2018

The Predictive Validity of the Oregon Kindergarten Assessment on Smarter Balanced Assessment Mathematics and English Language Arts Scores

Kelsey I. Weber

The Predictive Validity of the Oregon Kindergarten Assessment on Smarter Balanced Assessment Mathematics and English Language Arts Scores

by

Kelsey I. Weber

Faculty Research Committee

Chair: Susanna Thornhill, Ph.D

Committee Member: Scot Headley, Ph.D

Committee Member: Dane Joseph, Ph.D

Presented to the Faculty of the

Doctoral Department of the College of Education

George Fox University

in partial fulfillment for the degree of

DOCTOR OF EDUCATION

April 5, 2018

GEORGE FOX UNIVERSITY

COLLEGE OF EDUCATION

"THE PREDICTIVE VALIDITY OF THE OREGON KINDERGARTEN ASSESSMENT ON SMARTER BALANCED ASSESSMENT MATHEMATICS AND ENGLISH LANGUAGE ARTS SCORES," a Doctoral research project prepared by KELSEY WEBER in partial fulfillment of the requirements for the Doctor of Education degree in Educational Leadership.

This dissertation has been approved and accepted by:

4.5.18	[signature]	Committee Chair
Date	Susanna Thornhill, PhD	Associate Professor of Education
4/5/18	[signature]	
Date	Scot Headley, PhD	Professor of Education
04/05/18	[signature]	
Date	Dane Joseph, PhD	Assistant Professor of Education

Abstract

The purpose of this study was to examine the predictive validity of the 2013 Oregon Kindergarten Assessment (OKA) early literacy and early math scores on 2016-2017 third-grade Smarter Balanced Assessment (SBAC) English language arts/literacy and math scores in a semi-urban school district in Oregon. The 3,200 participants were fourth-grade students in the 2017-2018 school year; this cohort of students were kindergarteners when the OKA was first administered during the 2013-14 academic year and third-graders during the 2016-2017 academic year. This was the first cohort of students to be assessed using both the OKA and the SBAC. This study used a multiple linear regression model to examine the relationship between OKA scores and third-grade SBAC scores. In addition, this study used a hierarchical regression model to examine the extent to which OKA scores interacted with students' ethnicity, socioeconomic status (SES), English learner (EL) status, classification for special education services, and gender to predict their third-grade SBAC achievement scores. The findings of this study suggest that OKA early literacy and early math scores predicted third-grade SBAC ELA and math achievement to a modest extent. The hierarchical regressions indicated that, even when interacting with demographic variables, OKA scores were only a modest predictor of third-grade SBAC outcomes. Findings from the hierarchical regressions demonstrated that models that only utilize the demographic variables of ethnicity, SES, EL status, classification for special education services, and gender have a stronger relationship with SBAC ELA and math scores than models that only utilize OKA early literacy or early math scores.

Acknowledgements

The completion of this dissertation would not be possible without the hard work of a village of supporters. I am humbled by the generosity and care that my colleagues and loved ones have shown me throughout the dissertation process.

I am grateful for the support of the entire George Fox community. I would like to thank my classmates at George Fox for their encouragement and support over the past three years of our doctoral studies. I would like to thank my committee for their insightful critique and thoughtful encouragement. Thank you, Susanna Thornhill, for your positivity in the face of setbacks. Thank you, Dane Joseph, for your invaluable support in data analysis.

I would like to thank my family for their constant support and love. Thank you to my daughter for taking regular naps this past year, allowing me the time I needed to write the many pages of this dissertation. I hope to one day support you in achieving your educational aspirations. I would like to thank my husband for inspiring me to study this topic and for being my loyal cheering section and sounding board. Thank you to my mother for being my personal editor; your edits and comments helped tremendously. Thank you to my father for checking the data analysis details; you made a big catch that significantly changed my analysis. Many thanks to my parents and parents-in-law for their willingness to provide childcare at many critical times throughout this past academic year.

Finally, I would like to thank my parents for their unfailing support of their children's education and for instilling in me the belief in the value and power of education. From a young age I knew that I wanted to follow in your footsteps to one day earn a doctorate; all children should be so lucky to have parents who inspire them to achieve.

TABLE OF CONTENTS

List of Tables and Figures

Chapter 1

Introduction

In her annual message to the school board, the superintendent of the Shepherd Public Schools (pseudonym) announced that, unfortunately, the district's third graders had again scored below the state average on the Smarter Balanced Assessment (SBAC) English language arts/literacy (ELA) assessment. On the state level, 47.4% of third graders passed the SBAC ELA assessment and 47.5% passed the SBAC math assessment (Oregon Department of Education, 2016a). In Shepherd, only 38.5% of third graders passed the SBAC ELA assessment and 40% passed the SBAC math assessment. These results indicated that only 38.5% of Shepherd students were reading proficiently at the end of third grade according to grade level Common Core State Standards (CCSS) proficiency measures. The superintendent emphasized the problem with these results by citing a study by Hernandez (2012). The superintendent quoted Sparks' (2011) summary of Hernandez's finding stating the following:

> Students who can't read on grade level by third grade are four times less likely to graduate on time than a child who does read proficiently. Add poverty to the mix, and a student is thirteen times less likely to graduate on time than his or her proficient, wealthier peers. (para. 3)

These assessment results were certainly disappointing; moreover, they were frightening for parents, educators, and community members who naturally considered the low third-grade ELA pass rate in tandem with the district's high poverty rate (61%, according to the district website). These results could forecast dismal future outcomes for the district's students. Soon after these results were shared with the school board, district leadership began to ponder strategies for early identification of students who were not on a trajectory to read proficiently by the end of third grade. The district leadership was seeking a data source that would help them to identify these students in the primary grades. This call for early identification brought to my mind the one universal screening tool used in K-2 education, the Oregon Kindergarten Assessment (OKA). I wondered, could the OKA serve as a predictive tool for academic performance in Shepherd Public Schools?

A Brief History of OKA and SBAC

OKA and SBAC are two assessments that have emerged in recent years in response to the education reforms of the Race to the Top initiative and adoption of the CCSS. As stated above, in the state of Oregon, there is only one universal assessment tool that is administered in the primary grades, the OKA. In response to the Race to the Top Early Learning Challenge, House Bill 4165, a piece of Oregon state legislation passed in 2012, established the state's early learning standards for children ages 3-5 years. As part of this legislation, the state's Early Learning Council was directed to select, pilot, and implement a state kindergarten assessment (House Bill 4165, 2012). The primary motivation behind this effort was to answer the following questions posed by McClelland and Squires (2012):

> Are Oregon's children (as a population) arriving at kindergarten ready for school? Is their level of school readiness improving or declining over time? Are there disparities (geographical, cultural, racial, and socio-economic) between groups of children's kindergarten readiness that must be addressed? Are there particular domains of school readiness that Oregon should target? (p. 2)

The leaders of the work group also noted the importance of a kindergarten assessment for predicting future outcomes, stating, "predictive validity is also important because a central aim of school readiness assessments is to assess skills at kindergarten entry that significantly predict third-grade reading and math skills" (p. 22). To accomplish this undertaking, a workgroup known as the Kindergarten Readiness Assessment Workgroup was formed to analyze current research on kindergarten readiness assessments and gather and analyze stakeholder input.

This workgroup offered two recommendations for a statewide kindergarten readiness assessment. They suggested two options, "(1) A composite assessment based on the Child Behavior Rating Scale (CBRS) and [Easy Curriculum Based Measures] easyCBM Literacy and Math measures; [or] (2) A portfolio assessment using the modified Teaching Strategies Gold adopted by the state of Washington" (McClelland & Squires, 2012, p. 1). The composite assessment was the preferred recommendation and was the selected option. This composite assessment came to be known as the Oregon Kindergarten Assessment (OKA).

OKA was piloted in 2012 and implemented statewide in 2013 (Office of Teaching, Learning, and Assessment, 2017). At the time of this writing, predictive validity studies of OKA were not yet available in the literature, because the first cohort of students to take the OKA in the fall of 2013 only recently underwent their first year of standardized testing, taking the third-grade SBAC assessments in the spring of 2017. Data became available to carry out predictive validity studies of the OKA on SBAC scores during the 2017–2018 school year.

Another response to the Race to the Top initiative was the creation of the Smarter Balanced Assessment (Smarter Balanced Assessment Consortium, n.d.b). In 2010, the Department of Education awarded the Smarter Balanced Assessment Consortium a grant to develop an assessment system that aligned with CCSS as a part of the Race to the Top initiative. The summative assessments that came out of the consortium were pilot tested in the spring of 2013, field tested in the spring of 2014, and made available to member states, including Oregon, in the spring of 2015. Before the creation of the SBAC, Oregon students were assessed using the

Oregon Assessment of Knowledge and Skills (OAKS). With the adoption of CCSS, the SBAC ELA and math assessments replaced the OAKS reading and math assessments. At the time of this writing, the OAKS assessment was still utilized to assess mastery of Oregon science standards, as CCSS only consist of ELA and math standards. At the time of this writing, Oregon students first took the SBAC as third graders.

The recent development of the OKA and SBAC, both implemented in Oregon within the past four years at the time of this writing, has opened up many areas of research around these assessments, including predictive validity studies. Both the OKA and SBAC tests will be discussed further in chapters 2 and 3.

Emphasis on Third Grade

One important point of clarification regarding this study is the decision to focus on third-grade results. Third grade is a pivotal year in a student's academic life. It is the turning point in which a student is no longer in the primary grades. It is also the first year that students typically begin standardized testing. While educators have always known that third grade is a pivotal year, research also confirms the importance of the third-grade year. A study of developmental reading stages found that by the end of third grade, students are transitioning from the "learning to read" stage to the "reading to learn" stage (Chall, 1983), meaning that in the latter stage students begin using their literacy skills to access information in texts. Without proficient literacy skills, students may not be able to access information that their more proficient peers are able to access.

This lack of reading proficiency by the end of third grade has been found to have a negative impact on future student outcomes, including academic achievement in later grades and high school completion (Barry & Reschly, 2012; Hernandez, 2012). In Hernandez's (2012) seminal work on the importance of third grade reading proficiency, he identified that by the end of third grade, if a student is not yet reading on grade level, that student is four times less likely to receive a high school diploma than a student who is reading on grade level. End-of-third-grade reading proficiency has thus become a highly-scrutinized predictor of educational outcomes. One approach to improving outcomes for students is to provide interventions for those students who do not achieve reading proficiency by the end of third grade as a way of remediating reading difficulties (Wanzek et al., 2017). Another approach to this problem is to provide intervention before the end of third grade for students who do not appear to be on track to reach proficiency by the end of that year as a way to preemptively resolve potential reading difficulties (Berkeley, Bender, Gregg, & Saunder, 2009). Early intervention for reading difficulties is critical, yet formalized testing opportunities to identify the students who are in need of intervention earlier than third grade have been limited.

Students typically first experience standardized testing at the end of the third-grade year. The educational field is seeking ways to discern prior to third grade whether students are on-track for reading success. This preemptive strategy calls into question how educators can determine which students will likely not be reading proficiently or achieving grade-level proficiency in math by the end of third grade. Some studies have explored possible indicators in the years leading up to third grade, many beginning in kindergarten (Duncan et al., 2007; Goldstein, McCoach, & Yu, 2017; Jordan, Glutting, & Watkins, 2010; Judge, 2017; Lloyd & Hertzman, 2009; Salvador, Schoeneberger, & Tingle, 2017). One of the newest assessments in Oregon, the OKA, is one such effort. The OKA is the only universal assessment tool that is administered in the primary grades (K-2) in Oregon. The aim of this study was to determine if the OKA had predictive validity for third-grade SBAC scores and could, therefore, be used as a predictive tool. At the time of this writing, there were no studies in the literature on the predictive potential of the OKA as it pertains to SBAC; as stated above, the OKA and the SBAC were new assessments, and results for cohorts of students who have taken both assessments had only recently become available.

