg, C. C., Petkova, E., & Haritou, A. (1995). Statistical methods for comparing regression coefficients between models. *American Journal of Sociology*, *100*(5), 1261-1293.

Coleman, J. S., Campbell, E., Hobson, C., McPartland, J., Mood, A., Weinfeld, F., & York, R. (1966). *Equality of educational opportunity*. Washington D.C.: U.S. Department of Health, Education, and Welfare, Office of Education.

Data Recognition Corporation. (2015). *Technical report for the 2015 Pennsylvania system of school assessment*. Retrieved from http://www.education.pa.gov/ Documents/K-12/Assessment%20and%20Accountability/PSSA/Technical% 20Reports/2015%20PSSA%20Technical%20Report.pdf

Dembo, R., & Gulledge, L. M. (2009). Truancy intervention programs: Challenges and innovations to implementation. *Criminal Justice Policy Review*, *20*(4), 437-456.

Demler, O. V., Pencina, M. J., & D'Agostino Sr, R. B. (2012). Misuse of DeLong test to compare AUCs for nested models. *Statistics in Medicine*, *31*(23), 2577-2587.

DeLong, E. R., DeLong, D. M., & Clarke-Pearson, D. L. (1988). Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*, 837-845.

Dougherty, S. M. (2018). How measurement and modeling of attendance matter to assessing dimensions of inequality. *Journal of Education for Students Placed at Risk*, *23*(1-2), 9-23.

Dryfoos, Joy G (1991). *Adolescents at risk: Prevalence and prevention*. Oxford University Press.

Duncan, G. J., & Magnuson, K. (2011). The nature and impact of early achievement skills, attention skills, and behavior problems. *Whither opportunity*, 47-70.

Ehrlich, S. B., Gwynne, J. A., Stitziel Pareja, A., Allensworth, E. M., Moore, P., Jagesic, S., & Sorice, E. (2014). *Preschool attendance in Chicago public schools: relationships with learning outcomes and reasons for absences*. Chicago, IL: Consortium on Chicago School Research.

Engelbrecht, M. R., Jager, G. J., Laheij, R. J., Verbeek, A. L., Van Lier, H. J., & Barentsz, J. O. (2002). Local staging of prostate cancer using magnetic resonance imaging: a meta-analysis. *European Radiology*, *12*(9), 2294-2302.

Entwisle, A., Alexander, K. L., & Steffel, L. (1987). Olson. 1997. *Children, schools, and inequality*. New York, NY: Westview Press.

Every Student Succeeds Act (ESSA) of 2015, Pub.L. 114–95 (2015).

Faria, A. M., Sorensen, N., Heppen, J., Bowdon, J., Taylor, S., Eisner, R., & Foster, S. (2017). *Getting students on track for graduation: Impacts of the early warning intervention and monitoring system after one year*. Washington, D.C.: Regional Educational Laboratory Midwest.

Fryer Jr, R. G., & Levitt, S. D. (2006). The black-white test score gap through third grade. *American Law and Economics Review*, *8*(2), 249-281.

Garry, E. M. (1996). Truancy: First step to a lifetime of problems. *Juvenile justice bulletin*. Washington, D.C.: U.S. Department of Justice.

Gershenson, S., Jacknowitz, A., & Brannegan, A. (2017). Are student absences worth the worry in US primary schools?. *Education Finance and Policy*, *12*(2), 137-165.

Ginsburg, A., Jordan, P., & Chang, H. (2014). *Absences add up: How school attendance influences student success*. San Francisco, CA: Attendance Works.

Gönen, M. (2007). *Analyzing receiver operating characteristic curves with SAS*. Cary, NC: SAS Institute.

Gonzales, R., Richards, K., & Seeley, K. (2002). *Youth out of school: Linking absence to delinquency*. Denver, Colorado: National Center for School Engagement at the Colorado Foundation for Families and Children.

Goodman, J. (2014). *Flaking out: Student absences and snow days as disruptions of instructional time* (No. w20221). National Bureau of Economic Research.

