1-1-2018

Identifying Students at Risk of Dropping out: Indicators and Thresholds Using ROC Analysis

Susan E. Carlson
scarlson11@georgefox.edu

IDENTIFYING STUDENTS AT RISK OF DROPPING OUT: INDICATORS AND THRESHOLDS USING ROC ANALYSIS

by

SUSAN E. CARLSON

FACULTY RESEARCH COMMITTEE:

Chair: Dane C. Joseph, Ph.D.

Member: Susanna M. Thornhill, Ph.D.

Member: Linda Samek, Ed.D.

Presented to the Doctoral Department

and College of Education, George Fox University

In partial fulfillment of the requirements for the degree of

Doctor of Education

May 1, 2018

GEORGE FOX
UNIVERSITY

COLLEGE OF EDUCATION

"IDENTIFYING STUDENTS AT RISK OF DROPPING OUT: INDICATORS AND THRESHOLDS USING ROC ANALYSIS," a Doctoral research project prepared by SUSAN CARLSON in partial fulfillment of the requirements for the Doctor of Education degree in Educational Leadership.

This dissertation has been approved and accepted by:

05/15/18		Committee Chair
Date	Dane Joseph, PhD	Assistant Professor of Education
5-15-18	Linda Samek	
Date	Linda Samek, EdD	Provost
5.15.18		
Date	Susanna Thornhill, PhD	Associate Professor of Education

ABSTRACT

Each year a significant percentage of high school students in the United States do not graduate. School practitioners need accurate indicators for identifying potential dropouts in order to focus scarce intervention resources on students most in need. While the process of dropping out is complex, indicators measured at the end of students' ninth-grade year provide information regarding their future graduation outcome. The current study used receiver operating characteristic (ROC) empirical curves to assess the accuracy of three ninth-grade risk factors, GPA, credits earned, and over-age status, in predicting the likelihood that students sampled for the National Center for Education Statistics High School Longitudinal Survey of 2009, dropped out of high school. The results showed that all three gave better than chance predictions. GPA had a 74 percent probability of correctly distinguishing between dropouts and graduates. The cut point of GPA less than 1.7 identified 48 percent of the dropouts, 88 percent of the graduates, and had a false positive rate of 12 percent. The three indicators provide quantitative data for identifying students at the end of ninth grade who may benefit from strategies designed to keep them on track for graduation. School practitioners may want to conduct a similar analysis using their district data to assess the accuracy of the risk factors for their specific student population.

Keywords: High School Dropout, Receiver Operating Characteristic Curve, Area Under the Curve, GPA, Credits, Over-Age

ACKNOWLEDGEMENTS

This dissertation, like so many others, rests on the shoulders of those who came before. I am grateful for the tireless and passionate work of the researchers referenced in these pages. Their search to understand the dropout problem and the factors associated with early high school identification of students at risk inspired my own. Thank you to the National Center for Education Statistics for conducting comprehensive national longitudinal surveys and for making the collected data available for public use.

Thank you to my dissertation committee for sticking with me though a year-long leave of absence and welcoming me back with encouragement. Dane, as my dissertation chair, you asked timely questions, guided me through the technicalities of formatting, and prompted me to investigate more deeply the missing data issue. I always walk away from our conversations refreshed and inspired that my work is important. Thank you. Susanna, I am grateful for the qualitative perspective you brought to my committee. Your comments, questions, and request for clearer explanations made these pages not only read better but communicate what I was trying to say. Linda, back when I first met you as my statistics professor, I could not have imagined applying those principles to my own research. Thank you for being so interested in this project and taking time away from your busy schedule to support me.

These pages would not have been possible without the consistent support of my husband. Kris, you opened space in our lives so that I could focus on my research. Thank you for being fully present as I wrestled with concepts and ran ideas by you in the evenings after a long day at work. You knew I could do this and made me believe I could too.

Lastly, and most importantly, I acknowledge my Heavenly Father through whom all things are possible. It is his guiding Spirit whom provides ideas, illuminates connections, and

supplies the words. Thank you, Lord, for the gifts of a sound mind, curiosity, and persistence. May this work bring you glory.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

Introduction

Fourteen years ago, our school district, comprised of one comprehensive high school, realized there was a significant number of students each year teetering on the edge of dropping out. The district had been experiencing rapid growth due to the influx of high-tech companies and their affluent employees. Gone were the small-town days when most families worked for the mill and students grew up together. New schools were built, class sizes expanded and curriculum was enhanced to meet the expectations of students preparing for prestigious universities. In the midst of this growth, passionate teachers and concerned parents convinced the district to invest in struggling students by opening an alternative school in a wing of the main high school. Under the leadership of the assistant administrator, the small school began educating a group of students identified as needing intervention. Students with course failures, low credits, behavioral referrals, high detention hours, and high absenteeism were given the opportunity to attend. The administrator interviewed each potential student and handpicked the first three teachers. Under their model of a daily 100-minute academic block schedule, small class sizes, differentiated instruction and assessment, teacher support, and a caring community, the students thrived. Their engagement increased and behavior problems decreased. Students passed classes they previously failed and regained their missing credits. Best of all, students who never thought they would graduate walked across the stage in June confident in their plans for a brighter future.

Today, the school operates within its own building, which the district designed specifically for it. The enrollment has grown from 60 to 150 students. The original administrator still interviews each potential student and handpicks her team of eight certified staff. For two years, I had the privilege of teaching math in this alternative learning community. I came to

know, appreciate, and care deeply for students teetering on the edge of dropping out. I used to ask students to share their stories of why they chose our school. They would tell the same story, over and over. The students walking our halls just could not make it at the big high school on the hill. Some, like one student who found herself pregnant at 17, were pushed out. Others were pulled out by work, family responsibilities, or illness. Some were angry, others withdrawn. Some fell away due to losing interest in school, falling behind in credits, experimenting with drugs, or getting in trouble with the law. Others drifted away by missing so much class time it was impossible for them to recover. Still others were bullied, shamed, teased, or ostracized for being different. No matter what reasons they shared, all were wounded in some way.

At least that is how we teachers described them during lunch in the staff room. We always talked about our students, especially the new ones. With new students, our first order of business was to gain their trust, learn their story, and understand their unique learning needs. So, we shared what we learned and discussed the best ways to entice the new students to engage in this very different high school environment. Usually, it did not take long for our school's family atmosphere, air of respect, and culture of care to melt the protective walls students erected. Then the real work of learning would begin.

We teachers also strategized how we could get our students earlier, before their junior or senior year, before they picked up the burden of failures; before they lost confidence in their intelligence. How, we asked, could we help students, parents, and teachers identify students during their freshman year or earlier, and choose them to thrive in our alternative learning environment? That question became the foundation for my research study. School personnel need accurate indicators to identify students in need of intervention. Scarce resources necessitate

indicators that will identify students likely to drop out while not falsely labeling those who will likely not need intervention to graduate.

Rationale of the Study

This background indicates the beginnings of my quest to identify the most predictive indicators for students in need of support. A survey of the literature on the topic of high school dropouts informed me that the research is both extensive and deep. Longitudinal studies identify factors in early childhood, such as early home environment and quality of early care giving, that influence students' later decisions to drop out of high school (Jimerson, Egeland, Sroufe, & Carlson, 2000). Other studies dive deep into investigating the ways that dropouts differ from graduates resulting in checklists of factors such as the On-Track indicator used by the Consortium on Chicago School Research (Allensworth & Easton, 2005). Still others built models to describe life-course perspectives (Alexander, Entwisle, & Kabbani, 2001), students' mathematics developmental growth pathways (Muthan, 2004), and student engagement trajectories (Janosz, Archambault, Morizot, & Pagani, 2008), all of which hint at a causal relationship between student characteristics and graduation outcomes. While each research study added to my growing knowledge of predictive indicators, it was difficult to identify what research might be most useful to us as teachers at the alternative school. Which factors are most accurate, accessible, and likely to be responsive to intervention? Probabilities and log-odds ratios are meaningful for academia but fail to offer the detailed information that we can understand and put into practice.

Then I found Bowers, Sprott, and Taff's (2013) comprehensive review of 36 research studies containing 110 indicators of students at-risk of dropping out of high school. They compared the precision of each dropout flag, as they termed the indicators, to identify the

student-level characteristics that are most accurate, simple to obtain and usable for schools. Beyond that, their research identified factors that are under the influence of schools rather than demographics. Using the principles of receiver operating characteristics (ROC) analysis, they graphed the true-positive proportion (the percentage of dropouts with the flag) against the false-positive proportion (the percentage of graduates with the flag) of each indicator (see Figure 1). The relative position of each indicator to the reference point of perfect sensitivity indicated how accurately the flag identified dropouts without falsely identifying students who later graduated. In other words, a flag which correctly identified a higher proportion of dropouts without incorrectly classifying graduates as dropouts occupied a position closer to the upper left corner of the graph.

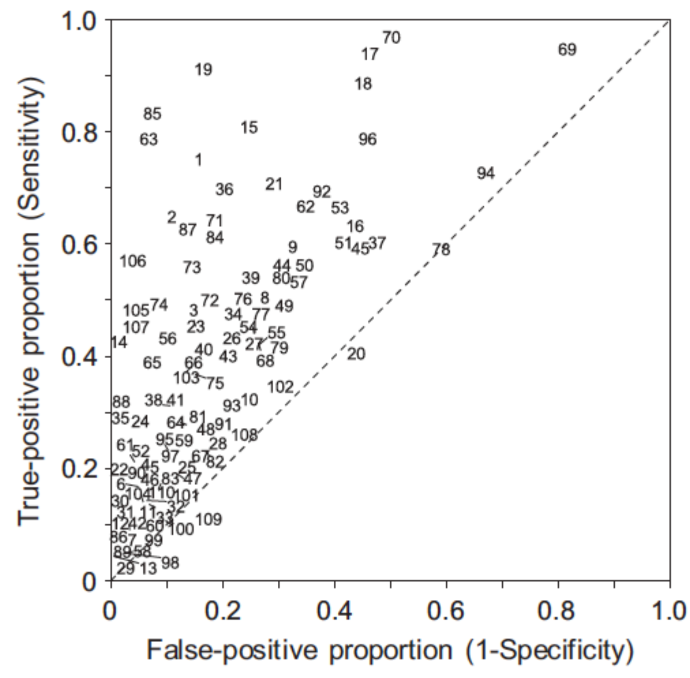

Figure 1: ROC of All Dropout Flags (Bowers, Sprott, & Taff, 2013, p. 94)

This review by Bowers et al. (2013) intrigued me for two reasons. First, they were searching for the type of indicators that could potentially be useful to us teachers at the alternative school. Second, they used the principles of ROC analysis but stopped short of conducting a full analysis which would have given threshold values for each indicator. Threshold values could give information regarding how many failed courses and how low of a grade point average (GPA) accurately predicts students' later dropout decisions. This information could help teachers identify students at the end of their ninth-grade year that are off-course for successful graduation.

This research study extended Bowers et al. (2013) by examining the top three dropout flags (GPA, credits earned, and over-age status) that are readily available to school personnel. It used ROC analysis to report the strength of each flag in predicting dropout. The resulting empirical ROC curves were used to identify potential threshold values to aid future decision-makers in allocating intervention resources.

Purpose of Study

The purpose of this study was to examine the accuracy and precision of ninth-grade indicators in predicting whether students will likely drop out of high school. Building on the theory that dropping out is a cumulative process of disengagement and withdrawal (Alexander et al., 2001; Jimerson et al., 2000), and extending the conceptualization of the diagnostic utility of ninth-grade indicators (Johnson & Semmelroth, 2010; Jordan, Glutting, Ramineni, & Watkins, 2010), this study used GPA, credits earned, and over-age status to predict the likelihood of students' future decisions to drop out. Over-age status was defined as the number of years a student was older than the typical ninth-grade age of 14 years old. Receiver operating characteristic (ROC) curves were used to analyze the accuracy of each indicator drawn from the

National Center for Education Statistics High School Longitudinal Study (HSLS:09) Base Year (2009) in predicting dropping out, as indicated in the HSLS:09 High School Transcript (2013) data. In the HSLS:09, a nationally representative sample of students were surveyed at the end of ninth grade and transcripts were collected the year following their expected four-year graduation date (Ingels et al., 2011; Ingels et al., 2015).

In addition, the ROC curves were evaluated for three threshold values based on typical decision criteria. The threshold values could be used by school staff, providing timely information in the form of possible cut scores, to identify students at the end of ninth grade who were likely to drop out. This study conceptualized identification of dropouts as a first step in a process of guiding all students toward successful graduation. The root causes of students' decisions to leave high school before earning a diploma are complex. The range of problems and issues require individualized intervention strategies tailored to the needs of diverse students (Dynarski & Gleason, 2002). Given the high cost of interventions, this study could give teachers a tool for allocating limited school resources to students most in need.

While the goal of this study was to produce threshold values significant to a nationally representative sample, it offers the analysis process as a tool for future research. This study could be extended by using logistic regression to build a model with all three indicators combined. ROC analysis could provide threshold values for a model significant to a nationally representative sample. Additionally, regionally-specific logistic regression models could be compared with the population model using area under the curve (AUC) analysis. If they differ significantly it would support previous research that threshold values vary by local context (Johnson & Semmelroth, 2010). Individual school districts may want to perform the analysis on

their student data to produce a specific high yield predictive model (Dynarski & Gleason, 2002) and threshold values unique to their student population.

Research Questions

1. How accurately do the ninth-grade indicators of GPA, credits earned, and over-age status predict the likelihood that students from the HSLS:09 dropped out of high school?
2. What are the threshold values for each indicator that optimize three criteria: maximum distance from chance, minimum distance to perfect prediction, and equality of sensitivity and specificity?

Significance of the Study

In the 2014-2015 school year 83 percent of students in the United States graduated with a high school diploma within four years of starting ninth grade (National Center of Education Statistics, 2017b). The other seventeen percent of students earned a GED, received a certificate of completion, were still enrolled, or dropped out. A separate report by the National Center of Education Statistics (2017a) states that in 2015, 5.9% of 38,491,000 individuals aged 16-24 were not enrolled in high school and lacked a high school credential. That translates to 2,270,969 persons who entered the workplace without the educational attainment research shows to be related to successful participation. Students who drop out of high school earn lower-wages (Rouse, 2007), have diminished health and shorter life-spans (Muennig, 2007), participate in more criminal activities (Moretti, 2007), and are more likely to live on welfare (Waldfogel, Garfinkel, & Kelly, 2007) than individuals who graduate.

School personnel need indicators to identify potential dropouts for intervention. For dropout intervention programs to improve student outcomes, the programs need to effectively

identify students at-risk and match interventions to students' characteristics (Dynarski & Gleason, 2002). "Even the highest quality dropout prevention programs will have little influence on the dropout problem if risk factors identify the wrong students (i.e., those who would not otherwise have dropped out)" (Gleason & Dynarski, 2002, p. 25).

While it might seem prudent to identify risk factors that over-capture students likely to drop out, limited school resources necessitate targeting students in real need of help who will most benefit from interventions. Thus, school staff need indicators that accurately identify students likely to drop out while not misidentifying students who would likely graduate. ROC analysis provides a statistical tool to balance this need for sensitivity and specificity (Zou, O'Malley, & Mauri, 2007).

This study focused on three indicators, GPA, credits earned, and over-age status, which meet Bower, Sprott, and Taff's (2002) criteria of being accurate, simple to obtain, usable by schools, and under the influence of schools rather than demographics. Research shows that not only are GPA and credits earned, which is a proxy for course failures, simple to obtain, but they are also good predictors of graduation outcomes when measured at the end of ninth grade (Allensworth & Easton, 2007). GPA and credits earned can be positively influenced by such in-school practices as tutoring and credit recovery. Over-age status is a proxy for grade retention and is easily measured from students' birthdates. Research shows that over-age students are more likely to drop out (MacIver & Messel, 2012). Even though schools cannot influence the prior grade retention of ninth-grade students nor the age at which they entered kindergarten, they can create caring and supportive environments in which older students feel accepted rather than labeled.

This study sought to provide a tool for identifying students who could possibly benefit from intervention. It assumed that identification does not necessitate labeling of students. It assumed that school personnel in general, and teachers in particular will use the tool with professional discretion and educational care as a first step in a research-based dropout prevention program. Teachers assess students' needs continually over the course of a class period. Student attitudes, behaviors, absences, assignment completion, quality of work, and test scores are a few indicators teachers utilize to measure students' educational growth, classroom engagement, and academic achievement. Teachers use these indicators to select students who may need extra instruction, targeted conversations, and/or referrals to additional in-school resources. After selecting students, teachers typically perform other assessments, as informal as conversations or as formal as additional testing, to determine the validity and extent of students' needs. None of the indicators perfectly identify students in need. Some, like students' attitudes and participation, are naturally assessed through teachers' own biases regarding their assumptions of what constitutes successful student behaviors. ROC analysis of GPA, credits earned, and over-age status could provide teachers with a quantitative set of indicators offering better-than-chance accuracy in identifying students who may need intervention.

This research offered an analysis of a recent national data set to the growing body of ROC education literature. By using ROC curves to analyze the diagnostic strength of three ninth-grade early warning indicators of the likelihood for students' future decisions to drop out of high school, I hoped to provide school staff with a set of easily accessible indicators with high predictive value to identify students who may be in need of intervention.

Key Terms, Delimitations, and Limitations

Early Warning Indicators. These were ninth-grade measures of high school outcomes (MacIver, 2013). The purpose of the indicators was to identify students in ninth grade who are showing signs of failure or withdrawal, who are off-track for on-time graduation in four years (Allensworth, 2013).

Graduates. Utilizing the definition used in the formula to calculate the adjusted cohort graduation rate (Curran, Reyna, & NGA, 2009; National Center for Education Statistics, 2017b), this study identified graduates as students who earned their high school diplomas within four years of ninth grade. In the HSLS:09 these students' transcripts indicated "Fall 2012-summer 2013 graduate," and "Pre-fall 2012 graduate." They received their high school diploma early or on-time, it was the gold standard (Zou, O'Malley, & Mauri, 2007).

Dropouts. For the purposes of this study dropouts were students who did not earn their diploma within four years of ninth grade. These were the students in the HSLS:09 whose transcripts indicated "Dropped out," "Left other reason," or "Graduation date unknown." Since students who were still enrolled may have later received their certificate of completion, or earned an alternative certificate, yet did not receive a diploma, they were also considered dropouts for the purposes of this study. Including these additional categories in drop-out statistics was consistent with research indicating that some students are pushed out, while others are pulled out (Doll, Eslami, and Walters, 2013), and still others opt out (Schubert, 2009).

GPA. Students' GPA was calculated across all credit-bearing courses taken in their ninth-grade year. Previous research indicates that ninth-grade GPA is a strong predictor of future on-time graduation (Allensworth & Easton, 2007; Johnson & Semmelroth, 2010; MacIver & Messel 2012; Stuit et al., 2016).

Credits earned. In high school, students earn credits for each class in which they receive a passing grade. Since failed courses result in unearned credits, credits earned served as a proxy for course failures. Course failures are directly tied to graduation because students need to accumulate a minimum number of credits to graduate. Previous research indicates that course failures and unearned credits give indications of students' lack of progress toward on-time graduation (Allensworth & Easton, 2007).

Over-age status. The typical age of students upon entry to ninth grade is 14, therefore, students age 15 and older are considered over-age. For the purpose of this study, over-age status served as a proxy for grade retention (Allensworth & Easton, 2007). It was measured in years and calculated by subtracting students' birth year from 2009, the year of the survey. According to MacIver and Messel (2013), this was a conservative representation. It included some students whose parents may have delayed entry into kindergarten. It excluded some students who may have been retained. For example, students who entered kindergarten a few months shy of their fifth birthdays, if retained, would have been a few months shy of 15 when they entered ninth grade. These retained students would not be captured by the variable over-age status.

Sensitivity. Sensitivity referred to the precision of the early warning indicator to identify true dropouts. It was the true-positive proportion, the number of true-positives divided by the total number of actual dropouts (Bowers et al., 2013). For example, among a sample of 10 students for whom their true graduation status was known, three were dropouts and seven were graduates. If a predictor of GPA = 0.5 correctly identified 2 as dropouts, then the sensitivity was 2/3 or 0.67. Sensitivity represented the probability that a student with GPA less than or equal to 0.5 will be correctly identified as a dropout (Gönen, 2007).

Specificity. Specificity referred to the precision of an early warning indicator to identify true graduates. It was the true-negative proportion, the number of true-negatives divided by the total number of graduates (Bowers et al., 2013). Continuing with the example above, if GPA = 0.5 correctly identified 6 of the 7 graduates then the specificity was 6/7 or 0.86. The specificity represented the probability that a student with GPA greater than 0.5 will be correctly identified as a graduate (Gönen, 2007).

1-Specificity. This term referred to the error in the predictive value of an early warning indicator. It was the false-positive proportion, the number of false-positives divided by the total number of graduates (Bowers et al., 2013). It was also referred to as the false-positive rate (Gönen, 2007). Continuing with the previous example, if GPA = 0.5 identified one student as a dropout whose known status was a graduate, then the 1-specificity was 1/7 or 0.14. It was the probability that a student with GPA less than or equal to 0.5 will be misclassified as dropout. In this example, there was a 14% probability that GPA will identify a student who has a GPA less than or equal to 0.5 as a dropout but who will likely graduate.

Receiver operating characteristic (ROC) empirical curve. The ROC empirical curve was drawn on a graph whose axes were 1-specificity (x-axis) and sensitivity (y-axis). It was a graph of “hits” versus “misses” (Bowers et al., 2013) at each possible cut point. It allowed the researcher to analyze the tradeoff at each threshold value of the indicator. Continuing with the above example in which GPA was being analyzed as a predictor variable in a sample of 10 students whose true graduation status was known, ROC analysis began with ranking students by GPA. Then the 1-specificty and sensitivity was calculated at each threshold value of GPA and the results were plotted on the ROC graph. Figure 2 contains the plot of 1-specificity and sensitivity at each of the 10 threshold values. The threshold value of GPA=0.5, with 1-specificity

(false positive rate) of 0.14 and the sensitivity (true positive rate) of 0.67, is designated with an arrow. A good indicator will produce a curve well above the diagonal chance line. The area under the curve (AUC) represented the probability of GPA correctly distinguishing dropouts from graduates. In this example the AUC was 0.86, 86% of the dropouts will have a lower GPA than graduates. The empirical curve was further analyzed for cut scores that optimized decision criteria. For example, the cut-score that was farthest from chance is indicated by the dashed circle in Figure 2.

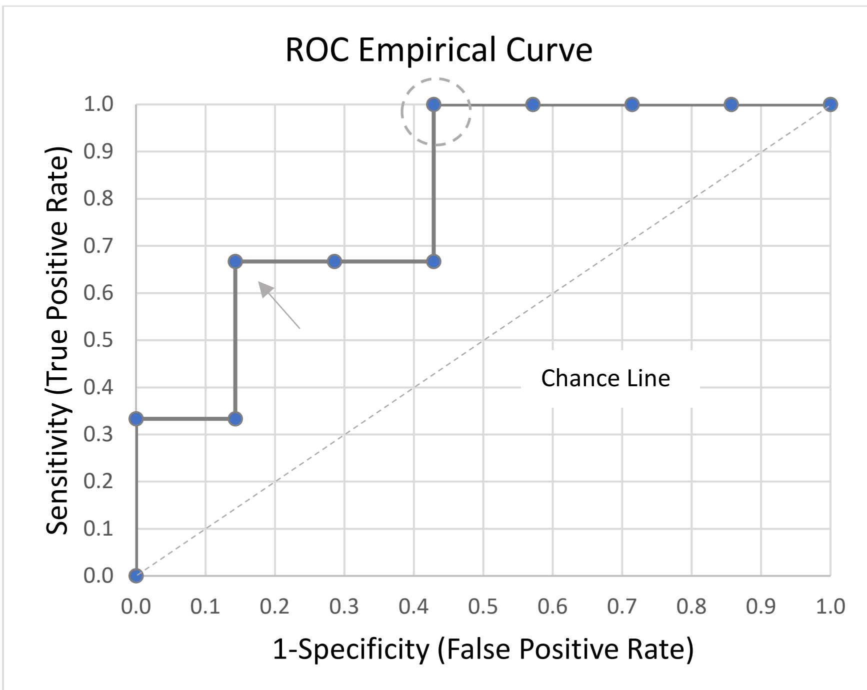

Figure 2: Example of ROC Empirical Curve

Limitations. There were potential limitations due to two sources of missing data. The National Center for Education Statistics High School Longitudinal Survey 2009 (HSLS:09) expended extra effort to realize the nationally-representative sample target of 800 schools (Ingles et al., 2011). However, of the 1,899 study-eligible schools, 945 refused to participate and additional 197 schools rescinded their participation after initial agreement. Of the 197 schools, 44 converted to participation after a personal visit from a HSLS:09 representative. The reasons

given for refusing to participate included concerns about the extra work for staff, the loss of instructional time, and over-testing of students. Forty percent of refusing schools cited time constraints and being too busy to take on the additional tasks required by the study. Thirty-one percent gave the general refusal of “Don’t want to participate.” While it was beyond the scope of this study to analyze these refusal reasons and rates, it is important to note that students from 50% of the eligible schools are not represented in the data. The result of this study were potentially limited to schools willing to participate in the study.