Statement of the Problem

A low proportion of students in the Shepherd Public Schools were meeting grade level proficiency standards by the end of third grade. This proportion was markedly lower than Oregon's statewide proportion of third graders who were meeting proficiency standards. The Shepherd Public Schools leadership was seeking a source of data that would enable identification of students in the primary grades who were not on a trajectory to achieve grade level proficiency in ELA and math by the end of third grade.

Purpose

The purpose of this study was to investigate the predictive validity of the 2013 OKA early literacy and early math scores on 2016-2017 third-grade SBAC ELA and math scores in the Shepherd Public Schools. This study was quantitative in nature using data from the OKA as well as the SBAC assessments. This study was an *ex post facto* repeated measures study utilizing secondary data from one cohort of students in one school district in the state of Oregon. The sample included 3,259 participants. Of these participants, 76.8% were considered to be living in poverty according to free and reduced lunch status, 18.8% were classified as receiving special education services, and 36.1% were ELs or transitioned ELs. This was the first cohort of students to have been assessed using both the OKA and the SBAC; they first took the OKA in the fall of 2013 and later took the SBAC in the spring of 2017.

Research Questions

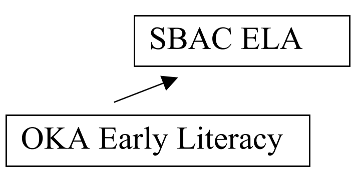
This study explored four research questions.

1. To what extent do Oregon Kindergarten Assessment (OKA) early literacy scores predict third-grade Smarter Balanced Assessment (SBAC) English language arts/literacy (ELA) achievement?
2. To what extent do OKA early math scores predict third-grade SBAC math achievement?
3. To what extent do OKA early literacy scores interact with students' ethnicity, socioeconomic status (SES), English learner (EL) status, classification for special education services, and gender to predict their SBAC ELA achievement scores?
4. To what extent do OKA early math scores interact with students' ethnicity, SES, EL status, classification for special education services, and gender to predict their SBAC math achievement scores?

The predictor variables that were examined to address the research questions are shown in Models 1-4b.

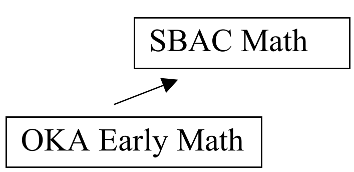
Model 1

Predictive variables of SBAC ELA Achievement



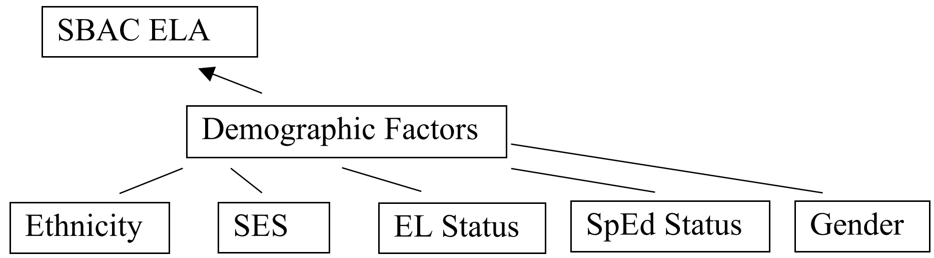
Model 2

Predictive variables of SBAC Math Achievement



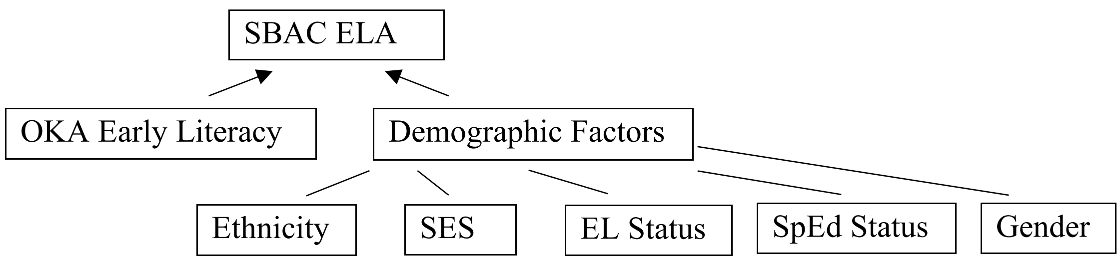
Model 3a

Demographic variables as predictors of SBAC ELA Achievement



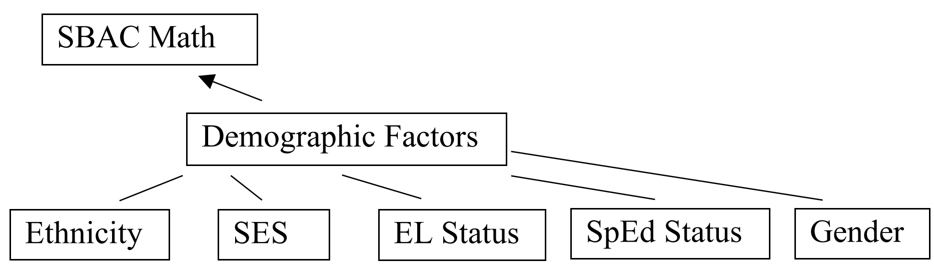
Model 3b

Demographic variables and OKA Early Litearcy scores as predictors of SBAC ELA Achievement



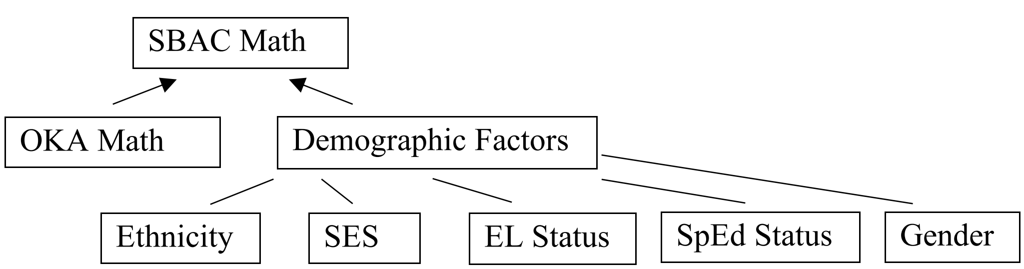
Model 4a

Demographic variables as predictors of SBAC Math Achievement



Model 4b

Demographic variables and OKA Early Math scores as predictors of SBAC Math Achievement



Significance/Contribution

The practical significance of this study is its functionality for informing early literacy intervention decisions for the Shepherd School District. It was the aim of the district leadership to identify students who were not on a trajectory to pass the third-grade SBAC with the intention of supporting those students with academic intervention. This study could offer new understandings of how to identify students in need of early intervention upon entering kindergarten. It was also in the interest of the school district to ensure that students who enter school on a trajectory to pass the third-grade SBAC continue on that trajectory; this study could be a useful tool in monitoring student progress. This research on the predictive validity of the OKA on third-grade SBAC results could also benefit Oregon teachers, principals, and school districts seeking a predictive tool for third-grade assessment outcomes.

Shepherd Public School District was a notable population to study as it had a high proportion of vulnerable populations of students. Students in poverty, ELs, and students receiving special education services are vulnerable populations (The Children's Hospital of Philadelphia PolicyLab, n.d.). According to the district website, Shepherd is one of the largest school districts in the state with over 40,000 students; 18% of students are ELs, well above the state EL proportion of 10%; 15% of Shepherd students receive special education services, above the state proportion of 13%, and 61% of Shepherd students are considered to live in poverty, well above the state proportion of 51% (Oregon Department of Education, 2016b). Shepherd has high proportions of these vulnerable populations. This study could offer new understandings of what variables may be useful to identify students from these vulnerable populations in need of early intervention. In addition, this study could give impetus for further study and replication in other districts.

The scholarly significance of this study is its addition to the volume of research on early childhood and elementary level assessments. Both OKA and SBAC were new assessment tools at the time of this writing; much research remained to be done on these assessments.

Limitations and Delimitations

One limitation of this study was that the sample only included one cohort of students. At the time of this writing, these were the only students who had taken both the OKA and the SBAC in the Shepherd district. In future years, the research could potentially be replicated for different cohorts of students as more data becomes available. In future years, it will also be possible to extend the research using SBAC results from fourth- through eighth-grade students as children who initially took the OKA continue to move through their elementary and middle school education. A second limitation is that this study utilized data from the first iteration of the OKA. The assessment has undergone changes since 2013 that may affect the predictive validity of the later iterations of the assessment. These changes to the assessment will be discussed further in Chapter 2.

A delimitation of this study was that the sample was drawn only from one school district. This narrowed the scope of the study to identify the predictive validity of the OKA specifically for the population of students in the Shepherd Public Schools. The findings of this study may also be informative for Oregon school districts with demographics similar to those of the Shepherd Public Schools. Another delimitation of this study was to limit the demographic variables that were utilized in the analyses to include ethnicity, SES, EL status, classification for special education services, and gender.

Definition of terms

The following terms are related to this study:

Common Core State Standards (CCSS):

Academic standards for end-of-grade knowledge and skills in English language arts/literacy (ELA) and math for grades K-12; adopted in 42 states

Kindergarten readiness assessment:

An assessment of behavioral and/or academic skills administered upon a child's entry into kindergarten; also known as a Kindergarten entry assessment

Oregon Kindergarten Assessment (OKA):

Kindergarten readiness assessment tool administered to all Oregon students enrolled in kindergarten within the first few weeks of the school year, beginning in 2013; this test includes three measures: early literacy, early math, and approaches to learning

Proficiency:

Meeting the standard of skill or knowledge that has been set; in the SBAC assessments, students who receive an achievement level score of 3 or 4 are considered to have achieved grade-level proficiency

Race to the Top:

United States Department of Education competitive grant fund designed to encourage and reward innovation in school improvement

Smarter Balanced Assessment (SBAC):

Standardized summative assessment aligned to the Common Core State Standards (CCSS) in English language arts/literacy and math for grades 3–8 and 11

Summary

Third-grade standardized assessment results are a focal point in K-12 education. As a way to identify students who were not on a trajectory to achieve proficiency on third-grade assessments, educational leaders in the Shepherd Public School District sought a source of data in the primary grades that could serve as a predictive tool for third-grade SBAC achievement. There is only one universal screening tool that, at the time of this writing, was used in the primary grades in the state of Oregon, the OKA. This study examined OKA scores as a predictor of SBAC results.

The next chapter is a review of the literature on the following topics: kindergarten readiness assessments, third-grade achievement assessments, and demographic influences on standardized test results. Chapter 3 is a description of the research methodology used in this quantitative study. Chapter 4 details the results of the data analysis. Chapter 5 is discussion of the findings of the study.

Chapter 2

Literature Review

The purpose of this literature review was to examine current research on kindergarten readiness assessments, third-grade achievement assessments, and demographic influences on standardized assessments.