Gottfried, M. A. (2009). Excused versus unexcused: How student absences in elementary school affect academic achievement. *Educational Evaluation and Policy Analysis*, *31*(4), 392-415.

Gottfried, M. A. (2010). Evaluating the relationship between student attendance and achievement in urban elementary and middle schools: An instrumental variables approach. *American Educational Research Journal*, *47*(2), 434-465.

Gottfried, M. A. (2011a). Absent peers in elementary years: The negative classroom effects of unexcused absences on standardized testing outcomes. *Teachers College Record*, *113*(8), 1597-1632.

Gottfried, M. A. (2011b). The detrimental effects of missing school: Evidence from urban siblings. *American Journal of Education*, *117*(2), 147-182.

Gottfried, M. A. (2013). Quantifying the consequences of missing school: Linking school nurses to student absences to standardized achievement. *Teachers College Record*, *115*(6), 1-30.

Gottfried, M. A. (2014). Chronic absenteeism and its effects on students' academic and socioemotional outcomes. *Journal of Education for Students Placed at Risk*, *19*(2), 53-75.

Gottfried, M. A. (2014a). The influence of tardy classmates on students' socio-emotional outcomes. *Teachers College Record*, *116*(3), 1-35.

Gottfried, M. A. (2015). Chronic absenteeism in the classroom context: Effects on achievement. *Urban Education*, https://doi.org/10.1177/0042085915618709.

Gottfried, M. A. (2017). Does Truancy Beget Truancy? Evidence from Elementary School. *The Elementary School Journal*, *118*(1), 128-148.

Gottfried, M. A., & Gee, K. A. (2017). Identifying the determinants of chronic absenteeism: A bioecological systems approach. *Teachers College Record*, *119*(7), 1-34.

Gottfried, M. A., & Kirksey, J. J. (2017). "When" students miss school: The role of timing of absenteeism on students' test performance. *Educational Researcher*, *46*(3), 119-130.

Gottfried, M., Stiefel, L., Schwartz, A. E., & Hopkins, B. (2019). Showing up: Disparities in chronic absenteeism between students with and without disabilities. *Teachers College Record*, *121*(8), 1-34.

Guare, R. E., & Cooper, B. S. (2003). *Truancy revisited: Students as school consumers*. Lanham, MD: Scarecrow Press.

Guttu, M., Engelke, M., & Swanson, M. (2004). Does the school nurse-to-student ratio make a difference? *Journal of School Health*, 74, 6-9.

Hajian-Tilaki, K. (2013). Receiver operating characteristic (ROC) curve analysis for medical diagnostic test evaluation. *Caspian Journal of Internal Medicine*, *4*(2), 627.

Hallan, S., & Åsberg, A. (1997). The accuracy of C-reactive protein in diagnosing acute appendicitis—a meta-analysis. *Scandinavian Journal of Clinical and Laboratory Investigation*, *57*(5), 373-380.

Hanley, J. A., & McNeil, B. J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, *143*(1), 29-36.

Harwell, M., & LeBeau, B. (2010). Student eligibility for a free lunch as an SES measure in education research. *Educational Researcher*, *39*(2), 120-131.

Heilbrunn, J. (2007). *Pieces of the truancy jigsaw: A literature review*. Denver, CO: National Center for School Engagement.

Henderson, C., Barghaus, K., Fantuzzo, J., Brumley, B., Coe, K., & LeBoeuf, W. (2018, May). *The classroom engagement scale: Validity evidence and implications for use*. Poster session presented at the National Research Conference on Early Childhood, Washington, D.C.

Heppen, J. B., & Therriault, S. B. (2008). *Developing early warning systems to identify potential high school dropouts. Issue brief*. Washington, D.C.: National High School Center.

Hibbett, A., Fogelman, K., & Manor, O. (1990). Occupational outcomes of truancy. *British Journal of Educational Psychology*, *60*(1), 23-36.