A second source of missing data came from the time lapse between receiving the ninth-grade student enrollment lists in early September and the scheduled session for students to take the survey. One thousand of the study-eligible students on the lists were not enrolled in school on the day of the student survey session (Ingels et al., 2011). These students were on the enrollment books but had never attended the school or had left the school prior to the survey session. This study was potentially limited to ninth-grade students enrolled in school at the time the student survey was administered.

Summary

Early warning indicators measured at the end of ninth grade offer information useful for identifying students likely to benefit from dropout intervention strategies. Prior research has identified GPA, credits earned, and over-age status as indicators that predict the likelihood of students dropping out of high school. Additionally, these three indicators are readily available to school personnel and open to the influence of school-base intervention strategies. This study applied receiver operating characteristic curves on data from the National Center of Education Statistics High School Longitudinal Survey 2009, to assess the accuracy of the indicators measured at the end of ninth grade, to predict students who later dropped out of high school. The

ROC curves were used to identify threshold values that optimized common decision criteria. Every day teachers use a variety of assessments to identify students in need of additional instruction and support. This study offered ROC analysis as an additional tool for teachers to use as a first step in a dropout intervention process.

CHAPTER 2

Literature Review

Students who decide to leave school before earning a diploma are called dropouts; however, "… the label never captures the complexity of the biography" (Cameron, 2012, p. xxv). While each student's decision is the result of a unique and complex variety of factors and influences, the first section of this chapter examines the dimensions of the dropout phenomena to provide a conceptual framework for this research. The critical importance, as described in the literature, of examining warning signs in students' transition to high school, typically measured at the end of their ninth-grade year, is discussed. Then, research regarding the development and predictive value of early warning indicators in general is reviewed, followed by an individual examination of GPA, credits earned, and over-age status indicators. This first section builds support for the central thesis that early warning indicators, specifically GPA, credits earned, and over-age status, measured at the end of ninth grade offer important information regarding students' progress toward on-time graduation four years later.

The second section of this chapter discusses the literature on the technical aspects of predicting dropout. Current educational research using ROC analysis is reviewed. Then the utility of ROC curves in assessing the accuracy of predictions is examined. An example of calculating the ROC is provided as a visual reference.

After support for early warning indicators is built and prediction techniques are examined, some other factors that influence students' decisions to leave high school before earning a diploma are reviewed. Students attend schools and live in communities that exert contextual pressures. Possible school and community level factors are examined as a reminder

that dropout decisions do not rest entirely on students' academic performance, instead ninth-grade indicators provide early warning signs of underlying influences.

Conceptual Framework

Students at risk of dropping out of high school often display identifiable warning signs in ninth grade. These early warning indicators, like red flags, signal potential danger, suggesting a student might be headed off the path to graduation. Three particular warning signs tied to students' academic performance and examined in this research were GPA, credits earned, and over-age status. They gave evidence of the quality, quantity, and progress of students' academic engagement.

Dropping out is defined as a student's decision to leave high school before completing the requirements for earning a diploma. Rumberger (2011) discusses three ways to view dropping out: status, event, and process. Dropping out as a current status means an individual may be considered a dropout today, but if that student decides to reenroll, his or her status would change. Dropping out as an event notes the moment in time a student leaves, either formally or by non-attendance. Dropping out as a process recognizes the ways a student's pattern of non-attendance, academic struggles, social difficulties, and behavioral problems accumulate over time. Each of these views reveal something important about the phenomenon.

Rumberger's dimensions of dropping out give a helpful framework for evaluating the literature. Viewing dropping out as a status leads to research into how the dropout status affects groups of students. For example, Sterns and Glennie (2006) examined reasons for dropout rates for high school students across grade level, age, gender, and ethnic groups. They analyzed a cross-section of data from students who dropped out of North Carolina public schools in 1998-1999. Sterns and Glennie classified the students' reasons for dropping out into six categories:

academic, disciplinary, employment, family, moving, and attendance. They found that academic reasons for dropping out increased by grade level. Older students tended to leave for academic reasons, whereas younger students tended to leave for disciplinary reasons. African American males were more likely to be pushed out of school by disciplinary actions, while Latina females were more likely to be pulled out by family reasons.

Viewing dropping out as an event leads to research into the timing of the dropout decision. For example, Batin-Pearson et al. (2000) compared five theories predicting students' early decisions to drop out (before 10th grade). They found that poor academic achievement fully mediated the other variables of deviance, low school bonding, antisocial peers, sexual involvement, low parental expectations, a parent's lack of education, gender, ethnicity, and socioeconomic status for early dropouts.

Viewing dropping out as a process leads to research into attitudes, behaviors, and academic performance that precede the decision to quit school. For example, Jimerson et al., (2000) analyzed longitudinal data collected on 177 children in Minneapolis. The first data were collected when the children's mothers were in their third trimester of pregnancy. The last data were collected when the children turned 19. They found that parent involvement, problem behaviors in the first grade, peer competence, problem behaviors when the child was 16, quality of caregiving, and academic achievements in first grade and at age 16 correctly classified 78 percent of the dropouts and 77 percent of the graduates.

While these three dimensions of dropping out are not mutually exclusive, my research focused primarily on the view that a student's decision to drop out of high school is the culmination of a process. Specifically, it examined academic performance indicators that precede students' decisions to drop out. Academic indicators, unlike demographics, attitudes, and

behaviors, are more open to the influence of school-base practices. For as Rumberger and Lim (2008) state, the academic indicators offer guidance to schools on how to develop intervention strategies to get students back on track to graduate.

The turbulent transition year. Although children begin their learning journey at birth, on the first day of kindergarten, they step onto their academic path to high school graduation. Knowledge gained, experiences gathered, and social interactions encountered cumulate over each school year and shape the attitudes, behaviors and achievements students bring to the next. In this way, each school year provides a snapshot of students' learning progress along a trajectory. There are points along this path, according to Neild, Stoner-Eby, and Furstenberg (2008), where students' successful completion of high school hangs in the balance and their educational trajectories can be reshaped. Ninth grade is one such point. Sometimes referred to as the rocky transition year, ninth-grade experiences present potentially treacherous waters that can be difficult for students to navigate (Neild, Balfanz, & Herzog, 2007; Neild & Farley, 2004; Neild, Stoner-Eby, & Furstenberg, 2008).

In the American school system ninth grade marks the beginning of a new system in which students must earn credits toward graduation. For most, it coincides with the physical changes of early adolescence, reduction of parental support, and an increase in peer influence (Neild, 2009). Students accustomed to the structured environment of middle school can experience ninth-grade shock (Allensworth, 2013) as they transition from eighth grade, where their teachers and parents held them accountable for attending class and completing assignments, to ninth grade where they experience increased freedom and responsibility. Allensworth describes a research study in which she and her team interviewed 52 students attending public schools in Chicago during their eighth-to-ninth grade transition. Preliminary results showed that

students did not feel that ninth grade coursework was more difficult. Many felt it was easier. In fact, students indicated they put less effort into their work due to the dramatic decline in monitoring by their teachers and reduced availability of supports.

Achievement loss in the transition to high school was examined by Alspaugh (1998). He grouped 16 small town and rural school districts into three groups based on the structure of the district. The first group had one K-8 elementary school and one high school. The second group had one each of elementary, middle, and high schools. The third group had two to three elementary schools, one middle school, and one high school. The results showed significant achievement loss for students transitioning from fifth to sixth grade (middle school transition) and transitioning from eighth to ninth grade (high school transition). Additionally, the dropout rate was significantly different between the districts with no middle school and the districts with a middle school. Alspaugh suggests that the instability and adjustments required of students when experiencing school-to-school transitions might be associated with increased dropout rates.

Recent research by Benner, Boyle, and Bakhriari (2017) supports this decline in student performance and engagement in the transition to ninth grade. They analyzed data from a longitudinal study of students in two ethnic minority-concentrated schools in a metropolitan area in the South. They used analysis of variance, repeated measures of covariance, and path analyses to document individual changes from eighth grade to ninth grade. The results showed that students' course grades declined significantly and their feelings of loneliness increased significantly over the transition period. However, school belonging appeared to play a buffering role, influencing a positive transition in relation to depressive symptoms, loneliness, school engagement, and grades. Additionally, Pharris-Ciurej, Hirschman, and Willhoft (2012) examined four cohorts of students enrolled in a West Coast metropolitan school district. Many students

who did relatively well in eighth grade failed courses in ninth grade. At the beginning of their 10th grade year, 14 percent were reenrolled in ninth grade and 14 percent had dropped out.

Some students enter ninth grade with skills below grade level. Others do not understand they must accumulate credits to graduate. Some have difficulty navigating the new school building with its new social relationships, school practices and routines, and a new teacher each class period. Add to the mix that freshman level classes tend to be taught by less experienced, newly-certified teachers. The result creates a turbulent transition for many students (Neild, 2009; Neild, Stoner-Eby, & Furstenberg, 2008). How well students weather the transition provides important information regarding their trajectory toward successful high school outcomes. Regardless of student performance in earlier grades, student work in ninth grade gives the strongest indication of their graduation outcome (Allensworth & Easton, 2007; Kemple, Segeritz, & Stephenson, 2013).

Early Warning Indicators. Enroll, attend, progress, graduate, these are the steps on the path to graduation (Rumberger, 2011). However, not all students follow the path. Some are pushed out for consequences of bad behavior, pulled out by employment, or fall out through disengagement (Doll, Eslami, & Walters, 2013). Others drop out because it is a matter of survival, a teacher encourages them to get their GED, the rules are too strict, or the pull of real life outside school makes it impossible to stay (Cameron, 2012). Regardless of the causes of or reasons for dropping out, there are factors associated with students' academic achievement (grades), educational persistence, and educational attainment (earning credits and being promoted to the next grade level) that predict graduation outcomes (Rumberger & Lim, 2008). These predictors are indicators associated with higher probabilities of dropping out (Carl, Richardson, Cheng, Kim, & Meyer, 2013).

Miller, Luppescu, Gladden, and Easton (1999) first developed the On-Track Indicator to describe ninth-grade students in a Chicago school district who graduated four years later. After analyzing student high school records for graduates in the years 1993-1997, they designated students as On-Track if they received no more than one F in a core course during the year and had earned enough credits to move to the next grade level on time. They found that being On-Track correlated significantly with graduation. Later in a study analyzing student records for six cohorts of students in Chicago public high schools in 1993-2000, Miller, Allensworth, and Kochanek (2002) found that 78 percent of On-Track students graduated four years later. Only 16 percent of Off-Track students graduated.

In succeeding years the Consortium on Chicago School Research investigated the predictive strength of the On-Track Indicator as well as other ninth-grade indicators. Allensworth and Easton (2005) found that students On-Track at the end of their freshman year were four times more likely to graduate than their Off-Track peers. However, given concerns that the On-Track Indicator, which is calculated at the end of the ninth-grade year, gives information too late to be useful, Allensworth and Easton (2007) compared the On-Track Indicator to other individual indicators (GPA, course failures, absences, and over age) measured at the end of the first quarter of ninth grade. They found GPA and course failures to be just as predictive as the On-Track Indicator, correctly identifying graduates versus non-graduates 80 percent of the time. Specifically, GPA was the most accurate in predicting dropouts. Allensworth (2013) reports that Chicago Public Schools is using these ninth-grade early warning indicators to develop strategies to improve student performance and lead more students in making progress toward graduation.

Similar to the research done by the consortium in Chicago, MacIver and Messel (2012) analyzed both eighth-grade and ninth-grade predictors for students in Baltimore City Public

Schools. Using the term early warning indicators, their predictors included attendance, behavior problems, and course failures. Specifically, they reported that 92 percent of the dropouts in the 2004-05 cohort had at least one early warning indicator in the ninth grade. In subsequent research MacIver and Messel (2013) used logistic regression to analyze a sample of 84,000 students in Baltimore City high schools. In addition to attendance, behavior problems, and course failures, they included over-age for grade, and GPA. Chronic absenteeism and course failures were stronger predictors than suspensions or demographic variables. Being male and over-age was a significant predictor of non-graduation when controlling for behavior. All ninth-grade indicators were stronger predictors than the eighth-grade indicators.

Referencing the work done by the consortium in Chicago, Kemple, Segeritz, and Stephenson (2013) developed an on-track indicator for a sample of 576,000 students in New York City Schools. The indicator included total attendance, number of credits, number of failed courses, and passing at least one end-of-course Regents Exam required for graduation. They found that credits earned, passing at least one Regents Exam, and attendance gave the best balance of correctly distinguishing between on-track and off-track students. It correctly identified 82 percent of the on-track students and 76 percent of the off-track students. Similarly, Carl et al. (2013) developed an indicator for analyzing a sample of 80,000 students in Milwaukee Public Schools. Looking to find an indicator that would convey information about the academic quality in addition to the quantity of credits, they defined Total Quality Credits (TQC) as a linear combination of credits earned and grades. They found that TQC correctly identified 85% of graduates.

The common threads in each set of indicators examined by these research studies are GPA, credits earned, and over-age for grade. This particular group of factors became the focus of

this research. Each gives information regarding a dimension of students' academic performance and the trajectory of future outcomes.

GPA. Grades are the symbols teachers assign to represent a composite measure of student performance in a course of study (Brookhart et al., 2016). Teacher-assigned grades in all courses are averaged and represented by a cumulative GPA. The literature describes grades as representing more than an assessment of students' academic knowledge and skills. Grades contain two components: academic knowledge and classroom achievement (Bowers, 2009). Grades take into account individual differences such as student effort and interest (Thorsen & Cliffordson, 2012). They represent how well students have fulfilled an implicit contract between the teacher and students (Willingham, Pollack, & Lewis, 2002). Grades emerge from an interactive process of teaching and learning that occurs daily between the teacher, the course material, and the student (Maué, 2016).

While most teachers view grades as something that students earn, a type of currency, or compensation for work completed (Brookhart, 1993), they often base their assignment of grades on a variety of factors. McMillan (2001) analyzed survey responses of 2,293 classroom teachers in middle and high schools in Virginia. Teachers reported using academic performance, performance compared to a set scale of percent correct, and specific learning objectives as the top three factors in determining grades. They also reported using, to some extent, the factors of student effort, ability level, quality of homework, and the degree of attention and participation. These teachers viewed grading as part of a larger philosophy of teaching and learning in which they made accommodations for individual student differences. Grading on effort was seen as a way to motivate students.

Effort, attendance, participation, and interest constitute affective classroom behavioral factors that teachers consider when assessing students and give grades multi-dimensionality. It is this dimensionality, Bowers (2011) suggests, that accounts for the difference between grades and test scores. In his analysis of data from the Education Longitudinal Study of 2002, Bowers compared grades to math and reading standardized test scores. He found that the non-cognitive classroom behavior dimension accounted for the differences in grades and test scores. The academic knowledge dimension accounted for the differences in how grades in core subjects were more closely associated with test scores than grades in noncore subjects. More than academic knowledge, grades are an assessment of students' overall performance across a range of classroom expectations.

> While teacher-assigned grades are not a pure assessment of academic knowledge, it appears that teachers may be adept at assessing a student's ability to perform at the social processes of the institution of schooling, in which academic knowledge is just one component of a much broader array of behaviors required by a student's community and school. (Bowers, 2011, p. 153)

Since successful graduation requires both academic knowledge and persistent engagement in the processes of school, this dimensionality of grades may account for their strong association with future educational outcomes, such has graduating or dropping out.

Regardless of the variation in teachers' inclusion and weighting of additional factors in their calculation of student grades (McMillan, 2001), teacher-assigned grades are strong predictors of future educational outcomes. Grades predict future academic success (Thorsen, 2014). They predict high school completion (Bowers, 2010) and higher education achievement (Cliffordson, 2008). Research shows that grades are strong predictors of life outcomes such as

wages, welfare, higher education, physical and mental health, and voting. (Borghans, Golsteyn, Heckman, & Humphries, 2016). Although Allen (2005) argues that grades are invalid assessments of student achievement because teachers include nonacademic factors, they are valid as assessment of students' overall schooling (Bowers, 2011) and as such give a strong indication as to students' future academic outcomes.

Credits earned. In American high schools students must earn a minimum number of credits to graduate. Credits are based on Carnegie Units, in which one unit represents a single subject taught for one classroom period for five days a week (U.S. Department of Education, 2008). The required minimum number of credits varies by state. California requires a minimum of 13 credits (local school boards may establish additional requirements); Texas requires 26 (National Center for Education Statistics, 2013a). Earning credits demonstrates students' progress toward graduation. Failing one course may signal a student's academic difficulties with the subject or the teacher, but multiple failed courses may signal a student's disengagement with school (Roderick & Camburn, 1999).

Neild, Stoner-Eby, and Furstenberg (2008) analyzed student data from a cohort of Philadelphia public schools. The results showed 46 percent of the dropouts were listed as 9th or 10th graders despite the fact that they had been enrolled in school for several years. The majority were behind in credits; 88 percent had earned no more than three credits during their time in high school. Logistic regression analysis results showed that a 20-percentage point increase in the number of failed courses (the equivalent of one extra failed course in semester) would increase the odds of dropping out by 40 percent. Roderick and Camburn (1999) analyzed data from students in Chicago Public schools. Their results showed the estimated probability of failing at least one course was 0.55 for a student who entered ninth grade two or more years below grade

level in reading and mathematics. Interestingly, they report that course failure did not appear to be limited to only those students who entered high school with low skills, the estimated probability of failing at least one course in ninth grade was 0.31 for a student who entered with grade-level skills.

Course failures accounted for the greatest difference of 10 percentage points between graduates and non-graduates in Robison, Jaggers, Rhodes, Blackmon, and Church's (2017) analysis of student data from the Louisiana Department of Education. In Allensworth and Easton's (2007) analysis of sample of Chicago Public High Schools, 53 percent of the students failed at least one course in their freshman year and fall semester course failures correctly identified 76 percent of dropouts. Pharris-Ciurej, Hirschman, and Willhoft (2012) report that more than a third of the sample of ninth graders began their second semester with a GPA less than 2.0, which indicated failing or almost failing one or more classes. When students fail courses, they do not earn credits. Thus, credits earned and courses failed can be conceptualized as two sides of the same dimension: students' freshman year educational attainment. Both indicate whether or not students have passed enough of their classes to earn enough credits to keep them on the path to graduation.

Over-age status. Being over-age is a term used to describe students who are older than their classmates. Students can be over-age in ninth grade for a variety of reasons. They may have entered Kindergarten older than their peers or they may have repeated an elementary or middle school grade level. While most parents enroll their students in Kindergarten at age five, some choose to delay schooling for their children. The prevalence of academic redshirting, or the delaying of entrance to Kindergarten, was four to five percent nationwide in 2006, Bassok and Reardon (2013) found in their analysis of the Early Childhood Longitudinal Study. Rates of

redshirting varied across gender, race and SES, with the likelihood of redshirting being higher among children whose birthdays fell in the months closest to the cutoff date, and among higher-SES families. The data suggests that the parents' decision to redshirt might be driven by concerns over their child's physical development rather than their cognitive or behavioral development.

Students unable to keep up with the rapid pace of calendar-driven schooling, who perform poorly in school, are often required to repeat a grade level in an effort to give them more time to master the curriculum (Martin, 2011). In 2015, 2.2 percent of the students in the U.S. were retained, or enrolled in the same grade level as the previous year (National Center of Education Statistics, 2015). While grade retention was designed to improve student learning by giving students additional time to catch up to grade level, Tingle, Schoeneberger, and Algozzine (2012), in their analysis of 125,000 students in the Southeast, found that year one and year two achievement of retained students was consistently below that of their promoted peers. Moller, Stearns, Blau, and Land (2006) in their analysis of the National Education Longitudinal Study 1988-1992 found that retention predicts students' academic trajectories. Students retained prior to eighth grade had initial math and reading achievement scores five points below their promoted peers and experienced fewer gains in achievement. Retention accounted for 23 percent of the variation in achievement and growth between the two groups of students. Grade retention at any stage of schooling elevates the risk of dropping out (Alexander, Entwisle, & Kabbani, 2001).

Grade retention is associated with negative effects on students. Students as early as third grade rated retention in their top five stressful life events; for sixth-grade students, it was rated as the top stressor (Anderson, Jimerson, & Whipple, 2005). Jimerson and Ferguson (2007) found that retained students were five times more likely to drop out of high school relative to students

who were recommended for retention but were promoted. Grade retention was a significant negative predictor of academic self-concept, academic motivation, academic engagement, general self-esteem, and homework completion in Martin's (2011) analysis of data from a sample of 3,261 students in grades 7 thru 12 in Australian schools. He examined the relationship between grade retention and academic and nonacademic measures. Grade retention was a significant positive predictor of maladaptive motivation and weeks absent from school.

Regardless of the reason for being one or more years older than their grade-level classmates, over-age students are more likely to drop out of school (Allensworth & Easton, 2007; Lee & Burkam, 2003; MacIver & Messel, 2012; Roderick 1994). Each additional year older a student is upon entering ninth grade doubles that student's odds of dropping out (Neild, Stoner-Eby, & Furstenberg, 2008). Sterns and Glennie (2006) found that 25 percent of dropouts with ninth grade standing were age 17 or older in a sample of North Carolina schools. Older students tended to get pulled out of high school by employment opportunities (Stearns & Glennie, 2006). Students entering ninth grade are entering a new social environment in which being older than their peers becomes more evident and important. Fifteen-year-old freshmen will be 19 if they graduate on time. If they take an extra year, they will be 20, the same age as their peers who will have completed two years of college or work experience (Neild, Balfanz, & Herzog, 2007). According to Roderick (1994) they may feel like failures and negatively self-conscious as they compare themselves with their peers. "…[B]eing over-age for grade, no matter when a youth was retained, has an impact on attitudes toward, and experiences in, school that may not be reflected in grades or attendance" (Roderick, 1994, p. 742). This impact of being over-age on students' attitudes toward school may add to the strength of over-age status in predicting future graduation outcomes.

In conclusion, identifiable warning signs displayed during students' turbulent transition year to high school offer early indicators of graduation trajectories. GPA, credits earned and over-age status are three indicators that predict the likelihood that students will dropout. Academic achievement and classroom performance, as represented by teacher-assigned grades, give an indication of students who are on-track to graduate. Passing classes and earning credits indicate students' progress toward their graduation goal. Being over-age for their grade level may provide additional information of students' past academic achievements. The predictive strength of these three indicators was analyzed in this study using ROC analysis. By using predictor variables, I hoped to provide to school personnel information regarding the severity of risk for students being off-track to graduation or to show that students are off-track in some areas but on-track in others (Kemple, Segeritz, & Stephenson, 2013).

Predicting Dropout

Accurate indicators correctly identify students at-risk of dropping out while not misidentifying as potential dropouts those students who will graduate. Scarce school resources drive the need for screening tools that target students most in need of interventions. Considering the potential for self-fulling prophesy, the work to accurately label students as at-risk must be done with as high a level of confidence as possible. One way to conceptualize early warning indicators is as diagnostic tests for dropping out. Similar to how medical personnel screen patients for the presence of a particular disease, GPA, credits earned, and over-age status could be viewed as diagnostic tests with the potential to predict the likelihood of dropout for students. When using early warning indicators to screen for possible at-risk students, school personnel would find information regarding their accuracy of the indicators useful.

Receiver operating characteristic (ROC). ROC analysis in educational research is relatively new. It is typically found in medical research where it is used to assess the accuracy of diagnostic tests. Since diagnostic tests are predictions of whether a patient has or does not have a disease, a ROC curve provides a statistical assessment of the accuracy of the prediction (Gönen, 2007). The curve gives a visual representation of the tradeoff between true positives and false positives at all values of the diagnostic test, thus giving medical decision makers a tool to set optimal threshold values (Fawcett, 2006).

Recently a few education researchers have conceptualized early warning indicators as diagnostic tests of student outcomes and have applied ROC analysis to assess the predictive value of both the individual predictors and logistic regression models of multiple predictors. Fewer researchers have extended the ROC analysis to determine optimal threshold values for the indicators. Munoz-Repiso and Tejedor (2012), McCluckie (2014), and Vivo and Franco (2008) employed ROC analysis to analyze the accuracy of first-year university student indicators. Munoz-Repiso used the area under the curve (AUC) to compare the strength of each indicator in predicting technology use. McCluckie, and Vivo and Franco used ROC curves to establish cut-scores for indicators predicting students' future academic success.