Kindergarten Readiness Assessments

The entrance into kindergarten is a key point of transition in a child's education, moving from an early learning setting to the K-12 system (Goldstein, McCoach, & Yu, 2017). Teachers administer kindergarten readiness assessments at this point of transition for three reasons, first to examine early childhood experiences, to establish a baseline for instruction, and to predict future outcomes. This study focuses on the usefulness of assessment for predicting of future outcomes.

Kindergarten predictors of later achievement. Results vary on which early childhood indicators have been found to be the strongest predictors of later achievement. Some studies have found academic skills upon school entry to be the strongest predictors of later academic success (Claessens, Duncan & Engel, 2009; Cleassens & Engel, 2013; Duncan et al., 2007). Other studies have identified teacher-observed kindergarten behavior as a strong predictor of later achievement (McClelland, Acock, and Morrison, 2006).

Learning-related skills. McClelland, Acock, and Morrison (2006) examined the impact of learning-related skills on reading and math outcomes throughout elementary school. They operationalized learning-related skills to include self-regulation and social competence and assessed learning-related skills using the work-related skills subscale of the Cooper-Farran Behavioral Rating Scales (Cooper & Farran, 1991). They found that kindergarten learning-related skills "uniquely predicted academic performance through sixth grade" (p. 482). Their eight analyses of correlations between kindergarten learning-related skills and reading level from kindergarten through sixth grade ranged from .38 to .50 (*p*s < .05). Their eight analyses of correlations between kindergarten learning-related skills and math level from kindergarten through sixth grade ranged from .41 to .49 (*p*s < .05) (p. 478).

One critique of kindergarten readiness assessments is the question of whether observed behavior can actually be standardized or objectively documented as a predictor of achievement, mainly because these skills are often documented in narrative notes by a child's teacher. Rimm-Kaufman, Pianta, Cox, Carolina, and Hill (2000) noted the influence that teachers' perceptions have on the rating of children's behavior using such scales. They found that teachers' perceptions of behavioral problems were related to school minority composition and district poverty level. They also found that teachers' ethnicity showed a relationship to their rates of reported problems, with nonminority teachers reporting higher rates of behavior problems for high-minority student populations in comparison to minority teachers' rates of reporting. These criticisms of behavior measures have led to the decision to exclude the approaches to learning segment of the Oregon Kindergarten Assessment (OKA) from the analyses of this study. Thus, the focus of this study is the early literacy and early math measures of the OKA.

Academic skills. Several studies have found that early math skills of kindergarteners are predictive of later academic achievement. Claessens and Engel (2013) found that early math skills predict several achievement measures from Kindergarten through eighth grade, including reading, math, and science achievement as well as grade retention. In a study of the predictive power of incoming kindergarten academic and socioemotional skills and fifth-grade achievement, Claessens, Duncan, and Engel (2009) found that "rudimentary math skills were the single most important set of kindergarten-entry skills…. followed by reading skills, and finally attention skills, which were consistently predictive of both math and reading outcomes" (p. 423). In a meta-analytic regression of six data sets, Duncan et al. (2007) found that the strongest predictors of later achievement, which the authors framed as grade completion and reading and math achievement, were school-entry math, reading, and attention skills, with early math skills having the greatest predictive power ($\beta = .33$).

The ties between academic skills in Kindergarten and later achievement are not limited to math skills. An experimental study of intervention in phonological awareness in kindergarten found a potential association between kindergarten phonological intervention and reading achievement in third, sixth, and ninth grades (Kjeldsen, Kärnä, Niemi, & Olofsson, 2014).

Oregon Kindergarten Assessment. McClellan and Squires (2012), the leaders of the workgroup that developed the OKA, pointed to the importance of predicting future outcomes, stating, "predictive validity is also important because a central aim of school readiness assessments is to assess skills at kindergarten entry that significantly predict third-grade reading and math skills" (p. 22); this statement points to the importance of conducting a predictive validity study for the OKA.

As mentioned in chapter 1, OKA is a composite assessment utilizing measures from the Child Behavior Rating Scale (CBRS) and Easy Curriculum Based Measurement (easyCBM) Literacy and Math measures. The easyCBM segments are known as *early literacy* and *early math* segments, while the CBRS segment is known as the *approaches to learning* segment. The approaches to learning segment has 15 items that the teacher scores "based on observation of the student in the classroom during regular classroom activities and routines" (Office of Teaching, Learning, & Assessment, 2017, p. 20), using a five-point scale ranging from *never exhibits behavior* to *always exhibits behavior*. The OKA early literacy segment includes letter and sound recognition. The early math segment includes counting, simple addition, and simple subtraction, and is multiple choice with three possible answers per question.

Evolution of the OKA. As noted in Chapter 1, the OKA has undergone changes in the years since its 2013 statewide implementation (Office of Teaching, Learning, & Assessment, 2016 & 2017). These changes include the addition of a Spanish letter sounds measure for Spanish-speaking children in 2016, though this change was excluded from the 2017 iteration with the reasoning that there is "no direct alignment of the Spanish Letter Sounds to the Common Core State Standards" (Office of Teaching, Learning, & Assessment, 2017, p. 10). Another change introduced in 2016 was a provision for Spanish-speaking students, which allowed for students to hear directions in Spanish and offer answers back in Spanish for the math measure; this provision was still included in the 2017 iteration. Another change involved the addition of geometry, measurement and data, and counting and cardinality in the 2017 iteration of the early math measure. Other early literacy measures were piloted in the fall of 2015 in response to the original assessment's floor effect (Office of Teaching, Learning, & Assessment, 2016), though no additional literacy measures have been added to the state-wide assessment at this time.

The changes in the OKA over the years could be perceived as a complication of this study's design. Because of these changes, in future studies, cross-comparison of OKA scores across testing years will not be possible as the assessment has changed from year to year. Nevertheless, the current iteration of the OKA is likely not the permanent version. Just as the Spanish Letter Sounds segment was added to the assessment and then soon removed, reverting the assessment to the original iteration of the early literacy assessment, other changes to the assessment may also be removed. It is not farfetched to imagine that the assessment may revert back to the original iteration. It is worthwhile, therefore, to explore the predictive validity of the original iteration of the OKA, particularly as it pertains to SBAC scores, which are a significant measure of academic progress in the state of Oregon.

Third-Grade Achievement Assessments

Third-grade assessments are under a great deal of scrutiny for several reasons. One such reason is that third-grade assessments are the first summative standardized assessments that most students encounter. While historically, some state tests have included assessment of second graders, those states no longer include second-grade assessments. Critics of the second-grade assessments pointed to the inappropriateness of standardized testing of young children. Considering young children's needs and abilities, a standardized test that lasts multiple hours is not viewed as developmentally appropriate. The Smarter Balanced Assessment (SBAC) is estimated to require 4 hours of testing (Smarter Balanced Assessment Consortium, n.d.a), a lengthy assessment for young students. SBAC and Partnership for Assessment of Readiness for College and Careers (PARCC), the assessments currently used in 21 states, begin at the third-grade level. Another reason for the emphasis on third-grade results stems from the policies enacted under the No Child Left Behind Act (NCLB) of 2001 (No Child Left Behind, 2002). These accountability policies mandated standardized assessments beginning in third grade.

Findings of correlations between third-grade assessment results and future student outcomes are another reason for the scrutiny of third-grade assessment results. Barry and Reschly (2012) developed models utilizing third-grade indicators that significantly predict high school completion. In Hernandez's (2012) seminal work, he identified that by the end of third grade, if a student is not yet reading on grade level, that student is four times less likely to receive a high school diploma than a student who is reading on grade level. Further studies have found that students who are in poverty and not reading proficiently by the end of third grade are thirteen times less likely to graduate on time (Center for Public Education, 2015).

National Assessment of Educational Progress. The National Assessment of Educational Progress (NAEP) was first administered in 1969. In the nearly 40 years it has been administered, the assessment has undergone a few, well-documented changes. The consistency of the assessment over this time period has allowed NAEP results to present a picture of academic progress over the years (National Center for Education Statistics, n.d.). Students are assessed as incoming fourth, eight, and 12th graders; thus, the results are viewed as a measure of student proficiency upon exiting third, seventh, and 11th grades. NAEP results are published each year in a report named the Nation's Report Card.

SBAC. As noted in chapter 1, the development of the SBAC was a result of Race to the Top Assessment grant funds. The assessment was first pilot tested in May 2012 and administered statewide in Oregon in the 2014-2015 school year (Smarter Balanced Assessment Consortium, n.d.b). As SBAC was being developed, the creators aimed to closely align SBAC achievement levels with NAEP achievement levels (Smarter Balanced Assessment Consortium, n.d.c). As SBAC was first being introduced, it was expected that fewer students would score at high achievement levels on the new assessment than on older assessments such as state tests and NAEP. The 2015 SBAC scores, however, were significantly lower (as much as 17 percentage points) than NAEP scores (Lombardo, 2015).

Rationale for choice of SBAC. For many years, education researchers have studied and reported on NAEP results. Hernandez's (2012) study, on which much of the emphasis on third-grade results is based, involved NAEP results and not state tests or SBAC results. There are many reasons for focusing on SBAC over NAEP in this study. First, NAEP is not universally administered. It is administered to only a sample of districts and schools across the nation. While this data offers insights into the nation as a whole and identifies national trends over time, it does not offer insights for an individual district or school. NAEP is not an assessment tool that districts and schools can regularly utilize in their analyses of student progress. Even for schools and districts who are part of the sample, NAEP does not provide results for individual students or for their schools. For that reason, an analysis involving SBAC results for an entire district was feasible, while an analysis using NAEP results was not. Utilizing the SBAC in this study was also a choice designed to benefit Oregon school districts. SBAC is administered throughout the state. SBAC results are available for individual students. SBAC data is also readily available to Oregon school districts. Another reason for conducting this study using the SBAC rather than the NAEP was the reality of present educational assessment practices. Oregon's accountability measures at the time of this writing were based on SBAC results, not NAEP results; SBAC was the assessment that would be utilized as the accountability measure for Oregon's state plan under the Every Student Succeeds Act (ESSA) (Oregon Department of Education, 2017). For these reasons, this study focused on SBAC results.

Demographic Influences on Standardized Assessments

Demographic factors such as ethnicity, English learner (EL) status, and socio-economic status (SES) (as indicated by free and reduced lunch (FRL) status), classification for special education services, and gender have been found to be associated with differences in scores on various achievement tests. The term "achievement gap" refers to disparities in outcomes for different groups (Thurlow, Wu, Lazarus, Ysseldyke, 2016); many data sources aim to identify these disparities.

A study by Mulligan, Hastedt, and McCarroll (2012), using data from the ongoing Early Childhood Longitudinal Study for the kindergarten class of 2010-2011 conducted by the National Center for Education Statistics, found differences in reading and math scores among subgroups on a kindergarten assessment. The Programme for International Student Assessment (PISA), is an international survey for evaluating education systems worldwide that tests skills and knowledge of 15-year-old students in the areas of science, mathematics, reading, collaborative problem solving and financial literacy (Organisation for Economic Cooperation and Development, n.d.). The Organisation for Economic Cooperation and Development (OECD), an international organization, analyzes the PISA results and reports notable trends including differences among subgroups of students. Even though NAEP data is not a part of this study, it is still widely-administered and offers insights into to the marked differences on test results that are associated with demographic factors. Recent data from the National Center for Education Statistics (2013) found differences in achievement on the fourth-grade NAEP across subgroups. These differences are summarized in Tables 1 and 3 below. Data from the 2014-2015 SBAC (Smarter Balanced Assessment Consortium, 2016) also show differences across subgroups. These differences are summarized in Tables 2 and 4 below. Findings of difference across subgroups by ethnicity, EL status, SES, classification for special education services, and gender on these assessments are summarized below.