Hickman, G. P., & Heinrich, R. S. (2011). *Do children drop out of school in kindergarten?*. Lanham, MD: Rowman & Littlefield Education.

Hill, C. J., Bloom, H. S., Black, A. R., & Lipsey, M. W. (2008). Empirical benchmarks for interpreting effect sizes in research. *Child Development Perspectives*, *2*(3), 172-177.

Howard, L. C. (2010). *The great equalizer? A five state study of compulsory school attendance age policy and administration* (Doctoral dissertation, University of Pennsylvania). Retrieved from ProQuest Dissertations and Theses database (UMI no. 3447486).

Huizinga, D., Loeber, R., Thornberry, T. P., & Cothern, L. (2000). Co-occurrence of delinquency and other problem behaviors. *Juvenile justice bulletin*, *11*, 1-7. Washington, D.C.: U.S. Department of Justice.

Huston, A. C., & Bentley, A. C. (2010). Human development in societal context. *Annual Review of Psychology*, *61*, 411-437.

Jacob, B. A., & Lovett, K. (2017). *Chronic absenteeism: An old problem in search of new answers*. Washington, DC: Brookings.

Jarjoura, G. R. (1993). Does dropping out of school enhance delinquent involvement? Results from a large-scale national probability sample. *Criminology*, 31(2), 149-172.

Jeynes, W. H. (2003). A meta-analysis: The effects of parental involvement on minority children's academic achievement. *Education and Urban Society*, *35*(2), 202-218.

Jordan, P. W., & Miller, R. (2017). *Who's in: Chronic absenteeism under the Every Student Succeeds Act*. Washington, D.C.: FutureEd.

Jordon, P.W., Fothergill, S., & Rosende, M. (2018). *Writing the rules: Ensuring chronic absenteeism data works for schools and students*. Washington, D.C.: FutureEd.

Jordan, N. C., Glutting, J., Ramineni, C., & Watkins, M. W. (2010). Validating a number sense screening tool for use in kindergarten and first grade: Prediction of mathematics proficiency in third grade. *School Psychology Review*, *39*(2), 181-196.

Kagan, S. L., & Kauerz, K. (2006). Making the most of kindergarten: Trends and policy issues. In D. Gullo (Ed.), *Teaching and learning in the kindergarten year* (pp. 161-170). Washington, DC: National Association for the Education of Young Children.

Karantonis, A., & Sireci, S. G. (2006). The bookmark standard-setting method: A literature review. *Educational Measurement: Issues and Practice*, *25*(1), 4-12.

Kearney, C. A. (2008a). An interdisciplinary model of school absenteeism in youth to inform professional practice and public policy. *Educational Psychology Review*, *20*(3), 257-282.

Kearney, C. A. (2008b). School absenteeism and school refusal behavior in youth: A contemporary review. *Clinical Psychology Review*, *28*(3), 451-471.

Kearney, C. A., & Graczyk, P. (2014, February). A response to intervention model to promote school attendance and decrease school absenteeism. *Child & Youth Care Forum*, 43(1), 1-25.

Klerman, L. V., & Glasscock, B. C. (1996). Features of children who do not attend school. In I. Berg & J Nursten (Eds.), *Unwillingly to school*, (25-37). London, UK: Gaskell.

Kotin, L., & Aikman, W. F. (1980). *Legal foundations of compulsory school attendance*. New York, NY: National University Publications, Kennikat Press.

Lamdin, D. J. (1996). Evidence of student attendance as an independent variable in education production functions. *The Journal of Educational Research*, *89*(3), 155-162.

Landes, W. M., & Solmon, L. C. (1972). Compulsory schooling legislation: An economic analysis of law and social change in the nineteenth century. *The Journal of Economic History*, *32*(1), 54-91.

Lee, V. E., & Burkam, D. T. (2002). *Inequality at the starting gate: Social background differences in achievement as children begin school*. Washington D.C.: Economic Policy Institute.