Liao, Yao, Chien, Cheng, and Hsieh (2014) and Jordan, et al. (2010) used ROC analysis on pre-school and kindergarten indicators. Liao et al. evaluated the ROC curve for optimal scores on a checklist used to predict developmental delays in preschool children in Taipei City. Instead of assessing the accuracy of a single cut point, ROC analysis allowed Liao et al. to provide a decision validity index on multiple cut scores. Jordan et al. assessed the accuracy of a Kindergarten Number Sense Screening Tool in predicting mathematics proficiency in the third

grade. They also used ROC establish diagnostic cut scores for predicting students likely to show mathematic weakness in third grade.

Of particular interest are the studies conducted by Nicholls, Wolfe, Besterfield-Sacre, and Shuman (2010), Stuit et al. (2016), and Johnson and Semmelroth (2010). Similar to my research, Nicholls et al. analyzed a complex data set to evaluate eighth-grade predictors of future educational outcomes. Using data from the National Educational Longitudinal Study of 1988 (NELS:88), they built two statistical models and then used ROC analysis to compare their predictive accuracy. They defined an integrated modeling approach that combined logistic regression, survival analysis, and sensitivity analysis in a step-wise progression. Starting with logistic regression, the outcome of one step was used as the input for the next step. They hypothesized that this integrated modeling approach would produce more accurate predictions than any single standard statistical approach. They began with 76 potential eight-grade variables and built a logistic regression model on the 22 significant predictors of a student’s future attainment of Science, Technology, Engineering, or Mathematics (STEM) university degrees. Next, they put the logistic regression model through the subsequent steps to produce an integrated model. Using ROC analysis to compare the two models, Nicholls et al. found that the logistic regression model alone produced more accurate and beneficial results than the integrated approach. They concluded the logistic regression model’s strength was due to the comprehensive nature of the complex data set, the large number of records, and the care taken in selecting and coding the variables.

In contrast, Stuit et al. (2016) and Johnson and Semmelroth (2010) used ROC analysis on regional samples. Stuit et al. used data from two cohorts of eighth- and ninth-grade students in three Ohio school districts. Johnson and Semmelroth used ninth- and tenth-grade data from two

suburban schools in the Northwest. Both sets of researchers built logistic regressions models from early warning indicators including GPA, credits earned, and failing grades. They analyzed the predictive strength of the models and determined optimal cut scores. Johnson and Semmelroth also calculated the relative risk: the ratio of the risk of a student dropping out with the indicator to the risk of dropping out without the indicator. In both studies, the indicators of GPA, credits earned and course failures consistently predicted dropping out, but the optimal cut scores varied by school and district.

Assessing accuracy of predictions. ROC analysis is a statistical tool which accesses the accuracy of predictions. Gönen (2007) defines a diagnosis as a prediction based on symptoms of what might be wrong with a patient. Accuracy of diagnosis is critical because it influences future evaluations and treatment. Diagnostic options also vary in cost and risk to patients so it is helpful to have a way to compare the accuracy of various diagnostic tests. ROC curves present a visual summary of the accuracy of predictions.

Originally developed during World War II to analyze the accuracy of detection radar in classifying signal from noise, ROC curves have been adapted for clinical use for evaluating diagnostic tests (Zou, O'Malley, & Mauri, 2007). They are widely recognized as a meaningful approach in quantifying the accuracy of diagnostic tests and decisions, according to Metz and Pan (1999). Fawcett (2006) describes the ROC graph as a technique for visualizing, organizing, and selecting classifiers based on their performance in distinguishing hits from false alarms. In the ROC space, the true positive rate is graphed against the false positive rate at each threshold value of the classifier. The ROC graph represents the tradeoff between true positives and false positives, benefits and costs. Decision makers can analyze the tradeoffs and choose an optimal threshold based upon specified benefits versus costs criteria. If the benefits of receiving

treatment outweigh the costs, a medical professional might choose a cutoff value for the diagnostic test which will over identify patients to receive treatment. ROC curves can also be used to compare the accuracy of two different classifiers. The area under the curve (AUC) gives a single scaler value representing performance, equivalent to the probability that the classifier will rank a randomly chosen positive instance higher than a randomly chosen negative instance. (Fawcett, 2006).

In order to use ROC analysis, the true disease status for each patient must be measured without error, which provides a gold standard (Zou, O'Malley, & Mauri, 2007). When selecting a gold standard, Zou, O'Malley, and Mauri warn that there is the potential for two types of errors: verification bias and measurement error. Verification bias can arise when the accuracy of the test is evaluated only on patients with a known disease status. Measurement error may arise when a true gold standard is unavailable or is an imperfect standard for comparison. In the case of this research study, receipt or non-receipt of a high school diploma within the typical four-year time frame represented the gold standard. The individual early warning indicators represented the diagnostic tests. The gold standard, receipt of a high school diploma within four years, was verified by students' official transcript records. However, the results may be limited by verification error because the accuracy of the diagnostic tests were evaluated on students with known graduation outcomes. To overcome this limitation bootstrap resampling would need to be used to test whether the predictive value of the indicators is consistent when applied to new data from the underlying population (Gönen, 2007; Stuit et al., 2016).

There are several advantages to using ROC curve analysis, according to Pandey and Jain (2016). The plotting of the ROC curve does not require a predetermined cut point. Since the curve plots the sensitivity and specificity at all cut points, an optimal cut point can be chosen

based on the cost versus benefits criteria of decision makers. The ROC curve is independent of the scale of the test, it is a rank-based measure. AUC values can be obtained to discriminate between and compare the accuracy of different tests. Lastly, the visual representation facilitates ease of interpretation. The ROC curves will provide the means to choose optimal cut point values for each test and compare the predictive strength of all the tests.

Visualizing ROC analysis. ROC analysis is useful for visualizing and assessing classifiers or diagnostic tests. As an example of how a ROC curve is calculated consider the following fictitious illustration adapted from Houts (2016). Suppose IQ is to be used to classify individuals as on-track or off-track for graduation. In a sample of 10 student records which include IQ scores and known graduation status, three students dropped out and seven students graduated. The students are put in rank order by their IQ scores so that a cut value can be identified that will correctly identify those who dropped out (white in Figure 3) without incorrectly identifying those who graduated (black in Figure 3). Where the cut-point is placed will determine the accuracy of the prediction.

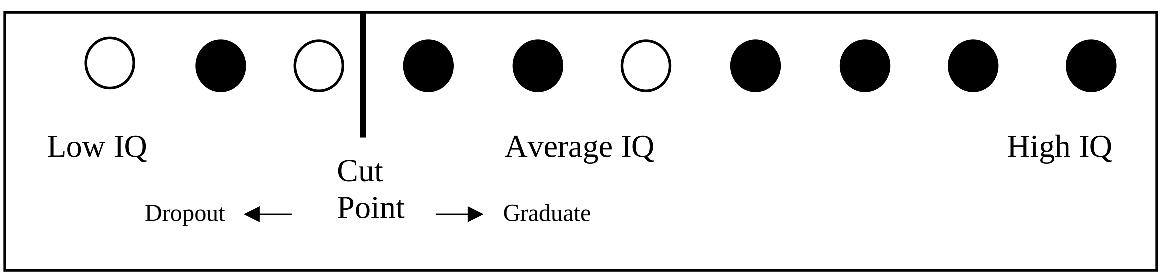

Figure 3: Cut Point for IQ to Classify Dropping Out of High School

If the cut-point is placed as indicated in Figure 3, then two dropouts and six graduates would be correctly classified. However, one graduate and one dropout would be incorrectly classified. For the cut point identified in Figure 3, the two-by-two matrix in Figure 4 is constructed. The matrix shows that the cut-point in Figure 3 correctly identifies two of the dropouts and six of the graduates, but it incorrectly identifies two individuals. The test classifies

as a dropout an individual whose true state is graduate, and it classifies as a graduate an individual whose true state is dropout. Using the matrix, ROC analysis extracts two values: the false positive rate and the true positive rate for the cut point. The false positive rate is the number of false positives divided by the total true state graduates. In Figure 4, the false positive rate is 1/7. The true positive rate is the number of true positives divided by the total true state dropouts. In Figure 4, the true positive rate is 2/3.

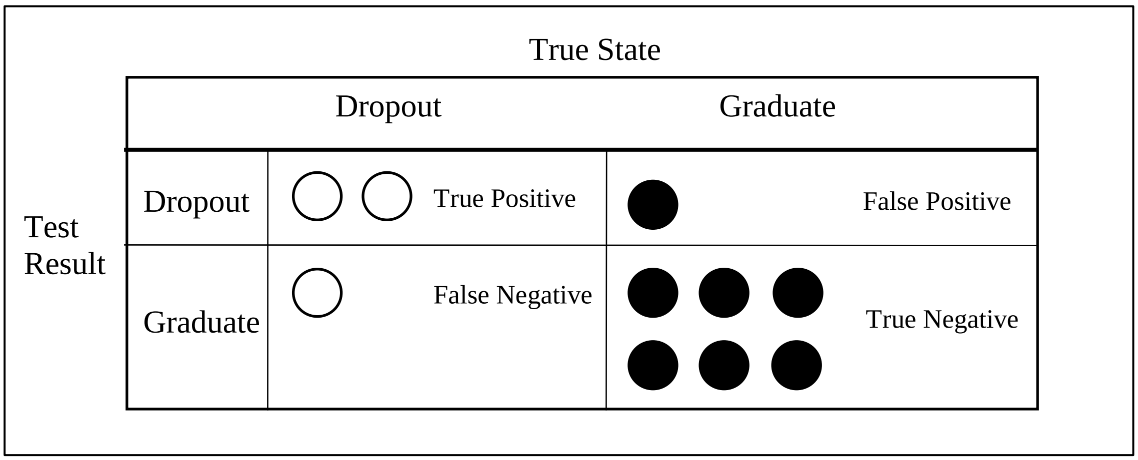

Figure 4: Two-by-two Matrix for the Cut Point in Figure 3

Next, consider a second cut point chosen at a higher IQ value than the first cut point such that it captures three additional individuals to the right of the first cut point. That second cut point value correctly classifies all three dropouts but captures an additional two false positives. This second cut point has a false positive rate of 3/7 and a true positive rate of 3/3.

The ROC graph in Figure 5 plots both of these cut points in the ROC space. The dotted diagonal line represents the line of a classifier that randomly guesses dropout or graduate status (Fawcett, 2006). In the ROC space classifiers above the chance line perform better than chance. Additionally, ROC analysis calculates the area under the curve (AUC). For the empirical curve in Figure 5 the AUC is 0.86. According to Houts (2016) our classifier is “Good.” Classifiers with AUC between 0.8 and 0.9 are “Good” and classifiers with AUC greater than 0.9 are “Excellent.” The AUC of 0.86 is interpreted as 86 percent of the time a dropout will have a lower IQ than a

graduate. In this illustration, the ROC curve is drawn for only two cut points. In a complete ROC analysis, the curve is drawn and the AUC is calculated using all possible cut points, and for this research, the SAS® statistical software calculated the weighted frequencies needed to calculate the standard error estimate and confidence interval for the AUC.

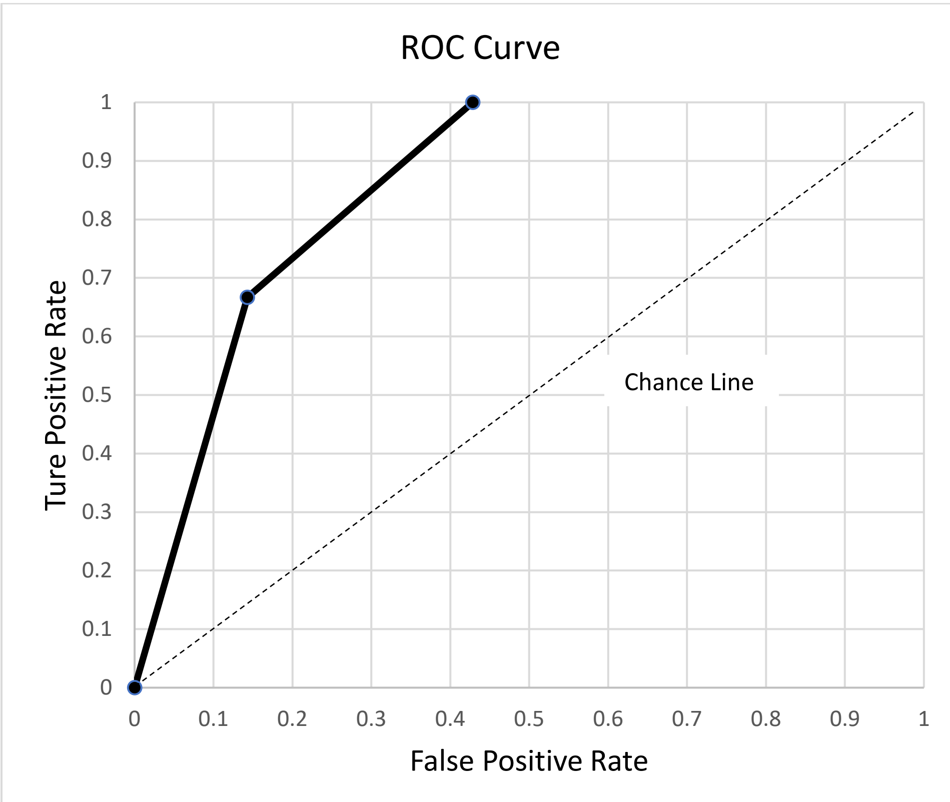

Figure 5: ROC Curve of Cut Points One and Two

Not only does the ROC analysis provide a visual for the classifier or diagnostic test, it also can be used by future decision makers interested in choosing an optimal cut point. This research study offered three typical cut point optimizations. A perfect classifier, one that correctly identifies dropouts and graduates with 100 percent accuracy, will have a true positive rate of one and a false positive rate of zero. It will be represented in the ROC space as the point (0,1). The cut point that is closest to the point (0,1), maximizes correct classification (Pandey & Jain, 2016). However, sometimes decision makers have other criteria to consider when choosing

an optimal cut point for a diagnostic test. A cost versus benefit assessment might lead a decision maker to choose a cut point with a greater sensitivity or true positive rate even if it means capturing some false positives. On the other hand, if sensitivity and specificity are equally important, a decision maker might choose a cut point that minimizes the difference between the two (Pandey & Jain, 2016). This research study offered three formulas for identifying cut points that maximized the distance from chance, minimized the distance to perfect prediction, and equalized sensitivity and specificity.

For my research, ROC analysis provided a visual assessment of the classifiers GPA, credits earned, and over-age status. Using AUC values the predictive accuracy of each classifier was assessed. The resulting ROC curves provided information regarding the classifiers' accuracy in correctly predicting the graduation status and for choosing cut points for the students sampled in 2009 as part of a national longitudinal survey.

In conclusion, the accuracy and predictive value of the three early warning indicators, GPA, credits earned, and over-age status, were assessed using ROC analysis. This visual representation of the tradeoff between true positive and false positives displayed threshold values for all cut points of a diagnostic test. It will allow future decision makers to choose a cut point that optimizes their cost versus benefits criteria. Using complex survey data from a nationally representative sample of ninth grade students enrolled in 2009-10 school year provided generalizable findings.

School and Community Influences

Identifying students at risk of dropping out was operationalized in my research study as three indicators measured during ninth-grade. GPA, credits earned, and over-age status are characteristics that give evidence of students who are potentially off the graduation track. Since

students must accumulate a minimum number of credits by earning passing grades in required courses and since students who do not earn enough credits for their grade level are retained, the three indicators provide information as to students' academic progress toward completing graduation requirements. Academic performance and engagement are student-level factors in which students are the agents, engaging in behavior that leads to their high school outcome. Examining student characteristics in isolation, without considering school and community level factors, frames dropping out as a student's personal decision, often based on a pattern of unwise decisions and low commitment to school (Lee & Burkam, 2003). It places the burden of graduation squarely on students' shoulders. However, the high school environment can push, and community expectations can pull students out of school. While students do make the final decision to cut short their high school journey, contextual factors can set the stage for eventual dropping out and sway teetering students either away from or toward their final dropout decision (Lessard et al., 2008).

Push factors. School context matters. To earn their diplomas students must successfully navigate the high school environment. They must follow the rules, relate to teachers, associate with peers, and complete assignments. Students need to participate. Yet, there are elements in the school environment, academic, disciplinary, and relational, that tend to push students out (Sterns & Glennie, 2006). School size, curriculum, and student-teacher relations either push out or hold in students at risk of dropping out (Lee & Burkam, 2003). Lee and Burkam, in their study of 190 high schools drawn from the National Educational Longitudinal study (NELS:88), found that large schools (student enrollment of 1,501 – 2,500) had the lowest SES, the lowest teacher-student relations, and the greatest proportion of dropout, when controlling for students' background characteristics. Yet schools, regardless of size, structured with a constrained

curriculum, defined as offering few low-level math classes or offering Calculus, had the lowest dropout rates. Academic organization of the school had a significant effect on dropout rates.

Students are pushed out when adverse situations in the school environment lead to consequences that result in dropping out (Doll, Eslami, & Walters, 2013). In an analysis of seven nationally representative studies from 1955 thru 2002, Doll, Eslami, and Walters (2013) found that beginning in the 1980's push factors were cited most often as dropout reasons. The top reason reported by 33.0% of students was poor grades. Not getting along with teachers (15.3%) or other students (5.6%) and being expelled or suspended (9.3%) were the other reported reasons for dropping out. These quantitative data give information regarding students' perspective on in-school factors preceding their dropping out.

School context mattered to the 12 Latino male dropouts in Halx and Ortiz's (2011) qualitative study. All wanted authentic relationships and conversation with school personnel. They cited this lack of personal connection as contributing to their disinterest in and hostile attitude toward school. School context was also cited as a factor in the decision to drop out for the 80 participants in Lessard's et al. (2008) research study. Some students reported sabotaging their educational journey by solving problems with their fists and turning away from school, but others reported suspension policies, teacher ostracism, and peer rejection for making them feel they did not belong in school. School context was also evident for participants in America's Promise Alliance dropout summits (Bridgeland, Legters, & Balfanz, 2010). The leading reason for leaving school cited by dropouts was not seeing a connection between classroom learning and real life. They spoke of sitting in classes learning what they would never use. Half were bored and stated classes were uninteresting.

Pull factors. Attending high school accounts for only part of students' day. When the dismissal bell rings, they step off the school grounds and into their communities. Family, friends, neighbors, and co-workers exert influence on students. Jimerson et al. (2000) describes a dynamic process in which background factors such as a student's SES, parents' education levels, peer relationships, and parental monitoring interact and build over time. This developmental process affects how students choose to engage, select, and interpret experiences based on previously established frameworks. Family context and the early caregiving environment can also shape whether students were on the path to high school completion or withdrawal. In Alexander, Entwisle, and Kabbani's (2001) life course perspective, home, school, and community are viewed as a social matrix shaping how children view themselves and how they enact their student role. Children living in these spheres of influence develop habits of conduct which influence school engagement or disengagement.

Lessard et al. (2008) view family context as setting the stage for students' educational journeys. Students who experienced family turmoil such as divorce, abuse, financial hardship, and increased mobility in elementary school faced increased challenges in secondary school due to the instability in their lives. One participant described how she was so preoccupied by family issues she could not focus on her student role. Family pulled her attention to such an extent she dropped out.

Pull factors can distract students away from school or attract them to the outside world (Doll, Eslami, & Walters, 2013). Financial worries, out-of-school employment, becoming pregnant, and marriage accounted for 37 percent of the dropouts in the Educational Longitudinal Study: 2002 (Doll, Eslami, & Walters, 2013). Similarly, Bradley and Renzulli (2011) report that pregnancy pulled out minority girls and marriage pulled out Latinas at greater rates than White

students. According to Ream and Rumberger (2008), family background such as low SES presented a cumulative resource disadvantage that may have deterred students from participating in homework and extracurricular school activities because they were obligated to work after-school jobs or care for siblings while parents worked. Additionally, friendship networks influenced students' educational decisions. Having at least one drop-out friend was associated with dropping out; students who reported having friends who valued education were more likely to complete high school on time (Ream & Rumberger, 2008).

Halx and Ortiz (2011) viewed the pull factors of economic pressures, personal emotional and family pressures, and social/cultural pressures as inputs that students must process in order to make sense of their educational journey. Students make meaning from the inputs and develop either a positive or negative sense of meaning for school as a whole, which in turn shapes their resiliency and desire to complete graduation requirements. Of the participants identified as dropouts and students on the brink of dropping out, none voiced a strong affection for or an intrinsic understanding of being an educated individual. Rather, they all expressed pride in their ability to work hard, even though they admitted that working interfered with school. "It was something everyone in their community did. They felt the need to work now, even if it meant they might be risking their future prosperity" (p. 426). Community pulled them in.

Students do not make their dropout decisions in isolation. Cameron (2012) fleshes out the push and pull influences in her qualitative study of dropouts. She interviewed seven participants and uncovered four themes driving their decisions to leave school: being known, being valued, purpose, and freedom. As students, these young people did not feel cared for or valued by their teachers, peers, or school staff. These former students lost sight of the purposes for staying in school. It was no longer relevant to them as the real world began pressing in with a good paying

job offer or a baby on the way. Lastly, Cameron's young people related experiences with oppressive restrictions and rules. As Ivan profoundly asks:

> [I]t is a miracle to be alive, it is a beautiful thing to be alive … then why am I filling my life, these prime years of my boyhood, … with this slavery, … sitting here in this ugly child factory? (p. 73)

For Cameron's dropouts the push and pull to leave was stronger than the value of receiving a diploma. High school did not entice them to stay.

These studies illustrate how academic performance is not the only force exerting pressure and influencing students. An adverse school environment can push students out. Feeling ignored by school personnel, unsupported by teachers, and unfairly treated for behavior problems contribute to student disengagement in school (Lessard et al., 2008). Family turmoil can pull students' attention away from fully participating in school. Work and marriage can pull away students who are eager to begin real life. While dropping out is the final step in a process of disengagement (Neild & Farley, 2004), students point to pivotal moments. These moments, which often include pregnancy, conflicts with teachers, automobile accidents, failing grades, and feeling burned out with balancing work and school, become instrumental in changing students' educational journeys (Lessard et al., 2008). For students teetering on the edge, these influences affected their academic performance and precipitated their decision to drop out.

Summary

Students exhibit identifiable warning signs in their ninth-grade year that may predict future dropping out. These academic indicators displayed in the transition year to high school give critical information regarding the trajectory of students' educational journeys. Research has shown that GPA, credits earned, and over-age status are predictive of students' eventual high

school outcome (Allensworth, 2013; Allensworth & Easton, 2005, 2007; Battin-Pearson et al., 2000; Bowers, 2010; Kemple et al., 2013). GPA, teacher assigned grades, measure academic knowledge and classroom achievement. They give an indication of how well students are learning the course content as well as navigating the classroom context (Bowers, 2010). Credits earned, which are tied to course failure, give information regarding students' progress toward fulfilling graduation requirements (Neild et al., 2008; Robison et al., 2017). Finally, being over-age, whether through grade retention or starting kindergarten comparatively late, is associated with dropping out (Alexander et al., 2001). Students older than their peers have additional challenges of negative self-concept and the draw of full-time work opportunities (Martin, 2011; Stearns & Glennie, 2006).

While research has shown that these three are good predictors of dropping out, few studies have examined the specificity and sensitivity of the predictors. Few examined the threshold values of each indicator. Decision makers in schools could benefit from knowing how low a GPA, how few credits, how many years over-age might rightly identify students likely to drop out without also misidentifying students likely to graduate. By viewing GPA, credits earned, and over-age status as diagnostic tests, ROC analysis gave a visual graph of the tradeoffs between sensitivity and specificity (Fawcett, 2006). Applying ROC analysis to data drawn from a nationally representative data set gave information regarding the accuracy of the indicators to predict dropping out and offered insight regarding the threshold values for the population of U.S. ninth-grade students in 2009.

The early warning indicators will likely give information regarding students' graduation status, but they also only represent one piece of the dropout puzzle. Dropping out is a complex process in which multiple factors are involved. Background characteristics and early childhood

experiences can set the stage for a difficult educational journey (Alexander et al., 2001; Lessard et al., 2008). School context can push, and family turmoil and community expectations can pull students who are teetering on the edge of dropping out (Bradley & Renzulli, 2011; Bridgeland et al., 2010; Doll et al., 2013; Halx & Ortiz, 2011; Lessard et al., 2008). Economic, family, and emotional pressures exert significant influence on students. Yet each of these factors are difficult to measure and typically beyond the scope of high school personnel. GPA, credits earned, and over-age status are readily accessible and open to the influence of school. This research could be used in the future by teachers and school staff to identify potential dropouts in a first step of an intervention strategy. Subsequent steps could address factors influencing students’ behavior and academic achievement. This research focused on the indicators for accurate and early identification of students in ninth grade.