Ethnicity. It is a well-documented phenomenon that there are differences in assessment outcomes across subgroups by ethnicity. Coleman (1966) first brought light to disparities across ethnic groups; these disparities are still evident today. In their study of kindergarten assessment scores, Mulligan, Hastedt, and McCarroll (2012) the following test score differences across ethnicities:

- Asian students had higher scores than any other ethnicity.
- White students had higher scores than Black, Hispanic, Native Hawaiian/Pacific Islander, and American Indian/Alaska Native students.
- Black students had higher reading scores than Hispanic students.
- Multi-racial students had higher scores in reading and math than Black, Hispanic, and American Indian/Alaska Native students and higher reading scores than Native Hawaiian/Pacific Islander students.
- Native Hawaiian/Pacific Islanders had higher math scores than Hispanic students.

Recent data from the National Center for Education Statistics (2013) found differences in achievement on the fourth-grade NAEP across groups by ethnicity. These differences are summarized in Table 1. Data from the 2014-2015 SBAC also found differences in achievement on third-grade SBAC across groups by ethnicity. These differences are summarized in Table 2. The most notable trend in each of these data sets is that Asian and White students outperform all other subgroups on both reading and math.

Table 1
NAEP 2013 Achievement of Proficiency by Ethnicity

Ethnicity			
White	Black	Hispanic	Asian/Pac. Is.
Reading proficiency			
46%	18%	20%	51%
Math proficiency			
54%	18%	26%	64%

Note: Asian/Pac. Is. = Asian/Pacific Islander

Table 2
SBAC 2014-2015 Third-Grade Achievement of Proficiency by Ethnicity

Ethnicity						
Amer. Ind.	Asian	Black/Af.Am.	Haw./Pac. Is.	Hisp/Lat	White	Two/More
Reading proficiency						
23.3%	63.8%	25.1%	30.8%	26%	55.9%	41.4%
Math proficiency						
25.9%	70.6%	24.2%	33.9%	28.1%	57.7%	43.5%

Note: Amer. Ind. = American Indian or Alaska Native, Af.Am. = African American, Haw./Pac. Is. = Native Hawaiian or Pacific Islander, Hisp/Lat = Hispanic/Latino, Two/More – Two or More Races

EL status. Students proficient in the English language (non-EL students) tend to outperform ELs on assessments in the United States. In their study of kindergarten assessment scores, Mulligan, Hastedt, and McCarroll (2012) found differences across EL status groups with non-ELs having higher scores in reading and math than EL students. On the 2013 NAEP assessment, there were differences between EL students' and non-EL students' scores. In reading, 7% of EL students were found to be proficient while 38% of non-EL students were found to be proficient. In math, 14% of EL students were found to be proficient while 46% of non-EL students were found to be proficient. The 2014-2015 SBAC data does not offer a comparison between EL and non-EL students; however, it is possible to compare EL proficiency percentages with total student proficiency percentages. On the third-grade ELA assessment, 17.7% of ELs were found to be proficient while 42.3% of all students were found to be proficient. On the third-grade math assessment, 22.5% of ELs were found to be proficient while 44.5% of all students were found to be proficient. Statistical analyses would be necessary to determine if these differences on NAEP and SBAC are statistically significant, however, the discrepancy does appear to be large enough to suggest statistical significance. On the 2013 PISA assessment, it was found that immigrant students were over twice as likely as non-immigrant students to perform below the baseline level of proficiency in the area of science (Organisation for Economic Co-operation and Development, 2016). This is important information as immigrant students in the United States are, for the most part, EL students.

SES. SES has been found to play a large role in educational outcomes. Sirin (2005) concluded that poverty accounts for up to 60 percent of standardized test score variance. In their study using data from kindergarten assessments, Mulligan, Hastedt, and McCarroll (2012) found differences across SES groups, most prominently finding that students in households with incomes below the poverty level had the lowest scores. From the 2013 PISA results, OECD researchers found that socioeconomically disadvantaged students were nearly three times less likely than socioeconomically advantaged students to attain the baseline level of proficiency in the area of science (Organisation for Economic Co-operation and Development, 2016). The 2013 NAEP assessment results also showed achievement differences across socioeconomic subgroups. In reading, 18% of students eligible for free lunch were proficient, 30% of students eligible for reduced lunch were proficient, and 51% of students not eligible for free or reduced lunch were proficient. In math, 24% of students eligible for free lunch were proficient, 38% of students eligible for reduced lunch were proficient, and 59% of students not eligible for free or reduced lunch were proficient. The 2014-2015 SBAC data does not offer a comparison between SES groups; however, it is possible to compare proficiency percentages for economically disadvantages students to total student proficiency percentages. On the third-grade ELA assessment, 27.6% of economically disadvantaged students were found to be proficient while 42.3% of all students were found to be proficient. On the third-grade math assessment, 30.1% of economically disadvantaged students were found to be proficient while 44.5% of all students were found to be proficient. Statistical analyses would be necessary to determine if these differences on NAEP and SBAC are statistically significant, however, at face-value the difference in proficiency proportions suggests the difference is statistically significant.

Special education. Students receiving special education services have been found to have lower achievement results on standardized assessments than students not receiving special education services (Thurlow, Wu, Lazarus, & Ysseldyke, 2016). This trend is evident in the 2013 NAEP and 2014-2015 SBAC. On the 2013 NAEP reading assessment, 11% of students receiving special education services were proficient while 38% of students not receiving special education services were proficient. On the math assessment, 18% students receiving special education services were proficient while 45% of students not receiving special education services were proficient. The 2014-2015 SBAC data does not offer a comparison between students receiving special education services and students not receiving special education services; however, it is possible to compare proficiency percentages for students served under the Individuals with Disabilities Education Act (IDEA) to total student proficiency percentages. On the third-grade ELA assessment, 17.8% of students served under IDEA were found to be proficient while 42.3% of all students were found to be proficient. On the third-grade math assessment, 20.8% of students served under IDEA were found to be proficient while 44.5% of all students were found to be proficient. Statistical analyses would be necessary to determine if these differences on NAEP and SBAC are statistically significant, though the seemingly large difference in proficiency proportions suggests the difference is statistically significant.

Gender. Gender has been found to play a role in assessment outcomes. Notably, females have been found to have higher outcomes on reading assessments while males have been found to have higher outcomes on math assessments.

Reading. On the 2013 NAEP assessment, in the area of reading, 32% of male students were found to be proficient while 38% of female students were found to be proficient (National Center for Education Statistics, 2013). On the 2014-2015 third-grade SBAC ELA assessment, 46.7% of females were found to be proficient while 38% of males were found to be proficient. Statistical analyses would be necessary to determine if these differences on NAEP and SBAC results are statistically significant.

Math. In a review of literature on gender differences in math achievement, Forgasz (2010) summarized findings from the 2000 and 2003 PISA assessments. On the 2003 assessment, 70% of all countries in the study showed statistically significant (95% confidence interval) male advantage in math achievement, the United States being one of these countries. On the 2013 NAEP assessment, in the area of math, 43% of male students were found to be proficient while 41% of female students were found to be proficient (National Center for Education Statistics, 2013). On the 2014-2015 SBAC math assessment, 45.2% of males were found to be proficient while 43.7% of females were found to be proficient. Statistical analyses would be necessary to determine if these differences on NAEP and SBAC are statistically significant.

The findings of differences among demographic subgroups reveal the possibility that demographic factors may be significant predictors of third-grade SBAC achievement. For this reason, demographic factors were examined as predictors of SBAC achievement along with OKA scores.

Table 3
NAEP 2013 Achievement of Proficiency by EL Status, Classification for Special Education Services, and Gender, and Free and Reduced Lunch Status

EL Status		SpEd Status		Gender		Free and Reduced Lunch Status		
EL	Not EL	SpEd	NotSpEd	Male	Female	Free	Reduced	Not
Reading proficiency								
7%	38%	11%	38%	32%	38%	18%	30%	51%
Math proficiency								
14%	46%	18%	45%	43%	41%	24%	38%	59%

Note: SpEd = receiving special education services

Table 4
SBAC 2014-2015 Third-Grade Achievement of Proficiency by Limited English Proficiency Status, Classification for IDEA Services, and Gender, and Economic Disadvantage Status

				Gender	
Total	LEP	IDEA	EconD	Male	Female
Reading proficiency					
42.3%	17.7%	17.8%	27.6%	38%	46.7%
Math proficiency					
44.5%	22.5%	20.5%	30.1%	45.2%	43.7%

Note: LEP = Limited English Proficiency, IDEA = qualifying for services under the Individuals with Disabilities in Education Act, EconD = economic disadvantage status

Summary

This chapter investigated studies of kindergarten readiness assessments, third-grade assessments, and demographic influences on standardized test results. The examination of kindergarten readiness assessments reviewed kindergarten readiness measures as predictors of later achievement as well as the OKA. This examination revealed that early math and literacy measures have been found to be correlated with future school outcomes. The examination also revealed that behavior measures are critiqued as being potentially biased, leading to the decision to exclude the approaches to learning segment of the OKA from the analyses of this study and, instead, focus on the early literacy and early math measures of the OKA. The examination of third-grade assessments reviewed reasons for the emphasis on third-grade assessments and reviewed the NAEP and SBAC assessments. This examination also revealed the significance of third-grade assessments, indicating that an analysis using NAEP would not be beneficial or feasible. These findings led to the decision to focus on third grade SBAC results. The examination of demographic influences on standardized test results reviewed differences among demographic subgroups on various standardized assessments. This examination revealed that demographic factors may be significant predictors of assessment results, leading to the decision to examine demographic factors as predictors of SBAC achievement along with OKA scores.

The following chapter is a description of the research methodology used in this quantitative study. Chapter 4 details the results of the data analysis. Chapter 5 is discussion of the findings of the study.

Chapter 3

Methodology

This chapter details the research methods of this study. It includes information on the sampling plan, data collection, data analyses, role of the researcher, and ethical considerations of this study.

Research Design

This study was an *ex post facto* repeated measures study using secondary data. It was a predictive validity study looking at Oregon Kindergarten Assessment (OKA) scores as a predictor of Smarter Balanced Assessment (SBAC) scores.

This study focused on answering four research questions:

1. To what extent do OKA early literacy scores predict third-grade SBAC English language arts/literacy (ELA) achievement?
2. To what extent do OKA early math scores predict third-grade SBAC math achievement?
3. To what extent do OKA early literacy scores interact with students' ethnicity, socioeconomic status (SES), English learner (EL) status, classification for special education services, and gender to predict their SBAC ELA achievement scores?
4. To what extent do OKA early math scores interact with students' ethnicity, SES, EL status, classification for special education services, and gender to predict their SBAC math achievement scores?

Sampling Plan

The population that was examined is the Shepherd Public Schools (pseudonym) elementary student population in the 2016-2017 school year; the sample was 2016-2017 third grade students. At the time of this writing, the district population consisted of over 41,900 students, over 19,000 of which were in grades K-5. Within the Shepherd student population, 61% were considered to be living in poverty, 18% of students were identified as ELs, and 15% received special education services. According to the district website, proportions of students identifying as specific ethnicities were as follows: Asian – 2%, African American – 1%, Hispanic – 37%, American Indian/Alaskan Native – 1%, Multi-Ethnic – 5%, Pacific Islander – 2%, White – 52%.