Lee, C. M. Y., Huxley, R. R., Wildman, R. P., & Woodward, M. (2008). Indices of abdominal obesity are better discriminators of cardiovascular risk factors than BMI: a meta-analysis. *Journal of Clinical Epidemiology*, *61*(7), 646-653.

Lehr, C. A., Sinclair, M. F., & Christenson, S. L. (2004). Addressing student engagement and truancy prevention during the elementary school years: A replication study of the check & connect model. *Journal of Education for Students Placed at Risk*, *9*(3), 279-301.

Lleras, C. (2008). Race, racial concentration, and the dynamics of educational inequality across urban and suburban schools. *American Educational Research Journal*, *45*(4), 886-912.

Logan, J. R., Minca, E., & Adar, S. (2012). The geography of inequality: Why separate means unequal in American public schools. *Sociology of education*, *85*(3), 287-301.

London, R. A., Sanchez, M., & Castrechini, S. (2016). The dynamics of chronic absence and student achievement. *Education Policy Analysis Archives*, *24*, 112.

Mac Iver, M. A. (2010*). Gradual disengagement: A portrait of the 2008-2009 dropouts in the Baltimore.* Baltimore, MD: Baltimore Education Research Consortium.

Manning, J. B., & Kovach, J. A. (2003). The continuing challenges of excellence and equity. In B. Williams (Ed.), *Closing the achievement gap: A vision for changing beliefs and practices*, (25-47). Alexandria, VA: Association for Supervision and Curriculum Development.

Maynard, B. R., Salas-Wright, C. P., Vaughn, M. G., & Peters, K. E. (2012). Who are truant youth? Examining distinctive profiles of truant youth using latent profile analysis. *Journal of Youth and Adolescence*, *41*(12), 1671-1684.

Maughan, D., & Moore, S. C. (2010). Dimensions of child neglect: an exploration of parental neglect and its relationship with delinquency. *Child Welfare*, *89*(4).

McFall, R. M., & Treat, T. A. (1999). Quantifying the information value of clinical assessments with signal detection theory. *Annual Review of Psychology*, *50*(1), 215-241.

McNeal Jr, R. B. (1999). Parental involvement as social capital: Differential effectiveness on science achievement, truancy, and dropping out. *Social Forces*, *78*(1), 117-144.

Metz, C. E., & Pan, X. (1999). "Proper" binormal ROC curves: theory and maximum-likelihood estimation. *Journal of Mathematical Psychology*, *43*(1), 1-33.

Mitzel, H. C., Lewis, D. M., Patz, R. J., & Green, D. R. (2001). The Bookmark procedure: Psychological perspectives. In G. Cizek (Ed.), *Setting performance standards: Concepts, methods and perspectives* (pp. 249–281). Mahwah, NJ: Erlbaum.

Mogulescu, S., & Segal, H. J. (2002). *Approaches to truancy prevention*. New York, NY: Vera Institute of Justice.

Morrissey, T. W., Hutchison, L., & Winsler, A. (2014). Family income, school attendance, and academic achievement in elementary school. *Developmental Psychology*, *50*(3), 741.

Mossman, D. (1994). Assessing predictions of violence: being accurate about accuracy. *Journal of consulting and clinical psychology*, *62*(4), 783.

Mueller, D., & Stoddard, C. (2006). Dealing with chronic absenteeism and its related consequences: The process and short-term effects of a diversionary juvenile court intervention. *Journal of Education for Students Placed at Risk*, *11*(2), 199-219.

NAEP District Profiles. (n.d.). Retrieved July 15, 2019, retrieved from https://www.nationsreportcard.gov/profiles/districtprofile?chort=1&sub=MAT&sj=XQ&sfj=NL&st=MN&year=2017R3.

Nauer, K. (2016). Battling chronic absenteeism. *Phi Delta Kappan*, *98*(2), 28-34.

Neild, R. C., & Balfanz, R. (2006). Unfulfilled Promise: The Dimensions and Characteristics of Philadelphia's Dropout Crisis, 2000-2005. *Philadelphia Youth Network*.