CHAPTER 3

Methodology

This methods section discusses the research design of the National Center for Education Statistics High School Longitudinal Survey 2009 and collection of the data that this study utilized. The advantages, theoretical framework and alignment with the research questions are examined. The population, sampling, and participants of the HSLS:09 are described in detail. Instrumentation appropriateness and measurement characteristics are explained. National Center for Education Statistics administration details such as collection procedures, assignment of analytic weights, and data imputation are described. This section specifies the analytic techniques that were used to answer the two research questions:

1. How accurately do the ninth-grade indicators of GPA, credits earned, and over-age status predict the likelihood that students from the HSLS:09 dropped out of high school?
2. What are the threshold values for each indicator that optimize three criteria: maximum distance from chance, minimum distance to perfect prediction, and equality of sensitivity and specificity?

Coding of variables, handling of missing values, data download procedures, and SAS® and ROC analyses are explained in the analytics section. Lastly, the researcher’s role and ethics are considered.

Design

This research study was a quantitative secondary analysis of data collected by the National Center for Education Statistics in the High School Longitudinal Study of 2009

(HSLS:09) that assessed the predictive value of three ninth-grade early warning indicators (GPA, credits earned, and over-age status).

Theoretical framework. According to the documentation of the research (Ingles et al., 2015), the HSLS:09 was the fifth research study in the Secondary Longitudinal Studies program. The primary aim of the program was to collect statistics and other data on nationally representative samples related to education in the United States and to make the data available to researchers. The HSLS:09 design was guided by a theoretical framework that took the student as the fundamental unit of analysis and attempted to trace the influences that factor into their education-related goal setting and decision making. The addition of high school transcripts provided a longitudinal record of courses taken, credits, grades, and graduation outcome. The documentation states that primary use of HSLS:09 data is for secondary longitudinal analyses that focuses on the high school years. The documentation suggests several topics to be investigated including the process of dropping out and the educational trajectory of students. The transcript files were designed to be combined with the questionnaire data for analysis and the data files contain composite variables generated to summarize reports from the academic files and questionnaire responses.

Alignment of the research questions to the HSLS:09. The research goals for this study aligned with the HSLS:09 for several reasons. The HSLS:09 provided longitudinal data necessary to analyze ninth-grade predictors of graduation outcomes. The ninth-grade and transcript data collections provided information on the variables of interest: GPA, credits earned, over-age status, graduate, and dropout. An appropriate sample size of 25,206 students was determined using power calculations increasing the probability of finding significant relationships if they exist (Newton & Rudestam, 1999). It provided data on graduation outcomes

for students who transferred or dropped out, which can be difficult to obtain. Lastly, the data was accessible for public use on the NCES website.

Sampling & Participants

HSLS:09 used a complex sampling design. The HSLS:09 documentation states that the Base Year surveyed, in the fall of the 2009-2010 school year, a random sample of 25,206 ninth-grade students from 944 public and private high schools in the United States (Ingles et al., 2011). Students, the sampling units, were selected by two-stage stratified cluster sampling design.

Advantages. This data set had several advantages. First, HSLS:09 provided a readily-accessible and nationally-representative sample of students. (Ingles et al., 2011). Students were surveyed in the fall of their ninth-grade year to capture information about their transition to high school. These ninth-grade students were the sole cohort across all rounds of data collection. Students were the unit of analysis and the primary use of the data was for longitudinal analysis that will focus on the high school years (Ingles et al., 2015). Data was collected from dropouts and transferred students as well. The transcript files were collected in 2013-14, the school year following the on-time graduation date for the cohort. The large sample and the longitudinal design allowed for analysis of precursors to outcomes (Greenhoot & Dowsett, 2012).

Second, using data from the complex sampling design of the HSLS:09 resulted in findings that may be generalized to the target population of U.S. ninth-grade students enrolled in study-eligible schools in Fall 2009. Included in the data set were sets of analytic weights constructed using the Balance Repeated Replicates method that were necessary to account for bias correction (Ingles et al., 2011). The weights accounted for oversampling, the unequal probabilities of student selection, and unit school and student non-response rates.

Sampling. The 2009 base year documentation (Ingels et al., 2011) described the sampling design. The HSLS:09 utilized a two-stage stratified cluster sampling design. First, schools were defined as the primary sampling unit, then in the second stage, students were randomly selected from the sampled schools. Each design stage had a target population. In the first stage the target population was regular public and private schools in the 50 United States and the District of Columbia that provided daily on-site instruction to students in both the 9th and 11th grades. Public charter schools were included. The sample schools were originally selected from two sources: the 2005-06 Common Core of Data (CCD) and the 2005-06 Private School Universe Survey (PSS). Later, after the start of recruiting schools, random sample schools were drawn from all new schools listed in the 2006-07 CCD and 2007-08 PSS to maximize coverage of the target population. The sampling strata in the first stage were defined by the interaction of school type (public, private-Catholic, private-other), region of the United States (Northeast, Midwest, South, West) and locale (city, suburban, town, rural). This produced a nationally-representative sample for school size and locale. In the sample, Catholic schools were oversampled in comparison to other private schools. This was in keeping with procedures for complex sampling strata where subgroups are sampled in proportion to the size in the population, with some subgroups being oversampled to ensure enough units for analysis of subgroups (National Center for Educational Statistics, n.d.b).

During this first stage of sampling schools, HSLS:09 recruiters expended extra effort to realize their sample target of 800 participating schools, 600 public and 200 private schools. (Ingles et al., 2011). Compared with previous National Center for Education longitudinal surveys, the HSLS:09 experienced higher rate of school refusals to participate. The common reasons for refusing to participate included concerns about the extra burden on school staff, loss

of instructional time, and the over-testing of students. Recruiters met personally with refusing schools to offer accommodations such as spit survey sessions, before and after school survey sessions, providing students with lunch or breakfast, and providing school-level results, designed to meet their concerns. Some schools accepted the accommodations others still decline to participate. Of the 1,889 eligible schools, defined as public and private schools in the U.S. with both 9th and 11th grades, the final school sample contained 944 schools and met the HSLS:09 sample target.

In the second stage, the target population was all ninth-grade students who attended study-eligible schools in fall 2009. Students were randomly selected from the pool of study-eligible and questionnaire-capable participants according to a stratified systematic sampling in four strata of race/ethnicity: Hispanic, Asian, Black, and other. Asian ninth-graders were oversampled for reasons explained shortly. Students were designated as study-ineligible if they transferred to another school or dropped out prior to the in-school data collection. All foreign exchange students were classified as study-ineligible. Since the questionnaire administered in school was electronic, students found incapable of completing the questionnaire were excluded. Students with physical limitations, cognitive disabilities, limited English proficiency, or emotional limitations were also excluded.

Appropriate number of participants. According to the HSLS:09 documentation (Ingels et al., 2011), power calculations were computed to determine the minimum sample size for students by race and ethnicity required for the analytic objectives of a design effect no larger than 2.0 and maximum correlation estimates for two waves of the study no larger than 0.6. Additionally, two-tailed statistical tests with 0.05 significance level and 80 percent power were conducted to determine sample sizes that would produce relative standard errors no larger than

2.5 and to detect a 5 and 15 percentage point change in the estimated means and proportions across the study waves. The power calculations determined that the sampling rate for both Catholic schools and Asian ninth graders needed to be increased so they were over sampled. The other ethnic group sampling rates were determined to be appropriate. The sample sizes from these calculations were inflated to account for ineligibility and nonresponse. The final 2009 Base Year sample included 26,305 students randomly selected from 944 schools, an average of 28.3 students per public school and 26.1 per private school. The HSLS:09 transcript collection collected transcripts in the 2013-14 academic year from all students who participated in the base year survey. However, six high schools had closed by then so in the final sample 938 base-year schools provided transcripts for 23,415 students.

Instrumentation & Administration

HSLS:09 was administered by the National Center for Educational Statistics (NCES) and made available for public use. For the past 40 years, the NCES has conducted a series of longitudinal surveys to “study the educational, vocational, and personal development of students at various stages in their educational careers, and the personal, familial, social, institutional, and cultural factors that may affect that development” (Ingels et al., 2011, p. 2). The purpose of the HSLS:09 was to monitor a national sample of young people through their high school experiences and to capture their transition to post-secondary education, the workforce, and other adult roles (Ingels et al., 2011). It is therefore a general-purpose data set, designed to serve a variety of policy objectives; it was not designed to test specific hypotheses.

The goal of HSLS:09 is to better understand the impact of earlier educational experiences (starting at 9^{th} grade) on high school performance and the impact of these experiences on

> the transitions that students make from high school to adult roles. (Ingels et al., 2015, p. 6)

It was designed to provide data for future researchers interested in investigating effective high schools, growth in academic achievement, the process of dropping out, and the educational and social experiences that affect those outcomes (Ingels et at., 2015).

The HSLS:09 data provided information on a nationally-representative sample consisting of a single cohort of high school students. Students were first surveyed in Fall 2009 (ninth-grade). Follow-up surveys were conducted with the same students in Spring 2012 (11th grade), Fall 2013 (after their expected graduation), and 2016. Transcripts were collected in 2013-14, the academic year following their expected graduation date. This study utilized data from the Base Year 2009 and the 2013 Transcript Collection. The large cohort sample, multiple collection points over five years, and high response rates (85.1 percent of 25,206 sampled students for 2009 Base Year and 79.3 percent of 23,401sampled students for 2013 Transcript Collection) provided rich data for quantitative analysis (Ingels et al., 2011; Ingels et al., 2015). NCES also provided a set of analytic weights that adjust for unit non-response, cluster sampling, and over-sampling so results may have greater generalizability. This data set provided access to high quality data that could not collected by a single researcher and utilized instruments constructed by expert survey developers (Smith, 2008).

This research utilized data from both the 2009 Base Year Survey instrument and the 2013 Transcript collection. The 2009 Base Year Survey provided the demographic variables and birthdate data. The transcript collection provided the GPA, credits earned, and graduation outcome data. Since demographic and birthdate data were utilized from the base year instrument, the next section gives a brief overview of the development of the survey and description of field

testing. The rest of the variables came from the 2013 Transcript collection which are detailed in a subsequent section.

2009 Base Year survey. The development process for the questionnaire administered in the HSLS:09 Base Year included a review of the literature, consultation with government offices and interest groups, circulating drafts in process, review by a technical panel, writing justifications for questionnaire items for the Office of Management and Budget, and field testing and revision (Ingels et al., 2011). Since the research objectives of the study were longitudinal in nature "…the first priority for the study questionnaire was to select the items that would prove most useful in predicting outcomes as measured in future survey waves" (Ingels et al., 2011, p. 13). The research question assessing the accuracy of ninth-grade predictors of future dropout aligned with the HSLS:09 objective of studying items predictive of future outcomes.

The HSLS:09 was the first NCES study to use a computerized survey instrument (Ingels et al., 2011). An advantage of using an electronic questionnaire was its sophisticated adaptive technology that routed subsequent questions according to first-stage responses. This eliminated the possibility of administrator error in scoring first-stage responses. The technology had the capability of branching questionnaire items, which was necessary to capture complex student decision-making pathways. Additionally, the instrument prompted responders to correct errors and omissions and provided helpful text when needed. The instrument was field-tested and evaluated. Item nonresponse and test-retest reliability were examined, scale reliabilities were calculated, and correlation between related measures was examined.

Students self-administered the questionnaire on a computer during school hours (Ingles et al., 2011). A telephone interview, with branching interview questions, was conducted with the same instrument with students unable to complete the survey in school. The instrument

contained nine sections labeled A thru I. Students were randomly assigned to two groups. Half of the students completed the sections in alphabetical order, the other half completed the sections in the following order: A, B, C, E, D, H, G, F, I. The section order was varied between the two groups to balance non-response for students who were unable to complete the entire survey. This research utilized data from the first section of the questionnaire which asked for demographic information including sex, race/ethnicity, and birthdate. Surveys were also completed by the students' parents, principals, math and science teachers, and the school's lead counselor, either by phone or on the web.

2013 Transcript Collection. Transcripts were collected using a web-based control system with two components: School Contacting System which stored data on students and tracked communication with schools, and the Data Receipt System which managed the transcript and catalog data (Ingels et al., 2015). The staff tasked with managing the web-based control system received a 2-day training on collecting transcript and catalog information, gaining cooperation from schools, problem resolution, working with the collection and receipt systems, and reviewing HSLS:09 information and confidentiality regulations.

In September of 2013, materials were sent to all base-year and transfer schools (Ingels et al., 2015). The materials provided information on preparing the transcripts and related documents. They directed schools to the secure study website and asked them to provide basic enrollment information, testing and course-taking information on each participating student, and information regarding the school's grading policies and graduation requirements. Schools were asked to upload transcripts to the secure website. Alternatively, they could fax them to a secure number, send by encrypted email, or send by FedEx.

Once received, the transcripts and course catalog information were coded and keyed into a web-based data entry application (Ingels et al., 2015). Course catalog information was keyed in first. HSLS:09 used the School Codes for the Exchange of Data to classify high school courses by subject with a five-digit code. Then transcripts were keyed into the system. One keyer/coder was assigned to a single course catalog or all transcripts from a single school. This allowed keyer/coders to become familiar with the data. Keyer/coders received training and were assigned a quality control supervisor. Quality Circle meetings were held weekly to provide additional training and discuss aspects of the keying and coding procedures.

Quality control of the keying and coding included three checks. The course catalog data coding was assessed with double-coding with arbitration, the Cohen’s Kappa statistic for inter-rater reliability was 0.95 for the 2-digit general category and 0.70 for the 5-digit specific course code which is considered “almost perfect agreement” and “substantial agreement” respectively (Ingels et al., 2015, p. 77). To evaluate the reliability of the keying and coding of transcripts, 10 percent of each keyer/coders’ transcripts were rekeyed by a different keyer/coder. The Cohen’s Kappa was 0.80. Additionally, all transcripts coded with “other, specify” were reviewed to determine if they could be coded with any other existing choices.

Analytic weights. HSLS:09 data contained a series of analytic weights computed to accommodate analyses specific to each round of the study (Ingels et al., 2015). The weights had bias correction properties and were needed to correctly calculate standard errors of population estimates. Since the research questions focused on students and drew data from the 2009 Base Year Survey and the 2013 Transcript Collection, the W3W1STUTR weights were used (Ingels et al., 2015). In the 2009 Base Year Survey, the final student weight was computed from a base weight which accounted for the non-equal probability of selection resulting from the two-stage

stratified sampling design. The base weight was adjusted for non-response, both parent refusal and student refusal to participate. The weight was then calibrated to adjust for the difference between the weighted sums and the sampling frame. The discrepancy was due to differences in student counts and race/ethnicity declarations in the time lag between the sampling records and the current records submitted by the schools. This became the final student weight for the 2009 Base Year data.

The W3W1STUTR base weight was constructed from the final student weight for the 2009 Base year (Ingels et al., 2015). The weight was adjusted for nonresponse of participants to prior rounds of the survey, specifically the base year, first follow-up and the 2013 Update. A second adjustment was made to account for all additional non-respondents associated with the weight. The weight was adjusted a third time for missing transcript data and then calibrated to ensure the weight was representative of the population defined in the base year. A nonresponse bias analysis initially showed a 35.8 percent bias in 67 comparisons of 18 test variables. After applying four nonresponse adjustments, the analysis showed no significant bias in any of the 67 comparisons. This final W3W1STUTR was used to construct the 200 Balance Repeated Replicate weights (W3W1STUTR001-200) this research used to calculate standard errors in the analysis.

Data imputation. HSLS:09 used variable imputation to resolve missing data issues (Ingels et al., 2011). In general, the survey did not suffer from high levels of item non-response, however 18 key analytic variables were identified for item imputation to facilitate complete-case analysis of all respondent records. According to the HSLS:09 documentation three of the student variables used for descriptive analysis have imputed values: sex, race/ethnicity, and SES (Ingels et al., 2011). Students’ sex was logically imputed from enrollment lists, gender-specific names,

parent responses, or student responses to other items in the questionnaire. Students' race and socioeconomic status was derived from other imputed variables. The imputed values were evaluated by three checks. The checks compared the distributions of the variables before and after imputation. Any differences greater than five percent were flagged and examined. Changes were made and checks were rerun. These imputed values, if close to the true value, produce results that were "… likely less biased than those produced with the incomplete data file" (Ingels et al., 2011, p. 162).

Analytics

SAS® University Edition software was used to analyze the HSLS:09 public use data, 2009 Base Year and 2013 Transcript collections, downloaded from the National Center for Education Statistics (2018) website. The SAS® University Edition (n.d.) was available by free download for academic and non-commercial use from the SAS® website. The download included a virtual machine that allowed users access to a remote server which accessed data stored on a laptop, performed the requested calculations, and returned the results to a laptop. All data, programming files, and results were stored on a personal, password-protected laptop.

SAS® University Edition contained a survey module for computing statistics for complex survey data.

> Many standard software packages calculate estimates under the assumption of a simple random sample design as in traditional mathematical statistics and do not account for the clustering of students within schools. This incorrect design assumption can lead to estimated variances and confidence intervals that are too small and can, therefore, lead to incorrectly rejecting the null hypothesis for statistical tests of differences. (Ingels et al., 2011, p. 131)

The SAS® survey module had an option to select the Balanced Repeated Replicates method of variance estimation for correctly calculating standard errors of complex survey data (Lewis, 2017). Lewis advises that variance estimates will be incorrect if the technique specified in the analysis software is different from the technique that was used to construct the weights.

Missing values. Missing responses in survey data lead to a loss in precision of statistical estimates (Brick, 2013). Although HSLS:09 imputed values for some variables and provided analytic weights to adjust for non-response bias, the variables of type, locale, region, GPA, credits earned, and over-age status contained item non-response. Dong and Peng (2013) define item non-response as incomplete information collected from a respondent. The respondent skips one or two questions but completes the rest of the survey. HSLS:09 coded variables with “Unit non-response” to indicate the data were not available because of unit non-response or the question did not apply (National Center for Education Statistics, n.d.b). Data were coded “Missing” to indicate an item that may apply but the respondent did not answer the question, or the question was not answered because the gate or introductory question was not answered. Missing data in this study was treated as item non-response, as defined by Dong and Peng. They indicate a missing rate of less than 5% as being inconsequential, but if data from more than 10% of respondents are missing then results are likely to be biased.

The data for variables type, locale, and region came from a branched introductory question: “[Have/Has][you/your teenager] earned a high school credential such as a high school diploma, certificate of attendance, GED or other high school equivalency?” (National Center for Education Statistics, 2014). If respondents answered “Yes,” they were routed to questions regarding the type of credential, date earned, and name and location of the school in that order. If they answered “No,” they were routed to questions regarding if they were still attending school,

their plans for earning a credential, and the name and location of the school they last attended. Although it is not possible to know the reasons why some respondents did not answer the survey question, the placement of the branching directly after a question about earning a diploma may have contributed to non-response. Regardless, the missing values for the three variables represented 5.0% (type), 5.5% (locale) and 5.5% (region) of the sample and, in reference to Dong and Peng, were not included in the analysis.

The missing data for the variable over-age status that was coded as “Missing” represents 0.2% of the sample and was not included in the analysis. The variables GPA, credits earned, over-age status and dropout were collected from student transcripts. GPA has both “Missing” and “Unit non-response” totaling 6.7% of the sample. Credits earned and dropout have 6.7% “Unit non-response.” Since these fell within Dong and Peng’s (2013) suggested parameter of less than 10% they were not included in the analysis.

Independent and dependent variables. The following is the list of independent and dependent variables that were tagged and downloaded from the HSLS:09 data files. The variable name is listed first. In bracketed parentheses is the name used in the HSLS:09 code book (Ingels et al., 2011). Next, is a short description of the variable. Lastly, the coding that was used for analysis is detailed.

Descriptive independent variables.

- *Sex* [X1SEX]. The sample members were asked their sex. Missing items were imputed from the base year student questionnaire, the parent questionnaire, information provided by the school, or by manual review of the student’s first name. *Sex* was coded as:
    - *Sex*=1 (male)
    - *Sex*=0 (female).

- *Race/Ethnicity* [X1RACE]. This was a composite measure based on students' answers to dichotomous questions regarding their race/ethnicity. Missing items were imputed from a school-provided roster or from the parent questionnaire. *Race/ethnicity* was coded as:
    - *Race/Ethnicity*=1 (White)
    - *Race/Ethnicity* =2 (Black)
    - *Race/Ethnicity* =3 (Hispanic)
    - *Race/Ethnicity* =4 (Asian)
    - *Race/Ethnicity* =5 (Other).
- *SES* [X1SESQ5]. This variable was the quintile coding of the composite measure used as a construct of socioeconomic status. It was calculated using the parent/guardians' education, occupation, and family income. For unit non-response, values were imputed according to the process described in the Data Imputation section above. *SES* was coded the same as X1SESQ5:
    - *SES*=1 (First quintile [lowest])
    - *SES*=2 (Second quintile)
    - *SES*=3 (Third quintile)
    - *SES*=4 (Fourth quintile)
    - *SES*=5 (Fifth quintile [highest]).
- *Type* [X3CONTROL]. This variable identified the student's last attended school as public, Catholic, or other private. The type of school was coded according to its designation in the 2005-06 CCD, 2005-06 PSS, 2006-07 CDD, and 2007-08 PSS files. In the public use files Catholic and other private were coded to one category. *Type* was coded the same as X3CONTROL:

  - *Type*=1 (Public)
  - *Type*=2 (Catholic or other private).

- *Locale* [X3LOCALE]. This variable characterized the "urbanicity" of students' last-attended school. The schools were coded according to their designations in the 2005-06 CCD, 2005-06 PSS, 2006-07 CDD, and 2007-08 PSS files. *Locale* was coded the same as X3LOCALE:
  - *Locale*=1 (City)
  - *Locale*=2 (Suburb)
  - *Locale*=3 (Town)
  - *Locale*=4 (Rural).
- *Region* [X3REGION]. This variable identified the geographic region where the sample members last attended school. The geographic regions were designated as Northeast (CT, MA, ME, NH, NJ, NY, PA, RI, VT), Midwest (IA, IL, IN, KS, MI, MN, MO, ND, NE, OH, SD, WI), South (AL, AR, DC, DE, FL, GA, KY, LA, MD, MS, NC, OK, SC, TN, TX, VA, WV), West (AK, AZ, CA, CO, HI, ID, MT, NM, NV, OR, UT, WA, WY). *Region* was coded the same as X3REGION:
  - *Region*=1 (Northeast)
  - *Region*=2 (Midwest)
  - *Region*=3 (South)
  - *Region*=4 (West).

Independent indicator variables.

- *GPA* [X3TGPA9TH]. This was the calculated grade point average for all credit-bearing courses reported on the transcript for students' ninth grade year. X2TGPA9TH was coded

into two categories either "Did not attempt" or with a value ranging from 0.25 to 4.00. *GPA* was coded as:

- *GPA*=0 (Did not attempt, values less than 1)
- *GPA*=1 (values greater than or equal to 1 and less than 1.7)
- *GPA*=2 (values greater than or equal to 1.7 and less than 2.7)
- *GPA*=3 (values greater than 2.7 or equal to and less than 3.7)
- *GPA*=4 (values greater than or equal to 3.7)

- *Credits Earned* [X3TCRED9TH]. This was the total Carnegie credits for all courses reported on the transcript for ninth grade. X3TCRED9TH was coded into two categories either "Zero" or with a value between 0.5 and 13.0. Students need to earn 6 or more credits in ninth grade to be on-track to graduate in four years (Allensworth & Easton, 2007) so *Credits Earned* was coded as:
  - *Credits Earned*=0 (Zero and values greater than or equal to 0.5 and less than 1)
  - *Credits Earned*=1 (values greater than or equal to 1 and less than 2)
  - *Credits Earned*=2 (values greater than or equal to 2 and less than 3)
  - *Credits Earned*=3 (values greater than or equal to 3 and less than 4)
  - *Credits Earned*=4 (values greater than or equal to 4 and less than 5)
  - *Credits Earned*=5 (values greater than or equal to 5 and less than 6)
  - *Credits Earned*=6 (values greater than or equal to 6).
- *Over-Age Status* [S1BIRTHYR]. This variable indicated the students' birth year from the 2009 Base Year questionnaire. S1BIRTHYR was coded as: 3 (1992 or earlier), 4 (1993), 5 (1994), 6 (1995), 7 (1996 or later). *Over-Age Status* was coded as:
  - *Over-Age Status*=1 (17 years old or older)

- *Over-Age Status* =2 (16 years old)
- *Over-Age Status* =3 (15 years old)
- *Over-Age Status* =4 (14 years old)
- *Over-Age Status* =5 (13 years old or younger).