The sample was a repeated measures sample of students in the Shepherd Public Schools. There were two inclusion criteria for participants to be included in the analysis. The first inclusion criterion was that participants must have been enrolled as third-grade students in the 2016-2017 school year, because students who were third graders in the 2016-2017 school year were the first group of students to have taken both the OKA and SBAC. The second inclusion criterion was that a record existed for either an OKA score or an SBAC math or ELA score. There was one exclusion criterion; students who took the Oregon extended assessment instead of the SBAC general assessment were excluded from the analysis. The Oregon Department of Education (n.d.) noted the following about the extended assessment:

> [The extended assessment was] designed specifically for students with the most significant cognitive disabilities…Because Extended Assessments are based on alternate achievement standards with content that is reduced in depth, breadth, and complexity, test results from these assessments are not comparable to results achieved on the state's general assessment, in spite of the similarity in performance category. (para. 1)

While ideally there would be a second exclusion criterion, which would be repetition of a grade level meaning that students who repeated Kindergarten would not be included in this sample, this information was not available in the data set. This was a limitation of the study that will be discussed further in chapter 5. The total sample size was 3,259 participants.

There were missing data, likely due to many factors such as students moving into the state of Oregon after the administration of the OKA during their kindergarten year, opting-out of SBAC testing, or illness resulting in missed test results. These missing data could affect the representativeness of the sample. For example, research indicates that students living in poverty are more likely to experience residential mobility (Herbers et al., 2012); the inclusion criteria may have led to a higher proportion of students in poverty being excluded from the analysis than students not living in poverty. These missing data could affect the power of the statistical analyses (Kang, 2013). I describe missing data analysis in the analysis section.

I chose Shepherd Public Schools as my sample for several reasons, the first being the convenience of sampling from a school district where I have access to the data needed for the analyses. The second reason is that Shepherd leadership was seeking a predictive tool for the performance of primary grade students; this study had the potential to be impactful for the school district. The third reason for choosing Shepherd was the unique population. As noted in chapter 1, according to the district website, at the time of this writing, Shepherd was one of the largest school districts in the state; 18% of students were ELs, well above the state EL proportion of 10%; 15% of Shepherd students received special education services, above the state proportion of 13%, and 61% of Shepherd students were considered to live in poverty, well above the state proportion of 51% (Oregon Department of Education, 2016b). Students in poverty, ELs, and students receiving special education services are vulnerable populations (Children's Hospital of Philadelphia PolicyLab, n.d.). Because Shepherd had high proportions of these vulnerable populations, it was an important population to study. This study could offer new understandings of which students from these vulnerable populations may be most in need of early intervention. This focus on a unique population in turn meant that there was low external validity, as the findings may not be generalizable to other populations with different proportions of population subgroups. This study, however, could give insight into the predictive validity of the OKA on SBAC results that could be meaningful for other school districts.

Instruments

As mentioned in Chapter 1, OKA was developed by McClelland and Squires (2012) using aspects of the Child Behavior Rating Scale (CBRS) (Bronson, Goodson, Layzer, & Love, 1990) and Easy Curriculum Based Measurement (easyCBM) (Alonzo, Tindal, Imer, and Glasgow, 2006) literacy and math assessments. As mentioned in Chapter 2, the easyCBM segments are known as *early literacy* and *early math* segments, while the CBRS segment is known as the *approaches to learning* segment. The approaches to learning segment has 15 items that the teacher scores "based on observation of the student in the classroom during regular classroom activities and routines" (Office of Teaching, Learning, & Assessment, 2017, p. 20), using a five-point scale ranging from *never exhibits behavior* to *always exhibits behavior*. The OKA early literacy measure includes letter and sound recognition; results have a total score range of 0-78. The OKA early math measure includes counting, simple addition, and simple subtraction, and is multiple choice with three possible answers per question; results have a score range of 0-16. The early literacy and early math segments are administered by teachers following policies and procedures outlined in Office of Teaching, Learning, and Assessment's (2017) test administration manual. The OKA scores used in this study include results in early literacy and early math; results from the approaches to learning segment of the OKA are not included in analyses. While validity testing of the OKA itself does not exist in the literature, there has been validity testing of easyCBM. In a study of the criterion-related validity of easyCBM, Lai (2013) found easyCBM to have correlations with comparator measures with a range of r =.39 to r =.86 for the kindergarten easyCBM reading measures. Comparator measures included the Dynamic Indicators of Basic Early Literacy Skills (DIBELS) (Good & Kaminski, 2002), The Comprehensive Test of Phonological Processing (CTOPP) (Wagner, 1999), and Test of Word Reading Efficiency (TOWRE-2) (Torgesen, Wagner, & Rashotte, 2012). Lai found that the kindergarten easyCBM had the strongest correlation with DIBELS (r =.86).

SBAC scores include scores in English language arts/literacy (ELA) and math. The Smarter Balanced Assessment was created by the Smarter Balanced Assessment Consortium. For the third-grade assessments, scores for each assessment were between 2300 and 2550. Student scores are then translated into achievement levels. There are four achievement levels:

1: Has *not met* the achievement standard and needs substantial improvement

2: *Nearly met* the achievement standard and may require further development

3: Has *met* the achievement standard

4: Has *exceeded* the achievement standard.

Table 5 shows the scaled scores for each achievement level for third-grade SBAC assessments. Data analysis for this study utilized participants' scores rather than achievement levels, though future predictive validity studies could utilize SBAC achievement levels; such a study would involve ordinal logistic regression.

Table 5
Third-grade SBAC Achievement Levels

ELA	Math
1. Not Met	
< 2367	< 2381
2. Nearly Met	
2367 – 2431	2381 – 2435
3. Met	
2432 – 2489	2436 – 2500
4. Exceeded	
> 2489	> 2500

In a report on the research agenda of the Smarter Balanced Assessment Consortium, Sireci (2012) laid out the framework validity studies of the SBAC, noting five sources of validity evidence: content validity, internal structure, relations with other variables, response processes, and testing consequences. A technical report that included findings of these studies using data from the 2014-2015 assessments was published in 2016 (Smarter Balanced Assessment Consortium, 2016). This report details the test construction procedures, administration procedures, scoring and scaling procedures that were undertaken in the development of the SBAC. It details a dimensionality study that was conducted in the pilot testing phase to "determine the factor structure of the assessments and the types of scales developed as well as the associated IRT [item response theory] models used to calibrate them" (p. 33). The report also includes information on reliability testing in the form of comparisons of marginal reliability of summative scores across subgroups of students. While analyses were completed at each grade level, I have focused on the analyses of third-grade scores. As a point of reference, the reliabilities for the SAT test usually range between 0.90 and 0.92 (Donlon & Agnoff, 1971), and a marginal reliability of 0.87 is "probably only marginally acceptable" (Sireci, Thissen, & Wainer, 1991). According to the SBAC technical report (Smarter Balanced Assessment

Consortium, 2016) for the math assessment of the SBAC, the marginal reliability for all subgroups ranged between 0.90 and 0.94. For the ELA assessment, lowest measures of marginal reliability were found for the subgroups American Indian/Alaska Natives (0.89), limited English proficient (LEP) (0.87), and students categorized as served under the Individuals with Disabilities Act (IDEA) (0.89). Marginal reliability for all other subgroups ranged between 0.90 and 0.92. The report pointed to a need for further studies of validity, noting the following:

> Usually, association with other assessment results requires a data set with a common set of either test items or examinees. The Consortium has not yet acquired a data set from its members so a study of this type is out of the scope of this manual. (p. 33)

This report plainly called for the need for ongoing research.

Data Collection

To conduct this study, I obtained information from a secondary data set. The Oregon Department of Education (ODE) collects student information including assessment and demographic data. These data can be obtained through a formal request. I requested student data for all students who were third graders in the Shepherd Public Schools in the 2016-2017 school year. I requested assessment and demographic data for these students. Assessment data included OKA scores and SBAC scores. Demographic data included ethnicity, SES, EL status, classification for special education services, and gender. In response to my request for data, the data governance committee from ODE suggested that I request the data directly from Shepherd Public Schools. I submitted a research proposal (see Appendix A) to the Shepherd Public Schools, which was approved; I received data from Shepherd Public Schools three weeks after requesting it.

Timeline

1. Obtained proposal approval – November 6

2. Submitted Institutional Review Board (IRB) form – November 7

3. Upon IRB approval, submitted request for data to ODE – November 28

4. Received notice from ODE to submit request to Shepherd Public Schools – January 31

5. Submitted request to Shepherd Public Schools – January 31

6. Received approval from Shepherd Public Schools – February 5

6. Received data set from Shepherd Public Schools – February 26

7. Data analysis using Statistical Packages for Social Sciences (SPSS) – February 27

8. Wrote results in chapter 4 – March

9. Wrote discussion and results in chapter 5 – March

10. Sent draft to dissertation chair for comments – March

11. Sent draft to dissertation committee –March

112. Defended dissertation – April 5

Data Analysis

Data analysis was completed using SPSS. The first data inspection was an exploratory descriptive analysis of the sample statistics for ethnicity, SES, EL status, gender, and special education classification. Though this analysis did not directly address a research question, it offered information on the overall composition of the sample.

I also conducted analyses of the missing data. First, I determined what proportion of the data was missing. Then I determined if the data was missing completely at random (MCAR), meaning that, “the probability that the data are missing is not related to either the specific value which is supposed to be obtained or the set of observed responses” (Kang, 2013, p. 402) using

Little's MCAR test (Little, 1988). Because the sample size was large, if data MCAR could be assumed, then listwise deletion of missing data, or deletion of all data that was missing, may have been a reasonable strategy (Kang, 2013), though pairwise deletion would also be option. If MCAR could not be assumed, I would have needed to utilize another method for handling the missing data, such as a pairwise deletion.

The following variables were included in the analyses. For the sake of quantitative analysis, categorical data was dummy-coded before being entered into the regression model.

Predictor Variables. The following variables were included as predictor variables in the analyses.

- Ethnicity – operationalized as a categorical variable based on groups defined by ODE including Asian, African American, Hispanic, American Indian/Alaskan Native, Multi-Ethnic, Pacific Islander, and White: Asian, African American, Hispanic, American Indian/Alaskan Native, Multi-Ethnic, or Pacific Islander (1), White (0).
- Socioeconomic status (SES) – operationalized as a categorical variable based on financial eligibility for free and reduced school lunch (FRL): yes (1), no (0).
- English learner (EL) status – operationalized as a categorical variable based on identification as an English learner or transitioned English learner – a student who previously was eligible for the English learner program but has been determined to have achieved English language proficiency: yes (1), no (0), transitioned (2).
- Classification for special education services – operationalized as a categorical variable based on identification as a student receiving special education services: yes (1), no (0).
- Gender - operationalized as a categorical variable based on male or female gender identification: female (1), male (0).

- OKA early literacy score – conceptualized as an incoming student's knowledge of literacy upon entering kindergarten; operationalized as a ratio variable based on the sub-score received on the early literacy portion of the OKA in September 2013: 0-78.
- OKA early math score – conceptualized as an incoming student's knowledge of math upon entering kindergarten; operationalized as a ratio variable based on the sub-score received on the early math portion of the OKA in September 2013: 0-16.

Criterion Variables. The following variables were included as predictor variables in the analyses.