No Child Left Behind Act of 2001, P.L. 107-110, 20 U.S.C. § 6319 (2002).

Ogilvie, G. S., Patrick, D. M., Schulzer, M., Sellors, J. W., Petric, M., Chambers, K., ... & FitzGerald, J. M. (2005). Diagnostic accuracy of self collected vaginal specimens for human papillomavirus compared to clinician collected human papillomavirus specimens: a meta-analysis. *Sexually Transmitted Infections*, *81*(3), 207-212.

Ou, S. R., & Reynolds, A. J. (2008). Predictors of educational attainment in the Chicago longitudinal study. *School Psychology Quarterly*, *23*(2), 199.

Patwardhan, M. B., McCrory, D. C., Matchar, D. B., Samsa, G. P., & Rutschmann, O. T. (2004). Alzheimer disease: operating characteristics of PET—a meta-analysis. *Radiology*, *231*(1), 73-80.

Penn Child Research Center (2017). *A comprehensive examination of the School District of Philadelphia's kindergarten classroom engagement scale (CES)*. Philadelphia, PA: Author.

Pennsylvania School Board Association (2015). *State board of education approves new PSSA cut scores*. Retrieved from https://www.psba.org/2015/07/state-board-of-education-approves-new-pssa-cut-scores

Phillips, M., Crouse, J., & Ralph, J. (1998). Does the Black-White test score gap widen after children enter school. In C. Jencks & M. Philips (Eds.), *The Black-White test score gap* (229-272). Washington, D.C.: Brookings Institution Press.

Pintea, S., & Moldovan, R. (2009). The receiver-operating characteristic (ROC) analysis: Fundamentals and applications in clinical psychology. *Journal of Cognitive & Behavioral Psychotherapies*, *9*(1).

Ready, D. D. (2010). Socioeconomic disadvantage, school attendance, and early cognitive development: The differential effects of school exposure. *Sociology of Education*, *83*(4), 271-286.

Reardon, S. F. (2011). The widening academic achievement gap between the rich and the poor: New evidence and possible explanations. In G. Duncan & R. Murnane (Eds.), *Whither opportunity* (91-116). New York, NY: Russel Sage Foundation.

Rice, M. E., & Harris, G. T. (2005). Comparing effect sizes in follow-up studies: ROC Area, Cohen's d, and r. *Law and Human Behavior*, *29*(5), 615-620.

Roby, D. E. (2004). Research on school attendance and student achievement: A study of Ohio schools. *Educational Research Quarterly*, *28*(1), 3-16.

Rodrigues, J. (2017). *Developing fraction screeners to identify children at risk for mathematics difficulties* (Doctoral dissertation, University of Delaware). Retrieved from ProQuest Dissertations and Theses database (Proquest no. 10616109).

Romero, M., & Lee, Y. S. (2008). *The influence of maternal and family risk on chronic Absenteeism in early schooling*. New York City, NY: National Center for Children in Poverty.

Sandy, J., & Duncan, K. (2010). Examining the achievement test score gap between urban and suburban students. *Education Economics*, *18*(3), 297-315.

Schoeneberger, J. A. (2012). Longitudinal attendance patterns: Developing high school dropouts. *The clearing House: A journal of Educational Strategies, Issues and Ideas*, *85*(1), 7-14.

School District of Philadelphia (2018). *Administrative procedures for attendance and truancy*. Retrieved from https://www.philasd.org/studentrights/wp-content/uploads/sites/67/2018/09/Attendance-Guidelines-Rev.-9-18-18.pdf

Sheldon, S. B. (2007). Improving student attendance with school, family, and community partnerships. *The Journal of Educational Research*, *100*(5), 267-275.

Sheldon, S. B., & Epstein, J. L. (2004). Getting students to school: Using family and community involvement to reduce chronic absenteeism. *School Community Journal*, *14*(2), 39-56.