Dependent variable.

- *Dropout* [X3OUTCOME]. This variable represented the outcome indicated on the transcript. X3OUTCOME was coded 1 (Fall 2012-summer 2013 graduate), 2 (Post-summer 2013 graduate), 3 (Pre-fall 2012 graduate), 4(Graduation date unknown), 6(Certificate of attendance), 8 (Dropped out), 9 (Transferred), 10 (Left other reason), 11(Still enrolled), 12 (Status cannot be determined). HSLS:09 was designed to follow students who transferred out of their base year schools, however the student designated as "Transferred" in this variable were those whose transcripts school staff designated as unobtainable. "Because dropouts occasionally were enrolled in a school for too brief a period to accumulate a coursetaking record, there is often little or no record of their origin or destination" (Ingels et al., 2015, p. 57). *Dropout* was coded as:
  - *Dropout*=0 (Fall 2012-summer 2013 graduate and Pre-fall 2012 graduate)
  - *Dropout*=1 (Post-summer 2013 graduate, Graduation date unknown, Certificate of attendance, Dropped out, Transferred, Left other reason, Still enrolled, and Status cannot be determined).

Data download procedures. HSLS:09 data was available to the public from the National Center for Education Statistics (2018) website. Clicking the link for the public-use data opened the Education Data Analysis Tool (EDAT) page and prompted users to agree to the terms and conditions of using the data. It was a confidentiality agreement in which users agreed to protect

the identity of the study participants and schools. Next, a page appeared in which users created a login account. Once an account was created users could login to EDAT and download data from the list of available surveys. Once a survey was chosen and the population of analysis was identified, the variable list window opened. Users could then tag the variables they would like included in the downloaded data set. Once the tag file was completed users could select the download tab and follow the instructions. The data file was downloaded to a laptop in the SAS® format with raw data ASCII layout.

SAS® and ROC analysis. SAS® University Edition software and Microsoft Excel for Mac (version 16.9) operating on MacBook Pro, macOS High Sierra (version 10.13.2) was used to analyze the downloaded HSLS:09 data file containing the tagged variables. To address the first research question (*How accurately do the ninth-grade indicators of GPA, credits earned, and over-age status predict the likelihood that students from the HSLS:09 dropped out of high school?*), ROC curves were used to analyze the predictive value of each independent variable (GPA, credits earned and over-age status) on the dependent variable (dropout). ROC analysis was appropriate because it assessed accuracy and provided a comprehensive and visually attractive way to summarize the accuracy of predictions (Gönen, 2007). The diagnostic accuracy of the indicator was derived from the area under the empirical ROC curve. It was not affected by decision criterion and was independent of the prevalence of the outcome variable (Hajian-Tilaki, 2013). A ROC curve assessed how effectively a predictor (categorical, ordinal, or continuous) distinguished a dichotomous outcome. In this study ROC curves were constructed for each of the ordinal independent variables (GPA, credits earned and over-age status) as predictors of the dichotomous dependent variable (dropout).

To construct each ROC curve, SAS® was programmed to use PROC SURVEYFREQ, the procedure in the complex samples module designed for bivariate analysis, to build a bivariate table of the predictor and outcome variables, which displayed the frequencies of dropout for each category on the ordinal scale of the predictor (Lewis, 2017). The table included the weighted frequencies calculated from the 200 BRR weights in the data set. SAS® was instructed to use PROC SURVEYFREQ to test for associations. For significant associations, the weighted frequencies were imported into an Excel worksheet on which two-by-two matrices were constructed for each threshold value and the sensitivity and 1-specificity of each threshold was calculated. Excel then plotted the empirical ROC curve. These ROC curves drawn using the SAS®-produced weighted frequencies accounted for the complex survey design.

It was necessary to import the weighted frequencies into Excel because the complex survey procedure package in SAS® did not currently have a command to construct the ROC curves directly (Agnelli, 2014). The workaround for drawing ROC curves with complex survey data was to first use PROC SURVEYLOGISTIC which created a logistic regression equation and output the weighted predicted probabilities to a file. Then analysts must instruct the standard PROC LOGISTIC, which assumed a simple random sample, to draw the ROC curve from that data file. ROC curves drawn in this manner were testing the accuracy of the logistic regression model not the predictive accuracy of the variable based on its score or ordinal scale (Agnelli, 2014).

Two steps were used to analyze the plot of each empirical ROC curve. First, the curve was visually compared to the chance diagonal. This 45-degree line represented the results of a predictor which discriminates by pure chance. A ROC curve that distinguished Dropouts from Graduates was well above the chance line. Second, the curve was analyzed statistically. The

diagnostic accuracy of the ROC curve was assessed using the quantitative area under the curve (AUC) index. AUC measured the probability that given a random pair of students the indicator correctly discriminated between dropout and graduate. The AUC could have been calculated directly by summing the areas of trapezoids drawn under the empirical ROC curve. However, because the independent variables were constructed as discrete ordinal categories, the "trapezoid rule tend[ed] to underestimate the area under what is in reality a smooth ROC curve" (Hanley & McNeil, 1982, p. 31).

Let: $x =$ an early warning indicator
$D =$ the population of Dropouts
$G =$ the population of Graduates
$x_D =$ an x value from an individual from D
$x_G =$ an x value from an individual from G
$n_D =$ sample size from D
$n_G =$ sample size from G

Make all $n_D \cdot n_G$ possible comparisons between the n_D sample x_D's and n_G sample x_G's scoring each comparison by the rule:

$$S(x_D, x_G) = \begin{cases} 1 \; if \; x_G > x_D \\ \frac{1}{2} if \; x_G = x_D \\ 0 \; if \; x_G < x_D \end{cases}$$

Average all the S's over the $n_D \cdot n_G$ comparisons:

$$W(AUC) = \frac{1}{n_D \cdot n_G} \sum_{1}^{n_D} \sum_{1}^{n_G} S(x_D, x_G)$$

Let: $Q_1 = prob$ (two randomly chosen Graduates will both be ranked higher than a randomly chosen Dropout)
$Q_2 = prob$ (one randomly chosen Graduate will be ranked higher than two randomly chosen Dropouts)

Then $$SE(W) = \sqrt{\frac{W(1-W)+(n_G-1)(Q_1-W^2)+(n_D-1)(Q_2-W^2)}{n_G n_D}}$$

Figure 6: Conceptualizing the Wilcoxon Statistic (Hanley & McNeil, 1982)

Hanley and McNeil (1982) demonstrate that the AUC is mathematically equivalent to the Wilcoxon statistic (W), the probability of an independent variable x correctly ranking a (dropout, graduate) pair. W reflects what probability of x values of an individual from the population of graduates are greater than the x values of an individual from the population of dropouts. For

example, for the variable GPA, W reflects the probability that the GPA of a graduate will be greater than the GPA of a dropout. The statistic does not depend on the actual values of x but the ranking. I utilized Hanley and McNeil's method (see Figure 6) to estimate the AUC because it produced both the W(AUC) and its associated standard error, $SE(W)$. "The quantity of Wcan be thought of as an estimate of θ, the 'true' area under the curve, *i.e.*, the area one would obtain with an infinite sample and a continuous rating scale" (Hanley & McNeil, 1982, p. 32). These two statistics W(AUC) and SE(W) provided the values needed to the test the null hypothesis that W(AUC) was the same as the 45-degree chance diagonal AUC_0=0.5 with the Z-score: $Z = \frac{W(AUC)-AUC_0}{SE(W)}$ (Hajian-Talaki, 2013).

To address the second research question (*What are the threshold values for each indicator that optimize three criteria: maximum distance from chance, minimum distance to perfect prediction, and equality of sensitivity and specificity?*), three measures of optimization were analyzed as detailed in Pandey and Jain (2016). The first, the Youden Index,

$$J = sensitivity + specificity - 1$$

maximized the distance between the threshold value and the chance line. It measured the maximum potential effectiveness of the predictor. The second, Distance,

$$D = \sqrt{(1 - sensitivity)^2 + (1 - specificity)^2}$$

minimized the distance to the point (0,1). It found the threshold closest the perfect prediction accuracy. The third, Sensitivity & Specificity Equality,

$$E = |sensitivity - specificity|$$

minimized the difference between sensitivity and specificity values. It found the threshold value that had equal sensitivity and specificity. If these three measures did not converge the practical

considerations of choosing one threshold value over the other were discussed in the results section.

Limitations

There were two important classes of limitations associated with this methodological approach. The first was due to the original survey design. The second was due to the planned data analysis. The HSLS:09 survey excluded key members of the target population of ninth-grade students attending study-eligible schools in Fall 2009. Students in ninth grade who transferred or dropped out before the data were collected were not surveyed.

> A little more than 1,000 sampled students were no longer enrolled at the high school on the day of the student session, likely because they were on the roster at the start of the school year but never attended the school, or attended at the time of rostering but had left the given school (e.g., transferred) prior to the student session. (Ingels et al., 2011)

There were likely students from the sampled schools who dropped out in the first weeks of their freshman year that were not surveyed. Since they dropped out before completing the semester their GPAs and credits earned would have been zero, had they been included in the survey. Thus, the results may underestimate the predictive strength of the indicators and the results were interpreted as representative of students attending high school at the time the Base Year survey was conducted. Additionally, the sample excluded students with physical limitations, cognitive disabilities, or limited English proficiency, thus the results are limited to students capable of completing the electronic survey or capable of understanding the questions posed by a telephone interviewer.

In addition to the excluded students, the HSLS:09 survey suffered from an unusually high rate of school refusals to participate as compared to previous National Center for Education

Statistics studies (Ingels, et al., 2011). Of the 1,899 study-eligible schools, 945 refused to participate. A large number of schools, 197, rescinded their participation after initially agreeing to participate. HSLS:09 sent representatives to those schools and because of that personal contact 44 more agreed to participate. The top refusal reasons (40%) cited by schools were concern about extra work for staff, loss of instructional time, and being too busy. Participation was also hindered by cutbacks in school staff and resources due to an economic downturn, and there was an influenza pandemic limiting the availability of school staff. Thus, the results are limited to high schools that were willing to take the time and expend resources to participate in the survey. While it was beyond the scope of this research to examine refusal rates and reasons, the question arises: Were these schools more likely to expend time and resources for students in need of intervention? If so, it was likely that students with low GPA, low credits earned, and/or students who were over-age received interventions and successfully graduated. Thus, the results may underestimate the strength of the indicators.

There was an additional design limitation associated with this study. There was not question on the survey asking students if they had received dropout intervention. Given the increased pressure to improve graduation rates and the current efforts of the U.S. Department of Education (n.d.) to fund grants for high schools to use for dropout prevention and intervention programs, it was likely that students classified as graduates participated in successful intervention programs. These resilient students may have had one or more of the ninth-grade indicators, been identified by their school, received intervention, and proceeded to get back on track to graduation.

The second important limitation involved the data analysis. The data analysis was limited by the capabilities of the SAS® program to handle complex survey data. At the time of this

research the software used logistic regression and the associated predicted probabilities to draw smooth bi-normal ROC curves rather than draw empirical ROC curves based on observed (sensitivity, 1-specificity) pairs (Gönen, 2007). Thus, there was not an available method to calculate the correlation coefficient of paired data (Hanley & McNeil, 1983). This correlation coefficient would describe the sampling variability when the same sample of students is being evaluated with two different indicators. It is a necessary component in the calculations to evaluate the statistical difference between two ROC curves for different indicators measured with the same sample. For example, the correlation coefficient could have been used to calculate the statistical difference between the AUC of the ROC curve drawn for GPA and the AUC of the ROC curve drawn for credits earned. Thus, this research reported the accuracy of each indicator using manual calculations produced in Excel but did not compare their predictive strengths.

Role of the Researcher

At the time of this research, I was a graduate student in the Doctor of Education program at George Fox University. This research was the final project for successful completion of the program. At that time, I was not employed, but for10 years (2005 – 2015) I taught high school mathematics in two different schools. For many of my students, their graduation was in jeopardy due to unsuccessful progress toward meeting their mathematics graduation requirements. During that time students were required to pass Algebra 2 (or an equivalent course) and the High School Proficiency (HSPE) math exam. For two years, I worked with students in a public alternative school who had failed the HSPE math test three times. I was tasked with preparing them to take the Alternative Assessment (a collection of evidence of the Algebra 2 proficiency) which they could substitute for the HSPE math requirement. Those students opened my eyes to the reality of high-stakes requirements and the effect of multiple failures on students' attitudes, motivation,

and self-efficacy. The relationships I built with them gave me a glimpse into the complex nature of their decisions to persist or leave. Several of my students had previously dropped out but subsequently made the decision to return and finish; one, in particular, returned at the age of 20. What a wonderful moment to see the confident and proud expression on her face as she received her diploma. My aim in this research was to gain a bit more understanding of the predictive strength of ninth-grade indicators in the hope that these could be used in the future by school staff to identify students most in need of intervention before they experience multiple failures.

Research Ethics

George Fox University required their Institutional Review Board (IRB) to approve research on human participants (see Appendices A and B). The secondary data set that was downloaded from the National Center for Education Statistics had been prepared for public release. It contained no identifying student or school information. Additionally, the EDAT system required consent to their confidentiality agreement, an agreement not to use the data intentionally to find the identity of participating students or schools, and if any identities were uncovered during research, EDAT must be notified immediately (see Appendix C). I was granted IRB approval on March 21, 2018.

Another ethical consideration is the interpretation of quantitative data. As a researcher, I observed the trends in the data as objectively as possible. However, I was aware of my vested interest in finding significant associations and significant predictors. I was committed to reporting all results with integrity, even those that did not support my preconceived notions.

CHAPTER 4

Results

This chapter discusses the results of analyzing the National Center for Education Statistics (NCES) High School Longitudinal Survey 2009 (HSLS:09) data set. The public use data file was downloaded from the NCES website in SAS® format. The variables for analysis were selected and recoded according to the process discussed in Chapter 3. This results chapter provides the details of the data download, filtering, and recoding. It describes the demographics of the sample and reports the frequency values of the independent and dependent variables. The bivariate analysis is explained and the tests for association statistics are reported. Lastly, the research questions are explored using receiver operating characteristic (ROC) analysis.

Data Recoding and Analysis

On March 27, 2018 the HSLS:09 data set in SAS® format was downloaded to a password-protected laptop. The file contained 23,503 total observations and 6,607 variables. Using SAS®, the file was filtered to include only the 211 variables needed for analysis: student identification number, five demographic variables, three independent variables, one dependent variable, one base analytic weight, and the 200 replicate weights. The filtered file was saved to the laptop. Next, the missing data pattern was examined. As discussed in Chapter 3, HSLS:09 utilized data imputation to resolve missing data issues (Ingels et al., 2011); however, variables in the downloaded data set still contained some missing values. Table 1 displays the missing data pattern for the four variables of interest: *GPA*, *credits earned*, *over-age status*, and *dropout*. Note that 93.3 percent of the HSLS:09 sample had observed values for *dropout*. These 21,928 observations became the analytic sample for this study.

Table 1

Missing Pattern for the Four Variables of Interest

GPA	Credits Earned	Over-Age Status	Dropout	Frequency	Percent (%)	Cumulative Frequency	Cumulative Percent (%)
0	0	0	0	20,039	85.26	20,039	85.26
0	0	1	0	1,889	8.04	21,928	93.30
1	1	0	1	1,405	5.98	23,333	99.28
1	1	1	1	170	0.72	23,503	100.00

Note. Missing represented with 1 and observed represented with 0.

The variables were recoded as described in Chapter 3 with the addition of using SAS®'s missing code of "." for missing values. Table 2 displays the frequency of missing observations in the analytic sample. Note that all percentages fall within Dong and Peng's (2013) suggested parameter of not more than 10 percent of total or categorical missing data indicating that statistical inferences are not likely to be biased. This modified file was saved to the laptop and used for analysis in SAS®.

Table 2
Frequencies of Missing Observations for all Variables in the Analytic Sample

Variable	Frequency	Percent (%)
Dropout	0	0.00
GPA	53	0.24
Credits Earned	0	0.00
Over-Age Status	1,927	8.79
Sex	5	0.02
Race/Ethnicity	924	4.21
SES (Quintile)	1,889	8.61
Type	1,006	4.59
Locale	1,010	4.61
Region	1,017	4.64

Note. N=21,928

HSLS:09 used two-stage stratified cluster sampling design and provided Balanced Repeated Replicate (BRR) weights with the data set. As discussed in Chapter 3, the 200 BRR weights account for complex sampling design, which is defined as any sampling selection other than simple random sampling (Lewis, 2010). According to Lewis, the nice feature of BRR weights is that they help protect the identities of participants in data sets made available for use by the public. When data sets include BRR weights, the sets do not need to contain strata and cluster information which often identify geographic regions of participants and decrease confidentiality.

SAS® used the 200 BRR weights to calculate variance estimates. "The variability among estimates calculated using each distinct replicate weight serves as the estimate of variability for the full-sample point estimate" (Lewis, 2017, p. 223). SAS® was instructed to use BRR (*varmethod=BRR*) and the replicate weights (*repweights W3W1STUTR001-W3W1STUTR200*) in each call to PROC SURVEYFREQ to perform descriptive and bivariate analyses.

The results from the bivariate analysis were entered into Excel worksheets designed to calculate the false positive and true positive rates at each cut point of the independent variables. These rates supplied data for Excel to draw ROC empirical curves. Additional formulas programmed into the Excel sheets calculated the Wilcoxon statistic (W), the estimate of the area under the curve (AUC), the standard error of W, and the three common optimization thresholds. Since the bivariate analyses showed significant association, the effect sizes were calculated at each identified threshold value.

Description of the HSLS:09 Sample

The sample used for analysis contained 21,928 participants, with a weighted total of 4,156,276 (see Table 5). The weighted total represents the estimated student population of

ninth-grade students in study eligible schools in the United States in 2009. The descriptive analysis is presented here to depict the representativeness of the sample. As shown in Table 3, 50 percent of the students were female and 48 percent of the students were nonwhite. The sample contained students evenly distributed among SES quintiles, with approximately 20 percent in each quintile. Ninety-three percent of the students attended public schools. Twelve percent of the sample attended schools in Town, while approximately 30 percent attended schools in City, Suburb, or Rural Locales. Geographically, 38 percent of the sample attended schools in the South. The other regions, Northeast, Midwest, and West were represented by approximately 20% of the sample.

Table 3

Frequencies of Demographic Variables in the Analytic Sample

Variable	Frequency	Weighted Frequency	Percent (%)
Sex			
Female	10,779	2,060,414	49.57
Male	11,144	2,095,857	50.43
Race/Ethnicity			
White	11,343	2,159,175	51.96
Black	2,252	564,843	13.59
Hispanic	3,503	910,046	21.90
Asian	1,830	146,523	3.53
Other	2,076	374,764	9.02
SES (Quintile)			
First	3,196	817,103	19.67
Second	3,452	837,127	20.15
Third	3,955	826,519	19.90
Fourth	4,248	823,395	19.82
Fifth	5,188	850,243	20.47
Type			
Public	17,663	3,766,044	93.03
Private	3,259	282,035	6.97
Locale			
City	5,808	1,249,868	30.88
Suburb	6,369	1,153,943	28.51
Town	2,692	489,378	12.09
Rural	6,049	1,153,996	28.51
Region			
Northeast	3,311	16,622	17.71
Midwest	5,591	893,432	22.08
South	8,506	1,525,444	37.70
West	3,503	910,471	22.50

Note. N=21,928

Independent Variables

This study examined the accuracy of three independent variables, *GPA, credits earned*, and *over-age status* in predicting the likelihood that students in the HSLS:09 dropped out. Table

4 contains the frequency, weighted frequency, and percent for each category of the independent variables. Nineteen percent of the students had a GPA less than 1.7, the equivalent of C- or less. Twenty-one percent of the students earned fewer than six credits. A typical ninth-grade course load is six credits. Forty-two percent were age 15 or older, representing one or more years over the typical ninth-grade age of 14 years old.

Table 4

Frequencies for Each Category of the Independent Variables

Variable	Frequency	Weighted Frequency	SE of Wgt Freq	Percent (%)	SE of Percent
GPA	(N=21,875)				
0	1,351	269,504	24,268	6.50	0.59
1	2,770	577,318	26,485	13.92	0.64
2	6,368	1,302,857	29,382	31.41	0.70
3	8,669	1,557,997	34,100	37.57	0.83
4	2,717	439,669	22,972	10.60	0.55
Credits Earned	(N=21,928)				
0	860	173,910	22,499	4.18	0.54
1	168	38,338	6,844	0.92	0.16
2	273	3,459	7,838	1.29	0.19
3	463	87,261	10,119	2.10	0.24
4	854	186,539	21,704	4.49	0.52
5	1,584	339,488	26,067	8.17	0.63
6 or more	17,726	3,277,280	42,408	78.85	1.03
Over-Age Status	(N=20,001)				
17 or older	160	36,892	7,753	88.94	0.19
16	844	83,618	6,425	4.43	0.40
15	7,557	1,556,128	31,683	37.52	0.76
14	11,346	2,352,799	37,496	56.72	0.90
13 or younger	94	18,363	3,323	0.44	0.08

Dependent Variable

The outcome variable in this study, *dropout*, was defined as students who did not earn their diploma within four years of entering ninth grade. In the sample, 23 percent of the students dropped out. The estimated number of students in the population who entered ninth grade in 2009 and did not earn a diploma, as indicated on their transcripts four years later, was 939,000 (SE=43,058).

Table 5

Frequencies for Each Category of the Dependent Variable

	Frequency	Weighted Frequency	SE of Wgt Freq	Percent (%)	SE of Percent
Dropout	5,199	939,000	43,058	22.59	1.04
Graduate	16,729	3,217,276	43,674	77.41	1.04
Total	21,928	4,156,276	9,552	100.00	

Tests for Association

SAS® was instructed to perform bivariate analysis of each independent variable with *dropout* and test for association. Table 6 shows the *GPA* by *dropout* analysis. Eleven percent of the students had a ninth-grade GPA of less than 1.7 and dropped out. Sixty-eight percent of the participants were graduates who had a ninth-grade GPA greater than or equal to 1.7.

Table 6

Bivariate Analysis of GPA by Dropout

GPA	Dropout	Frequency	Weighted Frequency	SE of Wgt Freq	Percent	SE of Percent
0 (GPA < 1)	0	460	101,981	18,653	2.46	0.45
	1	891	167,523	13,695	4.04	0.33
1 (1<=GPA<1.7)	0	1,297	297,569	18,825	7.17	0.45
	1	1,473	279,749	16,155	6.75	0.39
2 (1.7<=GPA<2.7)	0	4,789	1,009,267	25,310	24.34	0.61
	1	1,579	293,591	19,757	7.08	0.48
3 (2.7<=GPA<3.7)	0	7,634	1,392,632	34,979	33.58	0.86
	1	1,035	165,366	18,053	3.99	0.43
4 (3.7<=GPA)	0	2,514	408,984	22,279	9.86	0.53
	1	203	30,685	5,875	0.74	0.14
Total		21,875	4,147,347	11,791	100.00	

Table 7 shows the analysis of *credits earned* by *dropout*. Almost ten percent of the students earned fewer than six credits in ninth grade and later dropped out. Sixty-six percent of the students graduated who had six or more credits earned in their ninth grade.

Table 7

Bivariate Analysis of Credits Earned by Dropout

Credits Earned	Dropout	Frequency	Weighted Frequency	SE of Wgt Freq	Percent (%)	SE of Percent
0	0	409	92,612	18,498	2.23	0.45
	1	451	81,298	9,871	1.96	0.24
1	0	16	2,657	1,703	0.06	0.04
	1	152	35,681	6,841	0.86	0.16
2	0	35	5,107	2,041	0.12	0.05
	1	238	48,352	7,118	1.16	0.17
3	0	148	35,090	6,330	0.84	0.15
	1	315	52,171	7,235	1.26	0.17
4	0	454	108,554	17,101	2.61	0.41
	1	400	77,985	8,288	1.88	0.20
5	0	986	226,741	23,929	5.46	0.58
	1	598	112,747	12,369	2.71	0.30
6 or more	0	14,681	2,746,515	46,390	66.08	1.14
	1	3,045	530,766	34,839	12.77	0.84
Total		21,928	4,156,276	9,552	100.00	

The results of the bivariate analysis of *over-age status* by *dropout* is given in Table 8. Twelve percent of the students were 15 years or older when they entered ninth grade and dropped out. Forty-seven percent of the student entered ninth grade at age 14 or younger and later graduated.