- SBAC English language arts/literacy (ELA) score – conceptualized as an indicator of level of proficiency on the third-grade Common Core State Standards (CCSS) in the area of ELA; operationalized as an interval variable based on the score received on the SBAC ELA assessment in spring 2017: 2300-2500
- SBAC math score – conceptualized as an indicator of level of proficiency on the third-grade Common Core State Standards (CCSS) in the area of math; operationalized as an interval variable based on the score received on the SBAC math assessment in spring 2017: 2300-2550

In order to conduct the linear and hierarchical regressions, I operated under the following assumptions:

- One dependent variable measured at the continuous level – The dependent or criterion variable SBAC score was a continuous variable on the interval level (range 2300-2550).
- Two or more independent variables on the continuous or nominal level – There were several independent variables or predictor variables. Demographic variables of ethnicity, FRL status, EL status, classification for special education services, and gender were nominal variables.

OKA scores were continuous variables on the ratio scale (range 0-78 for early literacy and 0-16 for early math).

- Independence of errors – The independence of errors was inspected using the Durbin-Watson statistic.
- Linear relationship between the predictor variables and the criterion variable - The linear relationship was inspected by using partial regression plots between each continuous predictor variable and the criterion variable. In this analysis, there were only two continuous predictor variables, which were OKA early literacy score and OKA early math score.
- Homoscedasticity of residuals – Residuals were examined by plotting the studentized residuals against the unstandardized predicted values.
- No multicollinearity – Collinearity was examined through correlation coefficients and Tolerance/VIF values, as well as the correlation matrix values for the independent variables with the dependent variables. Values above r = .8 were flagged for possible collinearity.
- No significant outliers – Outliers were detected using casewise diagnostics.
- Errors are approximately normally distributed – Residuals were plotted on a histogram to verify approximate normal distribution (LAERD, n.d.).

Hypotheses

1. To what extent do OKA early literacy scores predict third-grade SBAC ELA achievement?

Ho_1: $\beta_1 = 0$, the coefficient of the OKA early literacy score slope was equal to zero

There was no significant prediction of third grade SBAC ELA scores by OKA early literacy scores.

H_1: $\beta_1 \neq 0$, the coefficient of the OKA early literacy score slope was not equal to zero

There was a significant prediction of third grade SBAC ELA scores by OKA early literacy scores.

I conducted a linear regression with OKA early literacy scores as a predictor of SBAC ELA scores. Through this analysis, I determined whether the linear regression between OKA early literacy and SBAC ELA was statistically significant ($p < .05$) and determined how much of the variation in the SBAC ELA was explained by OKA early literacy. Linear regression models the linear relationship between a predictor variable and a criterion variable where the predictor variable is predicting the criterion variable. With the predictor variable as X and the criterion variable as Y, a linear regression can be modeled through the following equation:

$$Y = \alpha + \beta X$$

In this equation α is the intercept and β is the slope coefficient. Predictions can be made based on the regression equation.

This first regression followed this model:

$$Y_1 = \alpha_1 + \beta_1 X_1$$

Y_1: SBAC ELA score

X_1: OKA early literacy score

β_1: slope coefficient of OKA early literacy score

α_1: intercept

2. To what extent do OKA early math scores predict third-grade SBAC math achievement?

H_{o2}: $\beta_2 = 0$, the coefficient of the OKA early math score slope was equal to zero

There was no significant prediction of third-grade SBAC math scores by OKA early math scores.

H_2: $\beta_2 \neq 0$, the coefficient of the OKA early math score slope was not equal to zero

There was a significant prediction of third-grade SBAC math scores by OKA early math scores.

I conducted a linear regression with OKA early math scores as a predictor of SBAC math scores. Through this analysis, I determined whether the linear regression between OKA early math and SBAC math was statistically significant ($p < .05$) and determined how much of the variation in the SBAC math was explained by OKA early math.

This second regression followed this model:

$$Y_2 = \alpha_2 + \beta_2 X_2$$

Y_2: SBAC math score

X_2: OKA early math score

β_2: slope coefficient of OKA early math score

α_2: intercept

3. To what extent do OKA early literacy scores interact with students' ethnicity, socioeconomic status (SES), English learner (EL) status, classification for special education services, and gender to predict their SBAC ELA achievement scores?

I carried out a hierarchical multiple-regression to assess the amount of variability in SBAC ELA scores that could be accounted for by OKA early literacy scores after controlling for demographic characteristics. Stockburger (2016) explained:

> Hierarchical regression adds terms to the regression model in stages. At each stage, an additional term or terms are added to the model and the change in R^2 is calculated. [A] hypothesis test is done to test whether the change in R^2 is significantly different from zero. ("Hierarchical Regression," para. 1)

With the predictor variables as X_1 through X_6 and the criterion variable as Y, a hierarchical regression can be modeled through the following two equations:

$$Y= \alpha + \beta_1X_1 + \beta_2X_2 + \beta_3X_3 + \beta_4X_4 + \beta_5X_5$$

$$Y= \alpha + \beta_1X_1 + \beta_2X_2 + \beta_3X_3 + \beta_4X_4 + \beta_5X_5 + \beta_6X_6$$

In a predictive validity study of the DIBELS assessment, Salvador, Schoeneberger, and Tingle (2017) controlled for demographic characteristics by first entering demographic variables in the first step in a hierarchical regression. For the second step in the hierarchical regression they added the DIBELS assessment scores variable to the regression. This two-step process allowed the researchers to first determine the amount of variance that demographic variables accounted for (31%) then determine the additional amount of variance that could be accounted for when including the assessment scores variable to the regression (an additional 23%).

In this analysis, variables were entered into the regression in two steps:

1. Including ethnicity, FRL status, EL status, classification for special education services, and gender
2. Adding OKA early literacy scores.

This two-step process allowed me to first determine the amount of variance in SBAC ELA scores that demographic variables accounted for (r^2_a), then determine the additional amount of variance that could be accounted for when including OKA early literacy scores as an additional variable in the regression (r^2_b- r^2_a), and if that additional amount of variance that could be accounted for was statistically significant ($p < .05$). I also determined the standard error of the estimate or standard error of the regression in each model.

The hierarchical regression followed this model:

$$Y_1 = \alpha_3 + \beta_{3a}X_a + \beta_{3b}X_b + \beta_{3c}X_c + \beta_{3d}X_d + \beta_{3e}X_e$$

$$Y_1 = \alpha_3 + \beta_{3a}X_a + \beta_{3b}X_b + \beta_{3c}X_c + \beta_{3d}X_d + \beta_{3e}X_e + \beta_{3x}X_1$$

Y_1: SBAC ELA score

X_a: ethnicity

X_b: FRL status

X_c: EL status

X_d: classification for special education services

X_e: gender

X_1: OKA early literacy score

β_{3a}: slope coefficient of ethnicity

β_{3b}: slope coefficient of FRL status

β_{3c}: slope coefficient of EL status

β_{3d}: slope coefficient of classification for special education services

β_{3e}: slope coefficient of gender

β_{3x}: slope coefficient of OKA early literacy score

α_3: intercept

4. To what extent do OKA early math scores interact with students' ethnicity, SES, EL status, classification for special education services, and gender to predict their SBAC math achievement scores?

I carried out a hierarchical multiple-regression to assess the amount of variability in SBAC math scores that could be accounted for by OKA early math scores after controlling for demographic characteristics. Variables were entered into the regression in two steps:

1. Including ethnicity, SES, EL status, classification for special education services, and gender
2. Adding OKA early math scores.

This two-step process allowed me to first determine the amount of variance in SBAC math scores that demographic variables accounted for (r^2_c), then determine the additional amount of variance that could be accounted for when including OKA early math scores as an additional variable in the regression ($r^2_d - r^2_c$), and if that additional amount of variance that could be accounted for was statistically significant ($p < .05$). I also investigated the standard error of the estimate or standard error of the regression in each model.

The hierarchical regression followed this model:

$$Y_2 = \alpha_4 + \beta_{4a}X_a + \beta_{4b}X_b + \beta_{4c}X_c + \beta_{4d}X_d + \beta_{4e}X_e$$

$$Y_2 = \alpha_4 + \beta_{4a}X_a + \beta_{4b}X_b + \beta_{4c}X_c + \beta_{4d}X_d + \beta_{4e}X_e + \beta_{4x}X_2$$

Y_2: SBAC math score

X_a: ethnicity

X_b: SES

X_c: EL status

X_d: classification for special education services

X_e: gender

X_2: OKA early math score

β_{4a}: slope coefficient of ethnicity

β_{4b}: slope coefficient of SES

β_{4c}: slope coefficient of EL status

β_{4d}: slope coefficient of classification for special education services

β_{4e}: slope coefficient of gender

β_{4x}: slope coefficient of OKA early math score

α_4: intercept

Role of the Researcher

At the time of this writing, I was a graduate student at George Fox University conducting this research as a requirement for the completion of a doctoral degree. I was also a kindergarten teacher; I administered the OKA to kindergarten students in September of 2017. In my career, I had worked closely with approximately 25 of the 3,374 participants. I had a personal interest in the findings of this study. I took precautions to be sure there was no way I would be able to identify individual students' scores, which I outline in the next section.

Ethics

The data for this study included existing secondary data. I, however, requested and received IRB approval as the data in this study included some personal information, such as individual test scores. To obtain the data for this study, I submitted a formal request to Shepherd Public Schools. In my request for data from Shepherd Public Schools, I requested that data be de-identified of names and student identification numbers. I stored all data on an encrypted drive and completed analyses on a password-protected computer. I stored the encrypted drive in a locked file cabinet. I will keep these files on the encrypted external drive in a locked cabinet for five years from completion of the study; retaining data for five years after the completion of research is a policy of George Fox University. At that point in time, the files will be deleted from the encrypted external drive.

I aimed to make ethical considerations in my discussion of the results of these analyses in chapter 5. I aimed to be cautious to not make sweeping policy recommendations from these results as the results may not be generalized to different contexts. I made efforts to clearly communicate the imprudence of interpreting the statistical results out of context. I recognized that the results of this study were unique; I aimed to be careful not to imply that policy should roll out from a single study; this study was the first of its kind for the particular school district; any results should be confirmed.

The following chapter details the results of the data analysis. Chapter 5 is discussion of the findings of the study.

Chapter 4

Results

The purpose of this study was to examine the predictive validity of the 2013 Oregon Kindergarten Assessment (OKA) early literacy and early math scores on 2016-2017 third-grade Smarter Balanced Assessment (SBAC) English language arts/literacy (ELA) and math scores in the Shepherd Public Schools (pseudonym). This study used a multiple linear regression model to examine the relationship between OKA scores and SBAC scores for the cohort of students who were kindergarteners in the Shepherd Public Schools during 2013-14 and third-graders in 2016-2017. In addition, this study used a hierarchical regression model to examine the extent to which OKA scores interacted with students' ethnicity, socioeconomic status (SES), English learner (EL) status, classification for special education services, and gender to predict their SBAC achievement scores. This chapter details the results of the data analyses. It includes information on the demographics of the sample, descriptive statistics, results of the analysis of missing data, results of the testing of assumptions associated with linear regression and hierarchical regression, and results of the linear and hierarchical regressions utilized to address the research questions.