Smink, J., & Heilbrunn, J. Z. (2006). *Legal and economic implications of truancy. Truancy prevention in action*. Anderson, SC: National Dropout Prevention Center/Network.

Smolkowski, K., & Cummings, K. D. (2015). Evaluation of diagnostic systems: The selection of students at risk of academic difficulties. *Assessment for Effective Intervention*, *41*(1), 41-54.

Spencer, A. M. (2009). School attendance patterns, unmet educational needs, and truancy: A chronological perspective. *Remedial and Special Education*, *30*(5), 309-319.

Spradlin, T., Cierniak, K., Shi, D., & Chen, M. (2012). *Attendance and chronic absenteeism in Indiana: The impact on student achievement*. Education Policy Brief, Volume 10, Number 3. Bloomington, IN: Center for Evaluation and Education Policy.

Sugrue, E. P., Zuel, T., & LaLiberte, T. (2016). The ecological context of chronic school absenteeism in the elementary grades. *Children & Schools*, *38*(3), 137-145.

Sutphen, R. D., Ford, J. P., & Flaherty, C. (2010). Truancy interventions: A review of the research literature. *Research on Social Work Practice*, *20*(2), 161-171.

Swets, J. A., Dawes, R. M., & Monahan, J. (2000). Psychological science can improve diagnostic decisions. *Psychological Science in the Public Interest*, *1*(1), 1-26.

Teasley, M. L. (2004). Absenteeism and truancy: Risk, protection, and best practice implications for school social workers. *Children & Schools*, *26*(2), 117-128.

Tyack, D. B. (1974). *The one best system: A history of American urban education* (Vol. 95). Cambridge, MA: Harvard University Press.

Tyack, D., & Berkowitz, M. (1977). The man nobody liked: Toward a social history of the truant officer, 1840-1940. *American Quarterly*, *29*(1), 31-54.

U.S. Department of Agriculture (2018). *Applying for free and reduced priced school meals*. Retrieved from https://www.fns.usda.gov/school-meals/applying-free-and-reduced-price-school-meals

U.S. Department of Education, Office for Civil Rights. (2016). *2013-2014 civil rights data collection: A first look*. Washington, DC: Author.

Utah Education Policy Center. (2012). *Research brief: Chronic absenteeism*. Salt Lake City, UT: Author.

West, T. C. (2013). *Just the right mix: Identifying potential dropouts in Montgomery County public schools using an early warning indicators approach*. Rockville, MD: Montgomery County Public Schools.

Wieman, G. (2007). *Third grade standardized tests as indicators and estimates of future performance on the ACT* (Master's Thesis, University of Michigan). Retrieved from https://commons.emich.edu/cgi/viewcontent.cgi?referer=https://www.google.com/&httpsredir=1&article=1057&context=theses

Williams, B. (2003). What else do we need to know and do. In B. Williams (Ed.), *Closing the achievement gap: A vision for changing beliefs and practices*, (13-24). Alexandria, VA: Association for Supervision and Curriculum Development.

Wilson, J., Olinghouse, N. G., McCoach, D. B., Santangelo, T., & Andrada, G. N. (2016). Comparing the accuracy of different scoring methods for identifying sixth graders at risk of failing a state writing assessment. *Assessing Writing*, *27*, 11-23.

Wright, S. E., & Wardle, M. (1996). Absences in a primary school. In I. Berg & J Nursten (Eds.), *Unwillingly to school*, (57-71). London, UK: Gaskell.

Youngstrom, E. A. (2014). A primer on receiver operating characteristic analysis and diagnostic efficiency statistics for pediatric psychology: we are ready to ROC. *Journal of Pediatric Psychology*, *39*(2), 204-221.

Zou, K. H., O'Malley, A. J., & Mauri, L. (2007). Receiver-operating characteristic analysis for evaluating diagnostic tests and predictive models. *Circulation*, *115*(5), 654-657.

Zweig, M. H., & Campbell, G. (1993). Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. *Clinical Chemistry*, *39*(4), 561-577.