Table 8

Bivariate Analysis of Overage Status by Dropout

Over-Age Status	Dropout	Frequency	Weighted Frequency	SE of Wgt Freq	Percent (%)	SE of Percent
17 or older	0	48	9,646	2,569	0.23	0.06
	1	112	27,246	6,901	0.66	0.17
16	0	370	87,671	10,151	2.11	0.25
	1	474	95,947	10,697	2.31	0.26
15	0	5,614	1,168,841	26,685	28.18	0.64
	1	1,943	387,287	22,383	9.34	0.54
14	0	9,202	1,931,417	42,492	46.56	1.03
	1	2,144	421,382	26,007	10.16	0.63
13 or younger	0	72	14,768	2,914	0.36	0.03
	1	22	3,594	1,379	0.09	0.08
Total		20,001	4,147,800	8,868	100.00	

Since HSLS:09 used complex sampling design and provided BRR weights for analysis, SAS® conducted the Rao-Scott Chi-Square test for association (see Table 9). The Rao-Scott statistic adjusts for the complex sampling design by dividing the standard Pearson Chi-Square by the design correction. The design correction supplies the factor by which the complex survey variance is estimated to be larger than a comparable simple random sample (SRS) design (Lewis, 2017). For the bivariate analysis of *GPA* by *dropout*, the design correction of 6.872 implies that the variance of HSLS:09 data is estimated to be 6.9 times greater than that of a SRS. The Rao-Scott Chi-Square of 499.8565 with four degrees of freedom (*df*) yielded an *F* value of 124.9641 and is significant at $p < .001$.Thus, *GPA* and *dropout* have a significant association. *Credits*

earned and *dropout* (F = 55.5812, df = 6, p < .001) and *over-age status* and *dropout* (F = 64.3561, df = 4, p < .001) are also significantly dependent.

The Rao-Scott Chi-Square tested for independence of the variables. However, large samples tend to detect significance more readily than small samples (Tanner, 2012). Since the sample size of 21,928 in this research is large, the strength of the association was calculated using C and *phi* values (see Table 9). The Coefficient of Contingency (C) measured the effect size of the association and the values of 0.16 (*GPA*), 0.14 (*credits earned*), and 0.12 (*over-age status*) represent small effects (Sprinthall, 2007). The Pearson phi (*phi*) coefficient provides the strength of the correlation of the variables and the *phi* values of 0.16 (*GPA*), 0.15 (*credits earned*), and 0.12 (*over-age status*) represent weak correlations (Tanner, 2012). These values imply that the results differ to a small degree from what would be expected if the associations were independent (Grissom & Lee, 2005). The effect sizes for each threshold value are discussed in a subsequent section.

Table 9

Summary of Bivariate Analysis Tests for Association of Each Indicator by Dropout

	Pearson Chi-Square	Design Correction	Rao-Scott Chi-Square	*df*	*F*	*p*	C	*phi*
GPA by Dropout	3,435.00	6.87	499.86	4	124.96	<.001	0.16	0.16
Credits Earned by Dropout	2,715.17	3.81	398.88	6	55.58	<.001	0.14	0.15
Over-Age Status by Dropout	879.12	3.42	257.42	4	64.36	<.001	0.12	0.12

Note. C is the Coefficient of Contingency. *Phi* is Pearson's phi coefficient.

Research Questions

Research Question 1: How accurately do the ninth-grade indicators of GPA, credits earned, and over-age status predict the likelihood that students from the HSLS:09 dropped out of high school? To answer the first research question, the weighted frequencies produced by the bivariate analysis in SAS® were entered into Excel worksheets. The weighted frequencies provided data for two-by-two matrices for each possible cut point. ROC empirical curves were drawn using the false positive rates and true positive rate calculated in the matrices. Figures 7, 8, and 9 show the ROC empirical curves for each indicator. Each was above the chance line, indicating that *GPA*, *credits earned*, and *over-age status* gave better than chance predictions of the likelihood that students from the HSLS:09 dropped out.

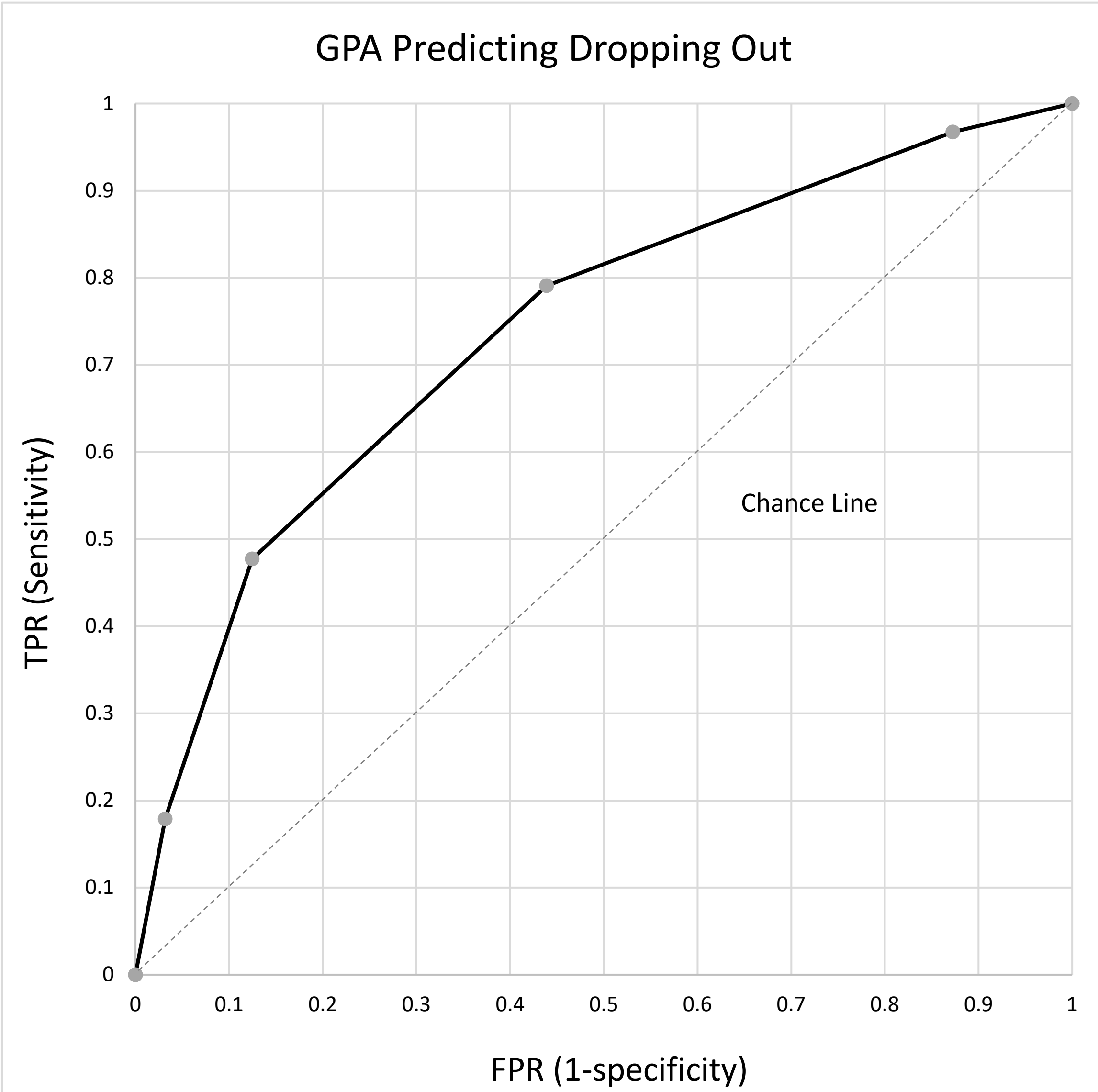

Figure 7: GPA by Dropout ROC Empirical Curve

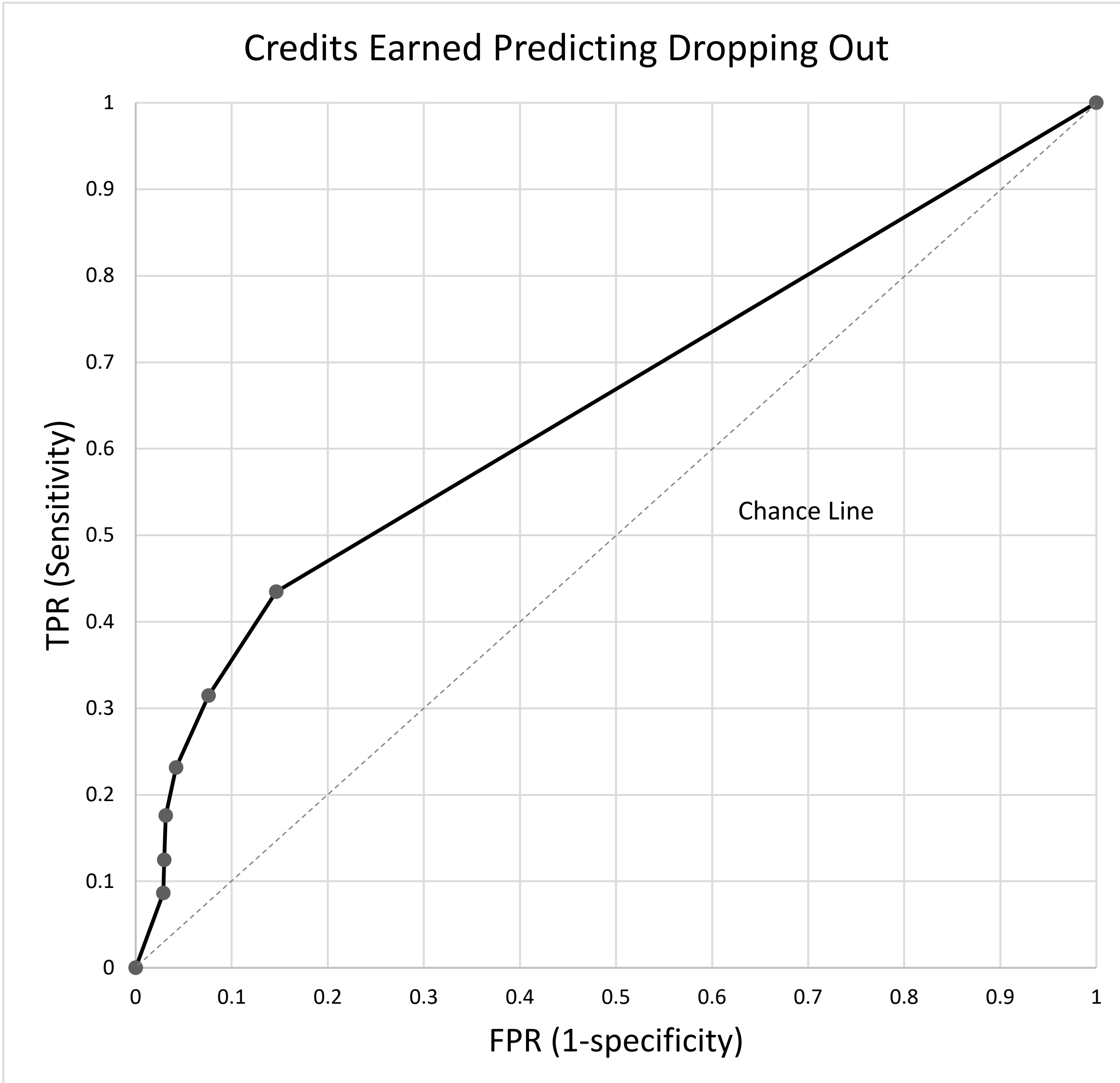

Figure 8: Credits Earned by Dropout ROC Empirical Curve

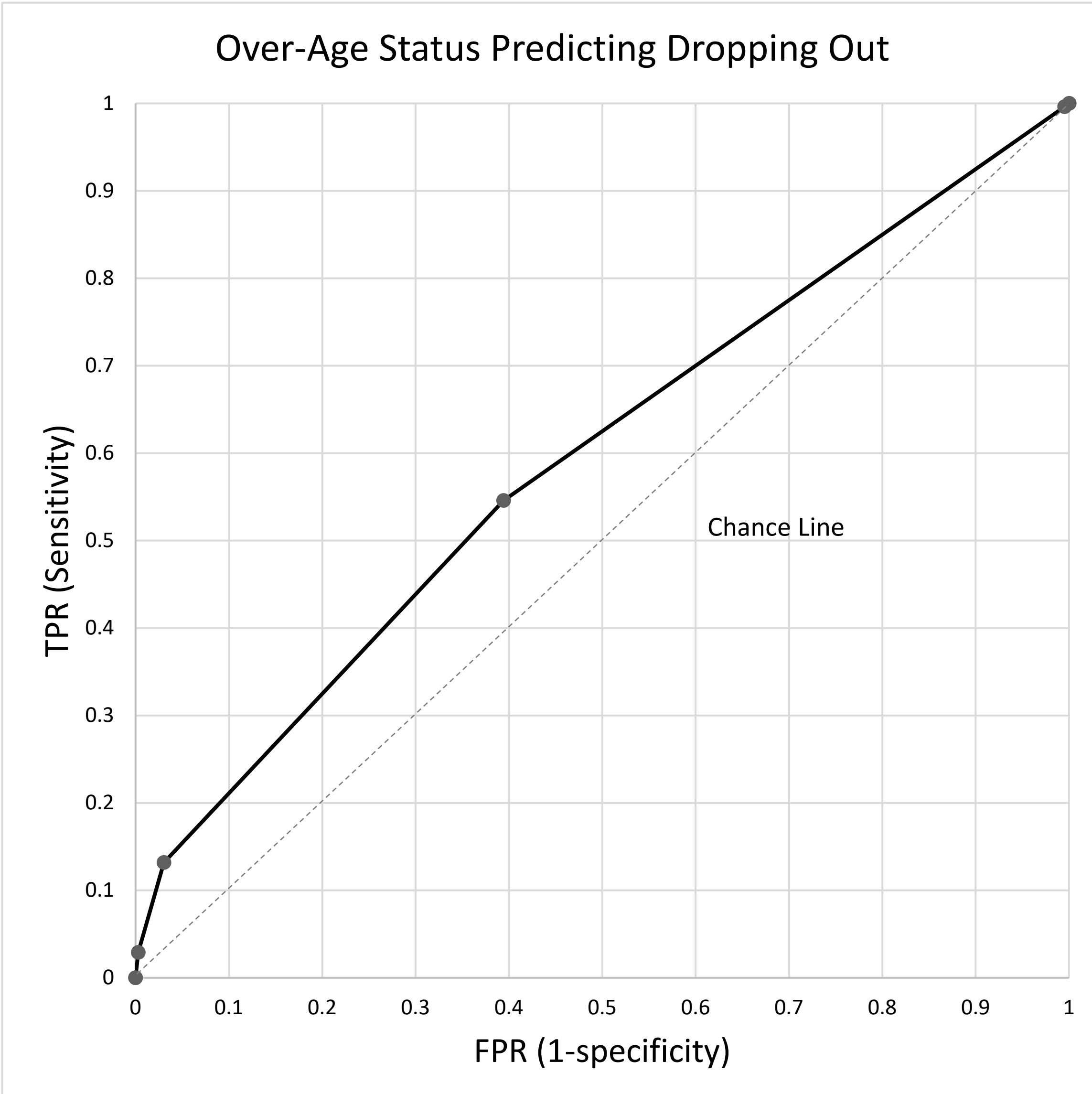

Figure 9: Over-Age Status by Dropout ROC Empirical Curve

After visually inspecting the curves for their position above the 45-degree diagonal chance line, the Wilcoxon statistic (*W*) was calculated. As discussed in Chapter 3, *W* estimates the area under the curve (AUC). It reflects the probability that the indicator correctly distinguishes between dropouts and graduates (see Table 10). The *Z* score tests the null hypothesis that the *W*(AUC) is the same as the chance line. All *Z* scores were significant. The results indicated that *GPA* had a 73.92 percent probability of correctly distinguishing dropouts from graduates and was significantly different from the chance line (*W*(AUC)=0.7392, $p < .001$,

99% CI [0.7386, 0.7397]). *Credits earned* had a 65.18 percent probability of correctly distinguishing dropouts from graduates and was significantly different from the chance line (*W*(AUC)=0.6518, *p* < .001, 99% CI [0.6513, 0.6524]). *Over-age status* had a 59.36 percent probability of correctly distinguishing dropouts from graduates and was significantly different from the chance line (*W*(AUC)=0.5936, *p* < .001, 99% CI [0.5930, 0.5942]).

Table 10

Wilcoxon Statistic for Area Under the ROC Empirical Curve for Each Indicator by Dropout

	W	SE(*W*)	99% Confidence Interval	*Z*	*p*
GPA by Dropout	0.7392	0.0028	[0.7386, 0.7397]	855.49	< .001
Credits Earned by Dropout	0.6518	0.0003	[0.6513, 0.6524]	521.18	< .001
Over-Age Status by Dropout	0.5936	0.0003	[0.5930, 0.5942]	295.32	< .001

Research Question 2: What are the threshold values for each indicator that optimize three criteria: maximum distance from chance, minimum distance to perfect prediction, and equality of sensitivity and specificity? To answer the second research question, formulas embedded in the Excel worksheets calculated the three common optimization criteria. As discussed in Chapter 3, the Youden Index,

$$J = sensitivity + specificity - 1$$

maximized the distance between the threshold value and the chance line. It measured the maximum potential effectiveness of the predictor. Distance,

$$D = \sqrt{(1 - sensitivity)^2 + (1 - specificity)^2}$$

minimized the distance to the point (0,1). It found the threshold closest to perfect prediction accuracy. Sensitivity & Specificity Equality,

$$E = |sensitivity - specificity|$$

minimized the difference between sensitivity and specificity values. It found the threshold value that had equal sensitivity and specificity.

For the *GPA* indicator, the optimization criteria did not converge to a single point (see Table 11). At cut point 2 which represented students with a GPA less than 1.7, the maximum *J* value of 0.35, indicated it was farthest from a chance prediction. Using a cut point of GPA less than 1.7 identified 48 percent of the dropouts and 88 percent of the graduates. It misidentified 12 percent of students (false positives - dropout was predicted but the students graduated). At cut point 3 which represented students with GPA less than 2.7, the minimum *D* value of 0.49 indicated it was closest to perfect prediction and the minimum *E* value of 0.23 indicated it had the minimum difference between sensitivity and specificity. Using GPA of less than 2.7 identified 79 percent of the dropouts and 56 percent of the graduates. It misidentified 44 percent of students (false positives - dropout was predicted but the students graduated). Thus, the cut point of GPA less than 1.7 had a lower false positive rate than GPA less than 2.7, but it identified fewer true dropouts.

For the *credits earned* and *over-age status* indicators, the three optimization criteria converged on a single cut point. For *credits earned* cut point 6 which represented students earning fewer than 6 credits the *J*, *D*, and *E* values indicated that this cut point was farthest from chance, closest to perfect prediction, and minimized the difference between sensitivity and specificity. Using credits earned of fewer than 6 credits represented 43 percent of the dropouts

and 85 percent of the graduates. It misidentified 15 percent of students (false positives - dropout was predicted but the students graduated). For *over-age status* cut point 3 which represented students one or more years over-age, the *J*, *D*, and *E* values indicated that this cut point was farthest from chance, closest to perfect prediction, and minimized the difference between sensitivity and specificity. Using over-age status of one or more years over-age identified 55 percent of the dropouts and 61 percent of the graduates. It misidentified 39 percent of students (false positives - dropout was predicted but the students graduated).

Table 11

Cut Points for Each Indicator that Optimized the Common Decision Criteria

	Cut Point	Interpre-tation	Sensitivity (TPR)	Specificity	1-Specificity (FPR)	*J*	*D*	*E*
GPA by Dropout	2	GPA less than 1.7	0.48	0.88	0.12	0.35	_	_
GPA by Dropout	3	GPA less than 2.7	0.79	0.56	0.44	_	0.49	0.23
Credits Earned by Dropout	6	Earned less than 6 credits	0.43	0.85	0.15	0.29	0.58	0.42
Over-Age Status by Dropout	3	1 year or more over-age	0.55	0.61	0.39	0.15	0.60	0.06

As a final analysis step, effect sizes were calculated for the two-by-two matrices identified as optimizing the common decision criteria. While the Pearson Correlation Coefficient (*phi*) measures the strength of the association between the ninth-grade indicator and the dropout

outcome, other measures of effect size were also calculated because they provide meaningful interpretations of the results for education practitioners. In reference to Grissom and Lee's (2005) assertion that the estimate of the association between the variables in a two-by-two table is the difference between two proportions – specifically the estimate between the probability of a given outcome in the two categories of the independent variable – this study calculated the risk difference for each identified cut point of the independent variables. Grissom and Lee caution against using this measure for variables that are continuous in nature but have been categorized arbitrarily by the researcher. However, in the case of this study the categories were not considered arbitrary because they were vetted through ROC analysis and had significant associations. The effect sizes give information as to the estimated risk difference for each specific cut point. In addition to Pearson's *phi* and the risk difference, this study also calculated the risk ratio and odds ratio at each identified threshold value (see Table 12).

Table 12

Effect Sizes and Practical Significance for the Cut Points of Each Indicator

	Cut Point	Interpre-tation	Probability a Student with the Indicator Dropped Out	Probability a Student without the Indicator Dropped Out	Risk Differ-ence	Risk Ratio	Odds Ratio	*phi*
GPA by Dropout	2	GPA less than 1.7	0.53	0.15	0.38	3.56	6.43	0.29
GPA by Dropout	3	GPA less than 2.7	0.34	0.10	0.25	3.51	4.83	0.29
Credits Earned by Dropout	6	Earned less than 6 credits	0.46	0.16	0.30	2.87	4.49	0.30
Over-Age Status by Dropout	3	1 year or more over-age	0.29	0.18	0.11	1.60	1.85	0.13

Note. Pearson Correlation Coefficient is represented by *phi*

The first effect size calculated was the risk difference. It is the difference of two estimated probabilities: the probability that a student with the ninth-grade indicator dropped out and the probability that a student without the ninth-grade indicator dropped out. Grissom and Kim (2005) refer to these probabilities as success proportions. Note that for each indicator the success proportion, the successful dropout prediction proportion, for a student with the indicator was greater than the success proportion for a student without the indicator. The risk difference is interpreted as the number of more students out of 100 with the indicator who dropped out than students who did not have that indicator and dropped out (Grissom & Kim, 2005). For example,

out of every 100 students, 38 more students with GPA less than 1.7 dropped out than students whose GPA was greater than or equal to 1.7. It can also be interpreted as the percent the dropout rate was higher among students with the indicator (Kline, 2004). For example, the dropout rate was 38 percent higher among students whose ninth grade GPA was less than 1.7 than students whose ninth grade GPA was greater than or equal to 1.7. The next two effect sizes calculated were the risk ratio and odds ratio. The risk ratio is interpreted as the risk of dropout among students with the indicator and the odds ratio explains the odds for dropout among students with the indicator (Kline, 2004). For example, the risk of dropout was 3.6 times greater for students with ninth grade GPA less than 1.7. The odds for dropout among students with ninth grade GPA less than 1.7 were 6.4 times the odds of dropout among students whose ninth grade GPA was greater than or equal to 1.7. The final statistic calculated was the population Pearson (*phi*) correlation between the students with - without the indicator and dropout – graduate dichotomies in the two-by-two matrices at each threshold value (Kline, 2004). For example, the correlation of the dichotomies for GPA at cut point 2, *phi* = 0.2945, can be interpreted as 8.67% (0.2945^2*100) of the information about dropout was contained in GPA.

Conclusion

For students participating in the HSLS:09, the indicators of GPA, credits earned, and over-age status were significant predictors of dropping out. All yielded better than chance identification of dropouts versus graduates. The AUC values indicated that GPA had a 74 percent probability of correctly identifying dropouts from graduates, credits earned had 66 percent probability and over-age status had 59 percent probability. The threshold values corresponding to three common optimization criteria were reported. GPA of less than 1.7 misidentified the fewest students. GPA less than 2.7 identified the greatest number of true

dropouts but had a false positive rate of 44 percent. Earning fewer than 6 credits was the threshold for credits earned that optimized all three criteria as did being one or more years over-age.

While the associations between the indicators and dropping out were statistically weak, they have practical significance. The risk of dropping out for students with ninth grade GPA less than 1.7 was three and a half times that for students with GPA greater than or equal to 1.7. The odds of dropping out for students who earned fewer than six credits in ninth grade was four and half times that of students who earned six or more credits. Additionally, out of every 100 students who enter ninth grade over the age of 14, eleven more will dropout than those who were 14 years old or younger. For students in the HSLS:09 who entered ninth grade in 2009, GPA, credits earned, and over-age status were early warning indicators with a better than chance prediction of future graduation outcomes.

CHAPTER 5

Discussion and Conclusions

Introduction

This research project found its roots in the two years I spent working with students on the brink of dropping out in an alternative high school. It began in the lunch room one day when one teacher pondered: What if we could get students to our school before they failed at the big high school on the hill? What if we could identify students who would benefit from the type of instruction, assessment, and intervention strategies offered at our school before they experienced the effects of getting off-track to graduation?