Sample Demographics

The population that was examined was the Shepherd Public Schools elementary student population in the 2017-2018 school year; the sample was comprised of fourth-grade students during the 2017-2018 school year. The sample included 3,259 participants; 25.1% of participants were ELs, well above the district proportion of 18% and state proportion of 10%; 18.8% of participants were classified as receiving special education services, above the district proportion of 15% and state proportion of 13%; 76.8% of participants were considered to be living in poverty according to their free and reduced lunch (FRL) status, well above the district proportion of 61% and state proportion of 51%. Table 6 offers descriptive statistics for the sample. The sample had large proportions of vulnerable populations. These proportions demonstrate a trend of vulnerable populations increasing in the Shepherd School District.

Table 6
Sample Demographics

Demographic characteristic	Count of sample (Sample percentage)	District percentage	State percentage
Gender			
Male	1593 (48.9)		
Female	1662 (51.0)		
Ethnicity			
White	1548 (47.5)	50	62.9
Ethnic Minority	1711 (52.5)		37
Asian	39 (1.1)	2	3.9
African American	50 (1.4)	1	2.3
Hispanic	1378 (40.8)	39	22.6
Native American/Alaska Native	24 (0.7)	1	1.4
Multi-Ethnic	175 (5.1)	5	5.9
Pacific Islander	90 (2.6)	2	0.7
SES			
FRL	2504 (76.8)	61	51
Not FRL	751 (23.0)		
Special Education Classification			
SpEd	614 (18.8)	15	13
Not SpEd	2641 (81.0)		
EL Status			
Not EL	2084 (63.9)		
EL	818 (25.1)	18	10
Transitioned EL	357 (11.0)		

Note: SpEd = Special Education. Some students receiving special education services were excluded from the sample as they were assessed using the Oregon extended assessment, not the SBAC.

Descriptive Statistics

The mean score for the sample's OKA early literacy assessment was 17.26 (*M* = 17.26, *SD* = 21.353, meaning that on average, incoming kindergarteners could identify a total of 17.26 uppercase letters, lowercase letters, and letter sounds out of a possible total of 78. The mean score for the OKA math assessment was 7.07 (*M* = 7.07, *SD* = 3.446), meaning that on average, incoming kindergarteners answered 7.07 counting, simple addition, and simple subtraction, problems correctly out of a possible total of 16. The mean score for the sample's SBAC ELA was 2394.13 (*M* = 2394.13, *SD* = 88.165). The mean score for SBAC math was 2409.14 (*M* = 2409.14, *SD* = 86.163). For third grade SBAC assessments, scores range between 2300 and 2550; the mean scores for SBAC ELA and math were on the lower end of the score range. Table 7 is a summary of descriptive statistics for participant scores on the OKA and SBAC assessments. Appendix C offers more descriptive statistics for OKA and SBAC assessments.

Table 7
Descriptive Statistics for OKA and SBAC Assessments

Assessment	*M*	*SD*	*N*
OKA Early Literacy (0-78)	17.26	21.353	2485
OKA Math (0-16)	7.07	3.446	2487
SBAC ELA Score	2394.13	88.165	2485
SBAC Math Score	2409.14	86.163	2487

Missing Data

I determined what proportion of the data was missing, meaning the proportion of the sample for which an SBAC score, an OKA score, or demographic characteristic was missing. I found that over 23% of data was missing for OKA early literacy and early math scores. This high proportion and the nature of Shepherd Public School's vulnerable populations points to the probability that many of these students moved into the state after the OKA testing window took place during their kindergarten year and before the SBAC testing window three academic years later; these students, therefore, did not have OKA scores. This scenario is in line with Oregon's high rates of migration into the state in recent years (Bechtoldt, 2017). Alternatively, other students may have been living in the state of Oregon at the time of the OKA testing window but may have enrolled in school after the testing window. Late enrollment in kindergarten is an ongoing issue that school districts and schools are working to mitigate, as evidenced by public campaigns that urge families to register on time (Mutnomah County Kindergarten Counts, n.d.).

No data were missing for SBAC math and 0.2% of data was missing for SBAC ELA. These missing data indicates that some participants took the SBAC math assessment and not the ELA assessment. A likely explanation for these missing data is extended illnesses, which can prevent students from taking one or both assessments. Families also have the choice to opt out of assessments, which could also account for some of these missing data. It should be noted that scores were excluded for students who took the Oregon extended assessment rather than the general assessment. As noted in chapter 3, the extended assessment data was not included in the analysis as it is a different assessment designed for students with significant cognitive disabilities. In the categories of gender, SES, and special education status, only 0.01% of data was missing, which could be a result of human error in the recording of this information in the student information database. Table 8 is a summary of the descriptive statistics of the missing data.

Table 8
Missing Data Descriptive Statistics

				Missing	
	N	*M*	*SD*	Count	Percent
SBAC ELA	3252	2395.52	89.310	7	0.2
OKA Literacy	2487	17.25	21.347	772	23.7
SBAC Math	3259	2409.22	86.681	0	0.0
OKA Math	2490	7.07	3.448	769	23.6
Gender	3255			4	0.1
Ethnicity	3259			0	0.0
SES	3255			4	0.1
SpEd Status	3255			4	0.1
EL Status	3259			0	0.0

I then determined if the data was missing completely at random (MCAR) using Little's MCAR test (Little, 1988). Table 9 is a summary of the findings from Little's test of MCAR. Figure 1 and Figure 2 in Appendix B also summarize the findings. The test resulted in a χ^2 value = 15.794 (df = 8, p = .045). This p-value was below 0.05, meaning that I rejected the null hypothesis that data was MCAR, and I assumed that the missing data was not MCAR. The sample size was large and MCAR could not be assumed; therefore, I selected a strategy of pairwise deletion of the missing data. This resulted in a sample of 2,549 participants.

Table 9
Results of Little's Test of MCAR

	SBAC ELA	SBAC Math	OKA Literacy	OKA Math
	EM Means*			
	2395.40	2409.22	17.39	7.08

	EM Covariances*			
SBAC ELA	7984.905			
SBAC Math	6136.049	7513.657		
OKA Literacy	1062.459	941.821	459.416	
OKA Math	136.877	142.483	36.938	11.937

	EM Correlations*			
SBAC ELA	1			
SBAC Math	.792	1		
OKA Literacy	.555	.507	1	
OKA Math	.443	.476	.499	1

*Little's MCAR test: χ^2= 15.794, *df* = 8, *p* = .045

Assumptions

The following assumptions were tested in order to conduct the data analyses: independence of errors, linear relationship between the predictor variables and the criterion variable, homoscedasticity of residuals, no multicollinearity, no significant outliers, and errors are approximately normally distributed. All assumptions were met, with the exception of a slight violation of homoscedasticity of residuals for the regression for SBAC ELA.

Independence of errors. The independence of errors was inspected using the Durbin-Watson statistic. For the ELA hierarchical model, there was independence of residuals, as assessed by a Durbin-Watson statistic of 1.893 For the mathematics hierarchical model, there was independence of residuals, as assessed by a Durbin-Watson statistic of 1.821. Independence of errors could be assumed. (See Table 14 and Table 16 for Durbin-Watson statistics)

Linear relationship between the predictor variables and the criterion variable. The linear relationship was inspected by using partial regression plots between the continuous predictor variables (the OKA scores) and the criterion variables (the SBAC scores) (Figure 1 and Figure 2 in Appendix B). The figures show that there was a slightly positive, linear relationship between the OKA scores and SBAC scores. The linear relationship could be assumed.

Homoscedasticity of residuals. Residuals were examined by plotting the studentized residuals against the unstandardized predicted values (Figures 3 and Figure 4 in Appendix B). The absence of a systematic pattern would suggest that homoscedasticity of residuals was met. For the plot of studentized residuals against unstandardized predicated values for the SBAC ELA (Figure 3 in Appendix B), there was a slight reduction in spread of the plot points. This pattern suggests that this data was in violation of the assumption of homoscedasticity of residuals. According to Statistics Solutions (n.d.), however, "….the violation of homoscedasticity assumption must be quite severe in order to present a major problem given the robust nature of [ordinary least squares] regression" (para. 4). This violation appears to be slight as the spread of the plot points reduces very little as it moves across the predicted values. For the plot of studentized residuals against unstandardized predicated values for the SBAC math (Figure 4 in Appendix B), there was no reduction in spread of the plot points moving across the predicted values. This constant spread suggests that homoscedasticity of residuals was met for the SBAC math regression.

No multicollinearity. Collinearity was examined through correlation coefficients and Tolerance/VIF values, as well as the correlation matrix values for the independent variables with the dependent variables. These values are summarized in Table 18 and Table 19 in Appendix B. VIF statistics less than 10 and tolerance measures greater than 0.1

satisfy the assumption that independent variables do not highly correlate with other predictors. No VIF statistics were greater than 10 and no tolerance measures were less than 0.1, which satisfied the assumption of no multicollinearity.

No significant outliers. Outliers were detected using casewise diagnostics. In ELA only 8 cases were identified out of 2485 cases, a percentage of 0.3, which was well below the approximate proportion of cases that would be expected by chance. In math only 17 cases were identified out of 2487 participants, a percentage of 0.3, which was below the approximate proportion of cases that would be expected by chance.

Errors are approximately normally distributed Residuals were plotted on a histogram (Figures 5 and 6 in Appendix B) to verify approximate normal distribution. The histograms show that the residuals were approximately normally distributed. Approximate normality could be assumed.

For further detail of the results of the testing of assumptions, see Appendix B.

Research Questions

1. To what extent do OKA early literacy scores predict third-grade SBAC ELA achievement?

I conducted a linear regression with OKA early literacy scores as a predictor of SBAC ELA scores. Through this analysis, I determined whether the linear regression between OKA early literacy and SBAC ELA was statistically significant ($p < .05$) and determined how much of the variation in the SBAC ELA score could be explained by the OKA early literacy score.

The r value, or Pearson coefficient, for the correlation between OKA early literacy and SBAC ELA was 0.550, which is a moderate, positive correlation. The result $F(1, 2483)$, $p < 0.001$ from the ANOVA test indicates that this correlation is statistically significant ($p < .05$), so I rejected the null hypothesis that there is no correlation between OKA early literacy score and

SBAC ELA achievement. The adjusted r^2 value is an estimate of the proportion of the variance of the dependent variable that can be expected to be explained by the independent variable in the population. The adjusted r^2 value for this regression was 0.302. This was a modest r^2 value, indicating that 30.2% of the variability in SBAC ELA outcomes could be accounted for by OKA early literacy scores. While there was a correlation that was statistically significant, the moderate adjusted r^2 value suggests that OKA early literacy scores were a modest predictor of SBAC ELA outcomes. Table 10 and Table 11 summarize the results of the regression analysis.

Table 10
OKA Early Literacy to SBAC ELA Regression Analysis

R	r^2	Adjusted r^2	*SE* of the Estimate	Durbin-Watson
.550	.302	.302	73.675	1.834

Table 11
OKA Early Literacy to SBAC ELA ANOVA

	SS	*df*	*MS*	*F*	*p*
Regression	5830489.852	1	5830489.82	1074.140	.000
Residual	13477862.940	2483	5428.056		
Total	193085352.793	2484			

2. To what extent do OKA early math scores predict third-grade SBAC math achievement?

I conducted a linear regression with OKA early math scores as a predictor of SBAC math scores. Through this analysis, I determined whether the linear regression between OKA early math and SBAC math was statistically significant ($p < .05$) and determined how much of the variation in the SBAC math could be explained by OKA early math.