It has been three years since I last walked the halls of that school building, since I last sat next to a struggling math student guiding her through the thinking necessary to solve a tricky word problem. Yet, the joy of watching students walk the stage in June to receive their diplomas, the memories of the hard work they accomplished to get there stays with me. Most of the original staff still work at the alternative high school and this past June the school was awarded the 2017 School of Distinction Award for improving graduation rates from 68.2 percent in 2012 to 86.1 percent in 2017(Camas School District, 2017).

The purpose of this study was to assess the accuracy of three ninth-grade early warning indicators in predicting the likelihood that students will drop out of high school. It sought to extend Bowers et al.'s (2013) analysis of dropout flags by conducting a full ROC analysis of GPA, credits earned, and over-age status. Each ROC empirical curve gave an estimate of the likelihood that students with the indicator would dropout. The sensitivity and specificity at each threshold of the indicator was used to determine cut points that optimized three common decision criteria. These cut points have practical significance to teachers and school staff interested in

allocating intervention resources. ROC analysis gives information regarding the accuracy of the cut points so that teachers and school staff can weigh the cost and benefits of identifying true dropouts without misidentifying students who would graduate without intervention.

Discussion of Findings

The results showed that all three indicators gave a better than chance prediction of whether students dropped out of high school. This study showed that the early warning indicators of GPA, credits earned, and over-age status provide quantitative information regarding whether students dropped out. The associations between the cut-points of GPA less than 1.7, earning fewer than 6 credits, being one or more years over-age, and dropping out were significant. The effect sizes were small, yet a difference of thirty-eight students out of one hundred has practical significance for teachers working to support all students toward successful graduation. The cut points have practical utility for teachers interested selecting students for participation in programs or strategies designed to keep them on track for graduation.

Dropout. The purpose of this study was to investigate factors that predict students being off-track for completing a regular high school program of study. Since a regular high school diploma is awarded to students who have completed some higher level of academic achievement as evidenced by meeting or exceeding a set of coursework and performance standards set by a state or school district (McFarland, Stark, & Cui, 2016), it was considered the gold standard for the ROC analysis. Thus, dropouts were defined as students who did not earn their regular high school diploma within four years of entering ninth grade. This definition aligns with the adjusted cohort graduation rate (ACGR) which U.S. Governors committed as of 2005 to use as a consistent measure in calculating graduation rates (National Governors Association, 2009). ACGR is defined as the percentage of first-time ninth graders who graduate in four years with a

regular high school diploma. It focuses on the receipt of a regular diploma by a single cohort of public high school students and is considered the most accurate measure for reporting on-time graduation rates (McFarland, Stark, & Cui, 2016).

The students in this study entered ninth grade for the first time in 2009. Using transcripts collected four years later in 2013, the results showed that 23 percent of the students dropped out. For the students in the HSLS:09, 23 out of 100 did not earn a regular diploma within four years. This finding is higher than an analysis of the Common Core of Data by the National Center for Education Statistics (McFarland, Stark, & Cui, 2016) which reported the ACGR as 81 percent in 2013. A possible reason for this is that there were students in the HSLS:09 who their schools reported as transferred, but the identified transfer school had no record of the student attending (Ingels et al., 2015). As explained in Chapter 3, these students were coded as dropouts because they had no record of still being enrolled. It is likely that some of these students enrolled in a different school and later graduated on time.

Note that students who earned an alternative or General Educational Development (GED) certificate are not considered in the ACGR. Similarly, these students were considered dropouts for the purposes of this study. Although data for alternative completers of high school is not readily available, data for the following year, 2014, showed that 92.4 percent of 18 to 24 year-olds held a high school diploma or alternative credential (McFarland, Cui, & Stark, 2018). This suggests that some students may have stayed enrolled to earn their regular diploma in the future or if they dropped out, may have found their way back into the educational system through alternative credential programs offered by their school district or earning their GED through community colleges (Pharris-Ciurej, Hirschman, & Wilhoft, 2012).

This study focused on Early Warning Indicators of being off-track in earning a regular high school diploma because prior research has shown that earning alternative certificates such as the GED are associated with similar negative outcomes as those experienced by students who drop out (Chapman, Laird, Ifill, & KewalRamani, 2011; Heckman & LaFontaine, 2006). Particularly, Ou (2008) found that graduates, GED recipients, and dropouts effectively functioned as three levels of educational attainment. The three groups varied significantly in the outcomes of earnings, incarcerations, mental health, and substance use. Although students wanting an alternative route to high school completion need different kinds of intervention than students headed for dropout, including these students for identification purposes may provide opportunities for school staff to convey the message to students that earning a GED is different from earning a regular diploma.

Early Warning Indicators predict dropout. Students' ninth grade GPA, credits earned, and over-age status are factors that may give information regarding their transition to high school, academic achievement and persistence, and their navigation of the social aspects of schooling. These indicators measured at the end of students' first year of high school may give information regarding their educational trajectory. Students with low grades may not be learning enough course content to sustain successful progress through the high school curriculum. Students with low credits at the end of their ninth-grade year may not be accumulating enough credits to graduate in four years. Students one or more years older than 14 may experience strong pulls to enter the workforce before finishing high school (Cameron, 2012; Halx & Ortiz 2011; Stearns & Glennie, 2006). The results of this research indicate that the factors of GPA, credits earned, and over-age status significantly distinguished dropouts from graduates. While the

associations were weak and the effect sizes small, the results support previous research and have practical significance to educational practitioners.

The results of this study showed GPA had a 74 percent probability of correctly distinguishing students in the HSLS:09 who dropped out from those who graduated. This supports Allensworth and Easton’s (2007) finding that GPA correctly identified non-graduates in Chicago Public Schools 73 percent of the time. The odds for dropping out were six times higher for students with GPA less than 1.7. Pharris-Ciurej, Hirschman, and Wilhoft (2012) examined GPA as a ninth-grade indicator and found that the odds of graduation for students in a West Coast metropolitan school district with GPA between 1.0 and 1.99 were one-seventh the odds of students with GPA greater than 3. For school staff and teachers working to assist all students toward successful graduation, perhaps the most practical interpretation is that this study showed that for every 100 students with GPA less than 1.7, thirty-eight more dropped out and for every 100 students with GPA less than 2.7, twenty-five more students dropped out than those with higher GPAs. Not only does a low GPA signal that a student is struggling academically, it provides information to school staff regarding where to target intervention strategies.

Students earn credits for successful completion of course content. Students who fail a course do not earn the credit. In this study credits earned had a 65 percent probability of correctly distinguishing students who dropped out from students who graduated. This is similar to Allensworth and Easton’s (2007) finding that failing at least one course correctly identified non-graduates 66 percent of the time. It is slightly lower than Kemple, Segeritz, and Stephenson’s (2013) finding that earning fewer than five credits in ninth-grade correctly identified 77 percent of the dropouts in New York City Schools. MacIver and Messel (2012) tied course failure with chronic absenteeism and found them to be stronger predictors than suspensions or demographic

variables for student in Baltimore City Schools. The results of this study showed that the risk for dropout is almost three times higher for students who earned fewer than six credits in ninth grade. These students with low credits would likely benefit from credit recovery programs.

Over-age status, in this study, refers to students older than 14 which is the typical age of students entering ninth grade. Students may enter ninth grade over-age due to grade retention in elementary school or a delayed start for kindergarten. This study found that over-age status had a 59 percent probability of correctly distinguishing dropouts from graduates. This slightly better than chance prediction supports Allensworth and Easton’s (2007) finding that being over-age added little to predict dropping out. However, it differs from MacIver and Messel’s (2012) finding that male students in Baltimore Public Schools who were over-age for ninth grade were significantly more likely to not graduate than females and non-over-age students by a factor of two. Future research could explore the effects of the two different types of over-age status, retained and academic red-shirted, in predicting graduation outcomes. Being over-age did not have a strong as predictive value as GPA and credits earned. However, the three factors taken together may give a better prediction than each taken individually. This could be explored in future research.

Receiver operating characteristic curves and threshold values. As the previous section demonstrates comparing ninth-grade indicators for predicting graduation outcomes across research studies is not straight forward. Researchers tend to define indicators as a group of factors such as the On-Track indicator (Allensworth & Easton, 2005, 2007; Miller, Allensworth, & Kochanek, 2002) which included accumulating five full credits and having no more than one course failure, and as the Early Warning Indicator (MacIver & Messel, 2012) which was defined as attendance, behavior problems, and course failures. Researchers consider students as off-track

for graduation if they exhibit one or more of the factors. Additionally, few researchers report the sensitivity and specificity of their results. In response to Bowers et al.'s (2013) recommendation, this research study examined the factors individually using ROC analysis and reported the sensitivity and specificity.

This analysis not only provided the overall predictive value for each indicator, it also provided threshold values for each based on common optimization criteria. Cut scores have practical utility for teachers and other school staff when making decisions on which students should receive limited intervention resources. For example, the indicator GPA had an overall predictive value of 74 percent, indicating it is a good predictor. Using the cut point of GPA less than 1.7 optimized the distance from the chance line and identified the most graduates without misidentifying students as dropouts who later graduated. Using the cut point of GPA less than 2.7 optimized the distance to perfect prediction and the balance between specificity and sensitivity. It identified the most dropouts, but it had the most false positives. Teachers and school staff may use this information to weigh the benefits of providing intervention resources to the greatest number of potential dropouts against the costs of misidentifying students and allocating resources for students who may graduate without intervention.

In contrast to GPA, credits earned had an overall predictive value of 65 percent and the threshold analysis converged on a single cut point that optimized all three criteria. The cut point of credits earned less than 6 identified 43 percent of the dropouts and 85 percent of the graduates. Likewise, the threshold analysis for over-age status converged to a single cut-point but the overall predictive value of 59 percent is only slightly better than chance. So, even though being 15 years or older upon entering ninth grade identified 55 percent of dropouts it also misidentified as dropouts 40 percent of students who graduated. Teachers and school staff should use the cut

points of earning fewer than 6 credits and being 15 years or older to identify students with caution. These indicators identify students as potentially off-track to graduation but practitioners will want to conduct further assessments to determine student individual needs for intervention to avoid false identification of students who will likely graduate.

Limitations

As discussed in Chapter 3, this study had two design limitations. The first was that the HSLS:09 suffered from a high rate of school refusals to participate (Ingels, et al., 2011). The question was raised in Chapter 3: were the schools willing to take the time to participate in the HSLS:09 more likely to expend time and resources for students in need of intervention? The results raise a second question: could it perhaps be the case that students with the ninth-grade indicators received intervention and successfully graduated? If so, then these results were likely an underestimation of the strength of the indicators. The second design limitation was that students who dropped out in their ninth-grade year before data were collected were not included in the study. These early dropouts, which in Stearns and Glennie's (2006) analysis of data from North Carolina Public Schools represented 8.7 percent of total dropouts, were not represented in this study. Thus, the results are limited to students who dropped out after their ninth-grade year.

There were two additional limitations. First, the results should be considered in light of the students who the HSLS:09 considered questionnaire capable. Students with limited English proficiency and/or limited cognitive abilities were excluded from participation, therefore the results are limited to students able to complete the on-line questionnaire. Practitioners are cautioned against using the early warning indicators of GPA, credits earned, and over-age status on students learning the English language and those with Individual Education Programs. The predictive accuracy for this specific student population was beyond the scope of this study.

Second, this study utilized information collected from students' transcripts. While this gave accurate information regarding whether students earned a diploma, it did not give information regarding whether students had ever dropped out. For example, it is conceivable that some students dropped out at some point in the four years of high school, but re-enrolled and graduated on-time. Also, there were likely students who had the ninth-grade indicators but received dropout intervention and graduated on-time. Future research could examine these resilient students who were counted in this research study as graduates.

Implications for Practitioners

ROC analysis provided an assessment of ninth-grade early warning indicators' strength in predicting students in the HSLS:09 who dropped out. The thresholds on the ROC empirical curves were further analyzed for optimizing three common decision criteria. The sensitivity and specificity of the resulting cut points gave information regarding the utility of that particular cut point in identifying true dropouts without misidentifying students who graduated. Practitioners may use the cut points to identify students potentially in need of intervention.

This study of ninth-grade indicators provides practitioners with quantitative indicators for identifying students in need. In contrast to feelings, prior experience, and personal beliefs regarding students likely to dropout, these indicators can be helpful in overcoming human biases when selecting students in need of intervention strategies. Additionally, the data is easily collected and school staff can examine patterns to address low student performance in strategic ways. Allensworth (2013) gives examples from high schools in Chicago that have been utilizing the indicators to identify students and for school improvement. She discusses the difference that knowledge about the ninth-grade indicators made in the types of conversations teachers had with each other, with students, and with parents. The resulting conversations led to improved student

performance. The schools developed strategies for creating on-track reports, systems for identifying students, and accountability measures for teachers in following through on intervention strategies.

The ROC analysis presented in this research can be duplicated by administrative staff using district-specific data. Utilizing historical ninth-grade data from a single cohort of students and transcript information four years later, the ROC empirical curves could be drawn and cut points identified. The information would give teachers and school staff information specific to students in their own district.

Suggestions for Future Analysis

This study focused on the three ninth-grade indicators of GPA, credits earned, and over-age status which met Bowers et al. (2013) criteria of being accurate, simple to obtain, and usable by school personnel. Additionally, Allensworth (2013) described these factors that are most directly tied to eventual graduation as being most malleable through school practices. “Not only can students be identified for intervention and support, but schools can use patterns in the indicators to address structural issues that make it more difficult for students to learn” (Allensworth, 2013, p. 69). Current research by Allensworth and the Consortium on Chicago School Research is exploring the relationship of student achievement and classroom instruction, specifically the combination of orderly, well-managed classrooms, challenging instruction, and sufficient support for students (University of Chicago, n.d.).

Future analysis could follow any of three possible strands. First, a deeper analysis of the predictive strength of the three factors of GPA, credits earned, and over-age status could be investigated. For example, are students with any two indicators more likely to drop out than those with just one? Or, is there a specific combination of indicators that is more predictive?

Combinations of the factors could be analyzed by creating logistic regression models and constructing ROC curves of the models. Some combination of the ninth-grade indicators may provide more accurate information in identifying students potentially at risk of dropping out without misidentifying those who would graduate. A sub-strand of this deeper analysis could investigate data from geographic regions or from various school districts. Perhaps ninth-grade indicators vary from district to district? As Jerald (2006) notes "[s]ome researchers now believe that the power of place can be as important as the power of time for unmasking the pathways students take to dropping out" (p. 9).

A second strand for future research could investigate a broader list of indicators. While this study focused on the top three factors identified by Bower's et al. (2013) literature review, there are additional factors identified in research that fit the criteria of being easy to obtain, malleable, and usable by schools such as attendance and behavior. MacIver and Messel (2012) included chronically absent and ever suspended in their group of Early Warning Indicators. They analyzed students with any one of the indicators. Future research could conduct ROC analysis of the individual factors of number of days absent and suspensions or behavioral referrals. These could also be analyzed in combination with GPA, credits earned, and over-age status.

A third possible strand for future analysis would be to widen the scope and analyze school, community, and family practices that support or hinder students' progress toward graduation, similar to the current research of Allensworth with the Consortium on Chicago School Research. Future analysis could target students in the HSLS:09 identified as false-positives. These are the students who had the ninth-grade indicator but did not drop out. The questions arise: How did over-age students with low GPA and/or low credits recover and graduate? What factors contributed to their resilience? Perhaps they received school-based

interventions. They might have benefited from supports such as community-based tutoring and mentoring programs. HSLS:09 conducted a follow-up survey during the students' 11^{th} grade and analysis of responses may give insight into factors which may have influenced students to graduate whom were predicted likely to drop out.

Conclusions

The Early Warning Indicators of GPA, credits earned, and over-age status gave better than chance identification of students in the HSLS:09 likely to dropout. These factors may give teachers readily available quantitative data to use in identifying students potentially in need of interventions. This study provides research-based support for the practice used in the alternative school where I previously taught of examining student grades and credits when identifying students likely to benefit from enrolling in our school. Rather than rely on factors such as behavior, probation status, truancy, and attitudes, which are open to human bias, school staff may focus on data which not only identifies potential dropouts but also provides information on where to target intervention strategies. Students with low GPA might benefit from in-school tutoring. Those with low credits might benefit from credit recovery programs. Over-age students might benefit from supportive social environments. Should I return to teaching at the alternative school this research will not only inform my practice but also prompt me to pursue permission to analyze our district data.

This research project has given me a deeper understanding of the dropout problem in general and specifically, the early indicators that potentially identify off-track students. During my time at the alternative school, I was so focused on the students in my classes that I did not take time to investigate the wealth of research being conducted on their behalf. We teachers would strategize ways of engaging the administrators at middle schools in our district to increase

awareness of our school as an alternative not just for struggling students but also for students looking for a different school environment. If I should return to the district, I would also suggest working with the comprehensive high school on the hill. Ninth-grade counselors could use these ninth-grade indicators to reach out to students who would potentially benefit from the educational environment at the alternative school. In the meantime, I find myself visiting the University of Chicago's Consortium on School research website. As I seek to learn more about their research agenda, I find they currently have plans to develop a national model for high school on-track networks (University of Chicago, n.d.). I am eager to follow their progress.

This research project started out in the lunch room at school and it ended in the basement room of my house. The person I was then has changed through the process. The search for answers turned into quiet conversations with the researchers in the journal pages saved to my laptop. I found that I not only understood what they were saying, but I could also see how my questions fit into the picture they were studying. I learned the very lesson that I used to teach my students: *You can do this*. When I realized that SPSS could not analyze data with BRR weights, I knew I would have to learn how to use a different statistics software. I dug deep into the recesses of my undergraduate computer programming training and found that I still remembered how to write code. The language was different but the core concepts similar enough to learn SAS®. Lastly, I have a new appreciation for data analysis. What I thought would be a straight forward process took twists and turn as I wrestled with missing data and effect size interpretations. It was not an easy task to make complex relationships accessible to readers new to ROC analysis. From now on I will look at tables and figures in quantitative research with respect for the depth of analysis behind such straight forward reporting.

Back when I started this journey of earning my doctorate, a wise professor assigned a book on writing. Little did I know then that Ann Lamott's (1994) words would visit me often in the writing of these pages. Her advice of breaking large overwhelming projects into small assignments pressed in on my thoughts each time my stress level rose. "You don't have to see where you're going, you don't have to see your destination or everything you will pass along the way. You just have to see two or three feet ahead of you" (p. 18). Wisdom that I tried to heed. *Just write chapter one,* I would tell myself. *Today, just find a statistics software package that will handle complex data. Right now, learn enough SAS® to write the program* for *bivariate analysis.* So, it went day by day. "That is all we're are going to do for now. We are just going to take this bird by bird. But we are going to finish this one short assignment" (Lamott, 1994, p. 20). Here I am writing these last words; this one short assignment is finished.

References

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

Allen, J. (2005). Grades as valid measures of academic achievement of classroom learning. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas,78*(5), 218-223. doi: 10.3200/tchs.78.5.218-223

Allensworth, E. (2013). The use of ninth-grade early warning indicators to improve Chicago schools, *Journal of Education for Students Placed at Risk (JESPAR), 18*(1), 68-83. doi: 10.1080/10824669.2013.745181

Allensworth, E. M., & Easton J. Q. (2005). *The On-Track Indicator as a Predictor of High School Graduation.* Chicago, IL: Consortium on Chicago School Research.

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.* Chicago, IL: Consortium on Chicago School Research.

Alspaugh, J. (1998). Achievement loss associated with the transition to middle school and high school. *Journal of Educational Research, 92*(1), 20-25

Anderson, G. E., Jimerson, S. R., & Whipple, A. D. (2005). Student ratings of stressful experiences at home and school: Loss of a parent and grade retention as superlative stressors. *Journal of Applied School Psychology, 21*(1), 1-20. doi: 10.1300/J370v21n01_01

Agnelli, R. (2014). *Examples of logistic modeling with the SURVEYLOGISTIC procedure.* Retrieved from: https://support.sas.com/resources/papers/proceedings14/SAS404-2014.pdf

Bassok, D., & Reardon, S. F. (2013). “Academic Redshirting” in Kindergarten: Prevalence, patterns, and implications. *Educational Evaluation and Policy Analysis, 35*(3), 283-297. doi: 10.3102/0162373713482764

Battin-Pearson, S., Newcomb, M. D., Abbott, R. D., Hill, K. G., Catalano, R. F., & Hawkins, J. D. (2000). Predictors of early high school dropout: A test of five theories. *Journal of Educational Psychology*, *92*(3), 568-582. doi:10.1037/0022-0663.92.3.568

Benner, A., Boyle, D., & Bakhtiari, A. (2017). Understanding students’ transition to high school: Demographic variation and the role of supportive relationships. *Journal of Youth and Adolescence*, *46*(10), 2129-2142. doi: 10.1007/s10964-017-0716-2

Borghans, L., Golsteyn, B., Heckman, J., & Humphries, J. (2016). What grades and achievement tests measure. *Proceedings of the National Academy of Sciences of the United States of America, 113*(47), 13354-13359. doi: 10.1073/pnas.1601135113

Bowers, A. J. (2009). Reconsidering grades as data for decision making: More than just academic knowledge. *Journal of Educational Administration, 47*(5), 609-629. doi: 10.1108/09578230910981080

Bowers, A. J. (2010). Grades and graduation: A longitudinal risk perspective to identify student dropouts. *The Journal of Educational Research*, *103*(3), 191-207. doi: 10.1080/00220670903382970

Bowers, A. (2011). What's in a grade? The multidimensional nature of what teacher-assigned grades assess in high school. *Educational Research and Evaluation, 17*(3), 141-159. doi: 10.1080/13803611.2011.597112

Bowers, A. J., Sprott, R., & Taff, S. A. (2013). Do we know who will drop out?: A review of the predictors of dropping out of high school – Precisions, sensitivity, and specificity. *High School Journal 96*(2), 77-100. doi: 10.1353/hsj.2013.0000

Bradley, C. L., & Renzulli, L. A. (2011). The complexity of non-completion: Being pushed or pulled to drop out of high school. *Social Forces,90*(2), 521-545. doi: 10.1093/sf/sor003

Brick, J. M. (2013). Unit nonresponse and weighting adjustments: A Critical Review. *Journal of Official Statistics, 29*(3), 329-353. doi:10.2478/jos-2013-0026

Bridgeland, J., Legters, N., & Balfanz, R. (2010). The new dropout challenge: Bridging gaps among students, parents, and teachers. *New Directions for Youth Development, 2010*(127), 101-110. doi: 10.1002/yd.366

Brookhart, S. M. (1993). Teachers' grading practices: Meaning and values. *Journal of Educational Measurement, 30*(2), 123-42.

Brookhart, S. M., Guskey, T. R., Bowers, A. J., McMillan, J. H., Smith, J. K., Smith, L.F., . . . Welsh, M.E. (2016). A Century of grading research. *Review of Educational Research, 86*(4), 803-848. doi: 10.3102/0034654316672069

Camas School District. (2017). *Report to our community: 2017.* Camas, WA: Camas School District.

Cameron, J. (2012). *Canaries reflect on the mine: Dropouts' stories of schooling.* Charlotte, NC: Information Age.

Carl, B., Richardson, J. T., Cheng, E., Kim, H. J., & Meyer, R. H. (2013). Theory and application of early warning systems for high school and beyond. *Journal of Education for Students Placed at Risk (JESPAR), 18*(1), 29-49. doi: 10.1080/10824669.2013.745374

Chapman, C., Laird, J., Ifill, N., and KewalRamani, A. (2011). *Trends in high school dropout and completion rates in the United States: 1972–2009* (Report No. NCES 2012-006). Retrieved from National Center for Education Statistics website: https://nces.ed.gov/pubs2012/2012006.pdf

Cliffordson, C. (2008). Differential prediction of study success across academic programs in the Swedish context: The validity of grades and tests as selection instruments for higher education. *Educational Assessment, 13*(1), 56-75. doi:10.1080/10627190801968240

Curran, B., Reyna, R., & NGA Center for Best Practices. (2009). *Implementing graduation counts: State progress to date, 2009.* Washington, DC: NGA Center for Best Practices. Retrieved from National Governors Association website: https://www.nga.org/cms/home/nga-center-for-best-practices/center-publications/page-archive/col2-content/graduation-counts.htm

Doll, J. J., Eslami, Z., & Walters (2013). Understanding why students drop out of high school, according to their own reports: Are they pushed or pulled, or do they fall out? A comparative analysis of seven nationally representative studies. *SAGE Open*, 1-15. doi: 10.1177/2158244013503834

Dong, Y., & Peng, C. J. (2013). Principles missing data methods for researchers. *SpringerPlus, 2*, 222. doi: 10.1186/2193-1801-2-222

Dynarski, M., & Gleason, P. (2002). How can we help? What we have learned from recent Federal dropout prevention evaluations. *Journal of Education for Students Placed at Risk, 7*(1), 43-69.doi: 10.1207/S15327671ESPR0701_4

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters, 27*(8), 861-874. doi: 10.1016/j.patrec.2005.10.010

Gleason, P., & Dynarski, M. (2002). Do we know who to serve? Issues in using risk factors to identify dropouts. *Journal of Education for Students Placed At Risk, 7*(1), 25-41. doi: 10.1207/S15327671ESPR0701_3

Gönen, M. (2007). *Analyzing Receiver Operating Characteristic Curves with SAS®*. Cary, NC: SAS Institute.

Greenhoot, A., & Dowsett, C. (2012). Secondary Data Analysis: An important tool for addressing developmental questions. *Journal of Cognition and Development, 13*(1), 2-18.

Grissom, R. J., & Kim, J. J. (2005). *Effect sizes for research: A broad practical approach.* New York, NY: Lawrence Erlbaum Associates.

Hajian-Talaki, K. (2013). Receiver operating characteristic (ROC) curve analysis for medical diagnostic test evaluation. *Caspian Journal of Internal Medicine 4*(2), 627-635.

Halx, M., & Ortiz, M. (2011). Voices of Latino male high school students on their disconnect with education: Perspectives of "drop-outs" and those on the brink. *Latino Studies, 9*(4), 416-438. doi: 10.1057/lst.2011.49

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. doi: 10.1148/radiology.143.1.7063747

Hanley, J. A., & McNeil, B. J. (1983). A method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiology 148*(3), 839-843.

Heckman, J., & LaFontaine, P. (2006). Bias-corrected estimates of GED returns. *Journal of Labor Economics*, *24*(3), 661-700.

Houts, R. (2016). A brief tutorial on ROC curves and area under the curve. Retrieved from: https://youtu.be/hnRBl9-BzjQ

Ingels, S. J., Pratt, D. J., Herget, D. R., Burns, L. J., Dever, J. A., Ottem, R., … LoGerfo, L. (2011). High school longitudinal study of 2009 (HSLS:09) Base-Year data file documentation (NCES 2011-328). Retrieved from National Center for Education Statistics website: http://nces.ed.gov/surveys/hsls09/hsls09_data.asp

Ingels, S. J., Pratt, D. J., Herget, D. R., Bryan, M. Fritch, L. B., Ottem, R., …Christopher, E. M. (2015). High school longitudinal study of 2009 (HSLS:09) 2013 update and high school transcript: Data file documentation (NCES 2015-036). Retrieved from National Center for Education Statistics website: http://nces.ed.gov/surveys/hsls09/hsls09_data.asp

Janosz, M., Archambault, I., Morizot, J., & Pagani, L. (2008). School engagement trajectories and their differential predictive relations to dropout. *Journal of Social Issues*, *64*(1), 21-40. doi: 10.1111/j.1540-4560.2008.00546.x

Jerald, C. D. (2006). *Identifying potential dropouts: Key lessons for building an early warning data system – A dual agenda of high standards and high graduation rates.* Washington, DC: Achieve.

Jimerson, S., Egeland, B., Sroufe, L. A., & Carlson, B. (2000). A Prospective Longitudinal Study of High School Dropouts Examining Multiple Predictors Across Development. *Journal of School Psychology*, *38*(6), 525-549.

Jimerson, S. R., & Ferguson, P. (2007). A longitudinal study of grade retention: Academic and behavioral outcomes of retained students through adolescence. *School Psychology Quarterly, 22*(3), 314-339. doi: 10.1037/1045-3830.22.3.314

Johnson, E., & Semmelroth, C. (2010). The Predictive Validity of the Early Warning System Tool. *NASSP Bulletin*, *94*(2), 120-134. doi: 10.1177/0192636510380789

Jordan, N., Glutting, J., Ramineni, C., & Watkins, M. (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), 181-195.

Kemple J. J., Segeritz M. D., & Stephenson, N. (2013). Building on-track indicators for high school graduation and college readiness: Evidence from New York City. *Journal of Education for Students Placed at Risk (JESPAR), 18*(1), 7-28. doi: 10.1080/10824669.2013.747945

Kline, R. B. (2004). *Beyond significance testing: Reforming data analysis methods in behavioral research.* Washington, DC: American Psychological Association.

Lamott, A. (1994). *Bird by bird: Some instructions of writing and life.* New York, NY: Anchor Books.

Lee, V., & Burkam, D. (2003). Dropping out of high school: The role of school organization and structure. *American Educational Research Journal, 40*(2), 353-393.

Lessard, A., Butler-Kisber, L., Fortin, L., Marcotte, D., Potvin, P., & Royer, E. (2008). Shades of disengagement: High school dropouts speak out. *Social Psychology of Education*, *11*(1), 25-42. doi: 10.1007/s11218-007-9033-z

Lewis, T. (2010). *Principles of proper inferences from complex survey data*. Retrieved from: http://support.sas.com/resources/papers/proceedings10/266-2010.pdf

Lewis, T. (2017). *Complex survey data analysis with SAS*. Boca Raton: Taylor & Francis.

Liao, H. F., Yao, G., Chien, C. C., Cheng, L. Y., & Hsieh, W. S. (2014). Likelihood ratios of multiple cutoff points of the Taipei City Developmental Checklist for Preschoolers, 2nd version. *Journal of the Formosan Medical Association*, *113*(3), 179-186. doi: 10.1016/j.jfma.2011.10.55

MacIver, M. (2013). Early warning indicators of high school outcomes. *Journal of Education for Students Placed at Risk (JESPAR)*, *18*(1), 1-6. doi: 10.1080/10824669.2013.745375

MacIver, M. A., & Messel, M. (2012). *Predicting high school outcomes in the Baltimore City Public Schools*. Washington, DC: Council of the Great City Schools.

MacIver, M. A., & Messel, M. (2013). The ABCs of keeping on track to graduation: Research findings from Baltimore. *Journal of Education for Students Placed at Risk (JESPAR)*, *18*(1), 50-67. doi: 10.1080/10824669.2013.745207

Martin, A. J. (2011). Holding back and holding behind: Grade retention and students' non-academic and academic outcomes. *British Educational Research Journal*, *37*(5), 739-763. doi: 10.1080/01411926.2010.490874

Maué, E. (2016). Achievement—and what else? The standardisation [*sic*]of semester grades due to the implementation of state-wide exit examinations. *Studies in Educational Evaluation*, *51*, 42-54.

McCluckie, B. (2014). Identifying students "At Risk" of withdrawal using ROC analysis of attendance data. *Journal of Further and Higher Education*, *38*(4), 523-535. doi:10.1080/0309877X.2012.726970

McFarland, F., Cui J., & Stark, P. (2018). *Trends in high school dropout and completing rates in the United States: 2014* (Report No. NCES 2018-117). Retrieved from National Center for Education Statistics website: https://nces.ed.gov/pubs2018/2018117.pdf

McFarland, J., Stark, P., & Cui, J. (2016). *Trends in high school dropout and completing rates in the United States: 2013* (Report No. NCES 2016-117). Retrieved from National Center for Education Statistics website: https://nces.ed.gov/pubs2016/2016117rev.pdf

McMillan, J. (2001). Secondary teachers' classroom assessment and grading practices. *Educational Measurement: Issues and Practice, 20*(1), 20-32. doi: 10.1111/j.1745-3992.2001.tb00055.x

Metz, C. E., & Pan, X. (1999). "Proper" binormal ROC curves: Theory and maximum-likelihood estimation. *Journal of Mathematical Psychology, 43*(1), 1-33.

Miller, S., Allensworth, E. M., & Kochanek, J.R. (2002). *Student performance: Course-taking, test scores, and outcomes*. Retrieved from Consortium on Chicago School Research. https://consortium.uchicago.edu/publications/student-performance-course-taking-test-scores-and-outcomes

Miller, S. R, Luppescu, S., Gladden, R. M., & Easton, J. Q. (1999). *How do Burton graduates perform in CPS high schools?* Retrieved from Consortium on Chicago School Research website: https://consortium.uchicago.edu/publications/how-do-barton-graduates-perform-cps-high-schools

Moller, S., Stearns, E., Blau, J. R., & Land, K. C. (2006). Smooth and rough roads to academic achievement: Retention and race/class disparities in high school. *Social Science Research, 35*(1), 157-180. doi: 10.1016/j.ssresearch.2004.08.001

Moretti, E. (2007). Crime and the costs of criminal justice. In C. Banfield, H.M. Levin, & Brookings Institution (Eds.), *The price we pay : Economic and social consequences of inadequate education* (pp. 142-159). Washington, DC: Brookings Institution Press.

Muennig, P. (2007). Consequences in health status and costs. In C. Banfield, H.M. Levin, & Brookings Institution (Eds.), *The price we pay : Economic and social consequences of inadequate education* (pp. 125-141). Washington, DC: Brookings Institution Press.

Munoz-Repiso, A. G., & Tejedor, F. J. T. (2012). The incorporation of ICT in higher education. The contribution of ROC curves in the graphic visualization of differences in the analysis of the variables. *British Journal of Educational Technology*, 901. doi:10.1111/j.1467-8535.2011.01270.x

Muthen, B. O. (2004). Latent variable analysis: Growth mixture modeling and related techniques for longitudinal data. In Kaplan (Ed.), *The Sage handbook of quantitative methodology for the social sciences* (pp. 345-370). Thousand Oaks, CA: Sage.

National Center for Education Statistics (n.d.a). *Getting started with the HSLS:09 data* [Module 2]. Retrieved from https://nces.ed.gov/training/datauser/

National Center for Education Statistics (n.d.b). *Statistical analysis of NCES datasets employing a complex sample design* [Common Module 5]. Retrieved from: https://nces.ed.gov/training/datauser/

National Center for Education Statistics. (2009). *The High School Longitudinal Study of 2009 (HSLS: 09) Base Year* [Data file and code book]. Retrieved from https://nces.ed.gov/surveys/hsls09/hsls09_data.asp

National Center for Education Statistics. (2013a). *Course credit requirements and exit exam requirements for a standard high school diploma and the use of other school completion*

credentials, by state (Table 234.30). Retrieved from https://nces.ed.gov/programs/digest/d13/tables/dt13_234.30.asp

National Center for Education Statistics. (2013b). *The High School Longitudinal Study of 2009 (HSLS: 09) High School Transcripts* [Data file and code book]. Retrieved from https://nces.ed.gov/surveys/hsls09/hsls09_data.asp

National Center for Education Statistics. (2014). *HSLS:09 2013 Update Instrument.* Retrieved from: https://nces.ed.gov/surveys/hsls09/pdf/2013_Student_Parent_Questionnaire.pdf

National Center for Education Statistics. (2015). *Number and percentage of elementary and secondary school students retained in grade, by sex, race/ethnicity, and grade level: 1994 through 2015* (Table 225.90). Retrieved from https://nces.ed.gov/programs/digest/d16/tables/dt16_225.90.asp

National Center for Education Statistics. (2017a). *Percentage of high school dropouts among persons 16-24 years old (status dropout rate) and number and percentage distribution of 16- to 24-year-olds, by selected characteristics: Selected years, 2005 through 2015* (Table 219.73). Retrieved from https://nces.ed.gov/programs/digest/d16/tables/dt16_219.73.asp

National Center for Education Statistics. (2017b). *Public high school graduation rates.* Retrieved from: https://nces.ed.gov/programs/coe/indicator_coi.asp

National Center for Education Statistics (2018). HSLS:09 Data Files. Retrieved from https://nces.ed.gov/EDAT/index.aspx?agrmnt=1

National Governors Association. (2009). *Graduation Counts.* Retrieved from: https://www.nga.org/cms/home/nga-center-for-best-practices/center-publications/page-archive/col2-content/graduation-counts.html

Neild, R. C. (2009). Falling off track during the transition to high school: What we know and what can be done. *The Future of Children*, *19*(1), 53-76.

Neild, R., Balfanz, R., & Herzog, L. (2007). An early warning system. *Educational Leadership*, *65*(2), 28.

Neild, R. C., & Farley, E. (2004). Whatever happened to the class of 2000? The Timing of dropout in Philadelphia's schools. In G. Orfield (Ed.), *Dropouts in America: Confronting the graduation rate crisis* (pp. 207-220). Cambridge, MA: Harvard Education Press.

Neild, R. C., Stoner-Eby, S., & Furstenberg, F. (2008). Connecting entrance and departure: The transition to ninth grade and high school dropout. *Education and Urban Society*, *40*(5), 543-569. doi: 10.1177/0013124508316438

Newton, R., & Rudestam, K. E. (1999). *Your statistical consultant answers to your data analysis questions*. Thousand Oaks, CA: Sage.

Nicholls, G., Wolfe, H., Besterfield-Sacre, M., & Shuman, L. (2010). Predicting STEM degree outcomes based on eighth grade data and standard test scores. *Journal of Engineering Education, 99*(3), 209-223. doi: 10.1002/j.2168-9830.2010.tb01057.x

Ou, S. (2008). Do GED recipients differ from graduate and school dropouts: Findings from an inner-city cohort. *Urban Education 43*(1), 83-117. doi: 10.1177/0042085907305187

Pandey, M., & Jain, A. (2016). *ROC curve: Making way for correct diagnosis.* Retrieved from: https://www.pharmasug.org/proceedings/2016/SP/PharmaSUG-2016-SP11.pdf

Pharris-Ciurej, N., Hirschman, C., & Willhoft, J. (2012). The 9th grade shock and the high school dropout crisis. *Social Science Research, 41*,709-730. doi: 10.1016/j.ssresearch.2011.11.014

Ream, R. K., & Rumberger, R. W. (2008). Student engagement, peer social capital, and school dropout among Mexican American and Non-Latino White students. *Sociology of Education*, *81*(2), 109-139.

Robison, S., Jaggers, J., Rhodes, J., Blackmon, B. J., & Church, W. (2017). Correlates of educational success: Predictors of school dropout and graduation for urban students in the Deep South. *Children and Youth Services Review*, *73,* 37-46. doi: 10.1013/j.childyouth.2016.11.031

Roderick, M. (1994). Grade Retention and School Dropout: Investigating the Association. *American Educational Research Journal,31*(4), 729-59.

Roderick, M., & Camburn, E. (1999). Risk and recovery from course failure in the early years of high school. *American Educational Research Journal,36*(2), 303-343.

Rouse, C. E. (2007). Consequences for the labor market. In C. Banfield, H.M. Levin, & Brookings Institution (Eds.), *The price we pay: Economic and social consequences of inadequate education* (pp. 99-124). Washington, DC: Brookings Institution Press.

Rumberger, R. (2011). *Dropping out: Why students drop out of high school and what can be done about it*. Cambridge: Harvard University.

Rumberger, R., & Lim, S. (2008). *Why students drop out of school: A review of 25 years of research* (Report No.15). Santa Barbara, CA: California Dropout Research Project.

SAS® University Edition (n.d.). Software Download. Retrieved from: https://www.sas.com/en_us/software/university-edition/download-software.html

Schubert, H. (2009). Foreword. In C.L. Carger, *Dreams deferred* (pp. xx-xi). Charlotte, NC: Information Age.

Smith, E. (2008). *Using secondary data in educational and social research*. Maidenhead, England: McGraw-Hill Education.

Sprinthall, R. C. (2007). *Basic statistical analysis* (8th ed.). Boston, MA: Allyn and Bacon.

Stearns, E., & Glennie, E. J. (2006). When and why dropouts leave high school. *Youth & Society, 38*(1), 29-57. doi: 10.1177/0044118X05282764

Stuit, D., O'Cummings, M., Norbury, H., Heppen, J., Dhillon, S., Lindsay, J., & Zhu, B. (2016). *Identifying early warning indicators in three Ohio school Districts* (REL 2016-118). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Midwest. Retrieved from http://ies.ed.gov/ncee/edlabs

Tanner, D. (2012). *Using statistics to make educational decisions.* Los Angele, CA: Sage.

Thorsen, C. (2014). Dimensions of norm-referenced compulsory school grades and their relative importance for the prediction of upper secondary school grades. *Scandinavian Journal Of Educational Research*, *58*(2), 127-146. doi:10.1080/00313831.2012.705322

Thorsen, C. & Cliffordson, C. (2012). Teachers' grade assignment and the predictive validity of criterion-referenced grades. *Educational Research and Evaluation, 18*(2), 153-172. doi: 10.1080/13803611.2012.659929

Tingle, L. R., Schoeneberger, J., & Algozzine, B. (2012). Does grade retention make a difference? *Clearing House: A Journal of Educational Strategies, Issues and Ideas*, *85*(5), 179-185. doi: 10.1080/00098655.2012.679325

U.S. Department of Education (n.d.). *High school graduation initiative also known at school dropout prevention program*. Retrieved from: https://www2.ed.gov/programs/dropout/index.html

U. S. Department of Education, International Affairs Office. (2008). *Structure of the U.S. education system: Credit systems.* Retrieved from U.S. Department of Education: https://www2.ed.gov/about/offices/list/ous/international/usnei/us/credits.doc

University of Chicago. (n.d.). Rigor and readiness [Web log post]. Retrieved from: https://consortium.uchicago.edu/page/rigor-and-readiness

Vivo, J., & Franco, M. (2008). How does one assess the accuracy of academic success predictors? ROC analysis applied to university entrance factors. *International Journal of Mathematical Education in Science and Technology, 39*(3), 325-340. doi: 10.1080/00207390701691566

Waldfogel, J., Garfinkel, I., & Kelly B. (2007). Welfare and the costs of public assistance. In C. Banfield, H.M. Levin, & Brookings Institution (Eds.), *The price we pay : Economic and social consequences of inadequate education* (pp. 169-174). Washington, DC: Brookings Institution Press.

Willingham, W. W., Pollack, J. M., & Lewis, C. (2002). Grades and test scores: Accounting for observed differences. *Journal of Educational Measurement,39*(1), 1-37.

Zou, K. H., O'Malley, A. J., & Mauri, L. (2007). Receiver-operating characteristic analysis for evaluating diagnostic tests and predictive models. *Statistical Primer for Cardiovascular Research.* Retrieved from http://circ.ahajournals.org/

APPENDICES

Appendix A: Human Subjects Research Committee Proposal

GEORGE FOX UNIVERSITY HSRC INITIAL REVIEW QUESTIONNAIRE **Page** 1

****NOTE:** Review carefully the full text of the Human Subjects Research Committee Policies and Procedures.

Date submitted: March 16, 2018 Date received: ______

GEORGE FOX UNIVERSITY
Human Subjects Research Committee

PROTECTION OF HUMAN SUBJECTS INITIAL REVIEW QUESTIONNAIRE

[Note: Dissertation, or other formal research proposal, need not be submitted with this form. However, relevant section(s) may need to be attached in some cases, in addition to filling out this form completely, but only when it is not possible to answer these questions adequately in this format. Do not submit a proposal in lieu of filling out this form.]

Title of Proposed Research:Identifying Students at Risk of Dropping Out: Indicators and Thresholds Using ROC Analysis
Principal Researcher(s): Susan E. Carlson
Degree Program: Doctor of Education
Rank/Academic Standing: Good
Other Responsible Parties (if a student, include faculty sponsor; list other involved parties and their role)
Dane Joseph, PhD., Susanna M. Thornhill, PhD., Linda Samek, EdD.

(****Please include identifying information on page 6 also**.)

(1) Characteristics of Subjects (including age range, status, how obtained, etc)

This research will use publicly available data collected by the National Center for Educational Statistics (NCES). The subjects in this research were selected to participate in the High School Longitudinal Survey of 2009 (HSLS:09) by a two-stage stratified cluster sampling design. The sample consists of 25,206 ninth-grade students from 944 public and private high schools in the United States who were surveyed in the fall of the 2009-2010 school year. Their transcripts were collected in fall of 2013.

(2) Describe Any Risks to the Subjects (physical, psychological, social, economic, or discomfort/ inconvenience):

The data has been collected by the NCES and made available for public use. Downloading the data poses no risk to the subjects.

GEORGE FOX UNIVERSITY HSRC INITIAL REVIEW QUESTIONNAIRE **Page 2**

(3) Are the risks to subjects minimized (i) by using procedures which are consistent with sound research design and which do not unnecessarily expose subjects to risk, and (ii) whenever appropriate, by using procedures already being performed on the subjects for diagnostic or treatment purposes? ☒Yes/☐No

Degree of risk: 1 (low)

(4) Briefly describe the objectives, methods and procedures used:

The objective of using this secondary data is to assess the accuracy the ninth-grade indicators of GPA, Credits Earned, and Over-age Status in predicting the liklihood that a student will drop out of high school. I will use Radio Operating Characteristic analysis on data downloaded from the public use files on the NCES HSLS:09 website. The NCES safeguards the identity of students in the public use files by using disclosure treatement methods such as variable recoding, suppressing and swapping. High risk variables are suppressed from the public use files. Before gaining access to the data, I must sign a confidentiality agreement in which I agree to protect the identity of the study participants and schools in the event their identities become known to me during my analysis.

(5) Briefly describe any instruments used in the study (**attach a copy of each**).

I will not be using any instruments to collect data. I will be downloading existing data for secondary analysis.

(6) How does the research plan make adequate provision for monitoring the data collected so as to insure the safety, privacy and confidentiality of subjects?

My research plan protects the safety, privacy and confidentiality of the participants by utilizing the public use data in which the NCES removed or suppressed all identifying information.

(7) Briefly describe the benefits that may be reasonably expected from the proposed study, both to the subject and to the advancement of scientific knowledge – are the risks to subjects reasonable in relation to anticipated benefits?

The proposed study will add to educational research on the accuracy of early warning indicators of potential high school dropouts. Accurate indicators will help teachers and school staff to identify ninth-grade students in need of intervention.

(8) Where some or all of the subjects are likely to be vulnerable to coercion or undue influence (such as children, persons with acute or severe physical or mental illness, or persons who are economically or educationally disadvantaged), what appropriate additional safeguards are included in the study to protect the rights and welfare of these individuals?

NCES safeguraded the rights and welfare of all individuals in the study. They excluded students with physical limitations, cognitive disabilities, or limited English proficiency. They protected the identities of all students including economically and educationally disadvantaged students.

(9) Does the research place participants "at risk?" ☐Yes/☒No If so, describe the procedures employed for obtaining **informed consent** (***in every case, attach copy of informed consent form; if none, explain***).

Appendix B: Human Subjects Research Committee Approval

GEORGE FOX UNIVERSITY HSRC INITIAL REVIEW QUESTIONNAIRE **Page 6**

Title: Identifying Students at Risk of Dropping Out: Indicators and Thresholds Using ROC Analysis

Principal Researcher(s): Susan E. Carlson

Date application completed: March 16, 2018

(The researcher needs to complete the above information on this page)

COMMITTEE FINDING: For Committee Use Only

☒ (1) The proposed research makes adequate provision for safeguarding the health and dignity of the subjects and is therefore approved.

☐ (2) Due to the assessment of risk being questionable or being subject to change, the research must be periodically reviewed by the **HSRC** on a basis throughout the course of the research or until otherwise notified. This requires resubmission of this form, with updated information, for each periodic review.

☐ (3) The proposed research evidences some unnecessary risk to participants and therefore must be revised to remedy the following specific area(s) on non-compliance:

☐ (4) The proposed research contains serious and potentially damaging risks to subjects and is therefore not approved.

Winston Seegobin 03/21/2018

Chair or designated member Date

Appendix C: National Center for Education Statistics Data Usage Agreement

NCES Data Usage Agreement

Secure | https://nces.ed.gov/ncesglobal/data_usage_agreement.asp?li...

Print Page

NCES DATA USAGE AGREEMENT

Under law, public use data collected and distributed by the National Center for Education Statistics (NCES) may be used only for statistical purposes. Any effort to determine the identity of any reported case by public-use data users is prohibited by law. Violations are subject to Class E felony charges of a fine up to $250,000 and/or a prison term up to 5 years.

NCES does all it can to assure that the identity of data subjects cannot be disclosed. All direct identifiers, as well as any characteristics that might lead to identification, are omitted or modified in the dataset to protect the true characteristics of individual cases. Any intentional identification or disclosure of a person or institution violates the assurances of confidentiality given to the providers of the information. Therefore, users shall:

- Use the data in any dataset for statistical purposes only.
- Make no use of the identity of any person or institution discovered inadvertently, and advise NCES of any such discovery.
- Not link any dataset with individually identifiable data from other NCES or non-NCES datasets.

To proceed you must signify your agreement to comply with the above-stated statutorily based requirements. This window will close and you can now download the file.

I agree to the terms above.

I do not agree. Close window.