The r value for the correlation between OKA early math and SBAC math was 0.472, which is a moderate correlation. The result $F(1, 2489)$, $p < 0.001$ from the ANOVA test indicates that this correlation is statistically significant ($p < .05$), so I rejected the null hypothesis that there is no correlation between OKA early math scores and SBAC math achievement. The adjusted r^2 value for this regression was 0.223. This was a modest r^2 value, indicating that 22.3% of the variability in the SBAC math outcomes could be accounted for by the OKA early math scores. While there was a correlation that was statistically significant, the modest adjusted r^2 value suggests that OKA early math scores were a modest predictor of SBAC math outcomes. Table 12 and Table 13 summarize the results of the regression.

Table 12
OKA Early Math to SBAC Math Regression Analysis

R	r^2	Adjusted r^2	*SE* of the Estimate	Durbin-Watson
.472	.223	.223	76.058	1.707

Table 13
OKA Early Math to SBAC Math ANOVA

	SS	*df*	*MS*	*F*	*p*
Regression	4135538.775	1	4135538.775	714.889	.000
Residual	14392749.181	2488	5784.867		
Total	18528287.955	2489			

3. To what extent do OKA early literacy scores interact with students' ethnicity, SES, EL status, classification for special education services, and gender to predict their SBAC ELA achievement scores?

I carried out a hierarchical multiple-regression to assess the amount of variability in SBAC ELA scores that could be accounted for by OKA early literacy scores after controlling for demographic characteristics. This two-step process allowed me to first determine the amount of variance in SBAC ELA scores that demographic variables accounted for (R^2_a), then determine the additional amount of variance that could be accounted for when including OKA early literacy scores as an additional variable in the regression ($R^2_b - R^2_a$), and whether the additional amount of variance that could be accounted for was statistically significant ($p < .05$).

The *R* value for the first model in the hierarchical regression, which only included demographic variables, was 0.593, indicating a moderate, positive correlation between demographic variables and SBAC ELA scores. With the addition of OKA early literacy scores, the *R* value increased to 0.672, a moderately strong, positive correlation. The standard error of the first model was 71.109, a large standard error for an assessment with a 250-point range. The standard error decreased slightly in the second model to 65.371, though this was still a moderately large standard error suggesting that actual scores varied greatly from predicted scores. The adjusted R^2 value for the first model, or the amount of variance in SBAC ELA scores that demographic variables accounted for in the population, was 0.349, a modest R^2 value. With the addition of OKA early literacy scores as an independent variable the adjusted R^2 value increased to 0.450, a moderate R^2 value. The change in R^2, which is the additional amount of variance that could be accounted for when including OKA early literacy scores as an additional variable, was 0.101, $F(1, 2477)$, $p < .001$, which was statistically significant ($p < .05$). This change in R^2, though statistically significant, was moderately low and indicates that OKA early literacy scores moderately added to the strength of the model and were a not a strong predictor of SBAC ELA outcomes. Table 14 summarizes the results of this analysis.

Table 14
Hierarchical Regression for Demographic Variables and OKA Early Literacy on SBAC ELA Scores

					Change Statistics					
Model	*R*	R^2	Adjusted R^2	*SE* of Estimate	R^2 Change	*F* Change	*df*1	*df*2	*p* of *F* Change	Durbin-Watson
1	.593	.351	0.349	71.109	.351	223.420	6	2478	.000	
2	.672	.452	0.450	65.371	.101	65.470	1	2477	.000	1.893

An unstandardized coefficient is the amount that the dependent variable changes with a one-unit change of an independent variable. The data analysis also allowed me to determine the unstandardized coefficients for all of the independent variables, including gender, ethnicity, SES, classification for special education services, EL status, and OKA early literacy score, with the dependent variable SBAC ELA scores. All variables, including gender, ethnicity, SES, classification for special education services, having EL status, having EL transitioned status and OKA early literacy score were all found to have statistically significant ($p < .05$) unstandardized coefficients.

Gender was found to have a statistically significant unstandardized coefficient ($\beta = 8.801$, $p < .001$), showing a trend that, when classified as female, participants' SBAC ELA scores were on average 8.8 points higher than if they were classified as male. Ethnicity was found to have statistically significant unstandardized coefficient (($\beta = -10.558$, $p < .001$), showing a trend that, when classified as an ethnic minority, participants' SBAC ELA scores were on average 10.5 points lower than if they were classified as White. Classification for special education services was found to have significant unstandardized coefficient ($\beta = -37.068$, $p < .001$), showing a trend that, when classified as receiving special education services, participants' SBAC ELA scores were on average 37.0 points lower than if they had been classified as not receiving special education services. This finding has little meaning for predicting third-grade SBAC achievement from kindergarten intake information, as few students are identified to receive special education services upon entering kindergarten and are often identified in the years between kindergarten and third grade. The variable of SES was found to have a statistically significant unstandardized coefficient ($\beta = -34.381$, $p < .001$) showing a trend that, when classified as eligible for FRL, participants' SBAC ELA scores were on average 34.3 points lower than if they had been classified as not eligible for FRL. The variable of having EL status was found to have a statistically significant unstandardized coefficient ($\beta = -41.182$, $p < .001$), showing a trend that, when classified as an EL, participants' SBAC ELA scores were on average 41.1 points lower than if they had not been classified as an EL. The variable of having EL transitioned status was found to have a statistically significant unstandardized coefficient (($\beta = 40.967$, $p < .001$), showing a trend that, when classified as a transitioned EL, participants' SBAC ELA scores were on average 40.9 points higher than if they were not classified as a transitioned EL. OKA early literacy score was found to have a statistically significant unstandardized coefficient ($\beta = 1.540$, $p < .001$), showing a trend that, for each additional point on the OKA early literacy score, participants would receive on average 1.5 additional points on their SBAC ELA scores. Table 15 below summarizes the findings of the unstandardized coefficients.

Table 15
Coefficients for Independent Variables with the Dependent Variable SBAC ELA Score

	Unstandardized Coefficients		Standardized Coefficients		
	β	*SE*	β	*t*	*p*
(Constant)	2409.111	4.147		580.970	.000
Gender	8.801	2.650	.050	3.321	.001
Ethnicity	-10.558	3.435	-.060	-3.074	.002
SES	-34.381	3.628	-.163	-9.477	.000
SpEd	-37.068	3.445	-.168	-10.759	.000
EL Status Yes	-41.182	3.964	-.209	-10.388	.000
EL Status Transitioned	40.967	4.708	.153	8.702	.000
OKA Literacy (0-78)	1.540	0.072	.373	21.333	.000

N = 2567

4. To what extent do OKA early math scores interact with students' ethnicity, SES, EL status, classification for special education services, and gender to predict their SBAC math achievement scores?

I carried out a hierarchical multiple-regression to assess the amount of variability in SBAC math scores that could be accounted for by OKA early math scores after controlling for demographic characteristics. This two-step process allowed me to first determine the amount of variance in SBAC math scores that demographic variables accounted for (R^2_c), then determine the additional amount of variance that could be accounted for when including OKA early math scores as an additional variable in the regression (R^2_d- R^2_c), and if that additional amount of variance that could be accounted for was statistically significant ($p < .05$).

The *R* value for the first model in the hierarchical regression, which only included demographic variables, was 0.572, indicating a moderate, positive correlation between demographic variables and SBAC math scores. With the addition of OKA early math scores, the R value increased to 0.613, a moderately strong, positive correlation. The standard error of the first model was 73.295, a large standard error for an assessment with a range of 250 points. The standard error decreased in the second model to 68.149, still a moderately large standard error suggesting that actual scores varied greatly from predicted scores. The adjusted R^2 value for the first model, or the amount of variance in SBAC math scores that demographic variables account for in the population, was 0.276, a modest R^2 value. With the addition of OKA early math scores as an independent variable the adjusted R^2 value increased to 0.374, still a modest R^2 value. The change in R^2, which is the additional amount of variance that could be accounted for when including OKA early math scores as an additional variable, was 0.095, $F(1, 2479)$, $p < .001$, which is statistically significant ($p < .05$). Though statistically significant, this change in R^2 was moderately low and indicates that OKA early math scores moderately added to the strength of the model and were not a strong predictor of SBAC math outcomes. Table 16 summarizes the results of this analysis.

Table 16
Hierarchical Regression for Demographic Variables and OKA Early Math on SBAC Math Scores

					Change Statistics					
Model	*R*	R^2	Adjusted R^2	SE of the Estimate	R^2 Change	*F* Change	*df*1	*df*2	*p of F* Change	Durbin-Watson
1	.572	.278	.276	73.295	.278	159.261	6	2480	.000	
2	.613	.376	.374	68.149	.095	389.645	1	2479	.000	1.821

The data analysis also allowed me to determine the unstandardized coefficients for all of the independent variables, including gender, ethnicity, SES, classification for special education services, EL status, and OKA early math score, with the dependent variable SBAC math scores. All variables, including gender, ethnicity, SES, classification for special education services, having EL status, having EL transitioned status, and OKA early math score were found to have statistically significant ($p < .05$) unstandardized coefficients.

Gender was found the have a statistically significant unstandardized coefficient (β = -9.285, $p < .001$), showing a trend that, when classified as female, participants' SBAC math scores were on average 9.2 points lower than if they had been classified as male. Ethnicity was found to have statistically significant unstandardized coefficient ((β = -12.659, $p = .001$), showing a trend that, when classified as an ethnic minority, participants' SBAC math scores were on average 12.6 points lower than if they were classified as White. The variable of SES was found to have a statistically significant unstandardized coefficient (β = -39.682, $p < .001$), showing a trend that, when classified as eligible for FRL, participants' SBAC math scores were on average 39.6 points lower than if they had been classified as not eligible for FRL. Classification for special education services was found to have a statistically significant unstandardized coefficient (β = -45.357, $p <.001$), showing a trend that, when classified as receiving special education services, participants' SBAC math scores were on average 45.3

points lower than if they were classified as not receiving special education services. This finding has little meaning for predicting third-grade SBAC achievement from kindergarten intake information, as few students are identified to receive special education services upon entering kindergarten and are often identified in the years between kindergarten and third grade. The variable of having EL status was found to have a statistically significant unstandardized coefficient (β = -32.240, p <.001), showing a trend that, when classified as an EL, participants' SBAC math scores were on average 32.2 points lower than if they had not been classified as an EL. The variable of having EL transitioned status was found to have a statistically significant unstandardized coefficient ((β = 28.574, p <.001), showing a trend that, when classified as a transitioned EL, participants' SBAC math scores were on average 28.5 points higher than if they were not classified as a transitioned EL. OKA early math score was found to have a statistically significant unstandardized coefficient (β = 8.362, p <.001), showing a trend that for each additional point on the OKA early math scores participants would receive on average 8.3 additional points on their SBAC math scores. Table 17 summarizes the findings of the unstandardized coefficients.

Table 17
Coefficients for Independent Variables with Dependent Variable SBAC Math Score

	Unstandardized Coefficients		Standardized Coefficients		
	β	*SE*	β	*t*	*P*
(Constant)	2406.976	5.101		471.905	.000
Gender	-9.285	2.763	-.054	-3.361	.000
Ethnicity	-12.659	3.576	-.073	-3.540	.000
SES	-9.682	3.689	-.192	-10.729	.000
SpEd	-45.357	3.62	-.210	-12.733	.000
EL Status Yes	-32.240	4.042	-.168	-7.976	.000
EL Status Transitioned	28.574	4.862	.109	5.877	.000
OKA Math (0-16)	8.362	0.424	.334	19.739	.000

N = 2549

The following chapter is discussion of the findings of the study. It includes a discussion of the findings, implications of these findings, limitations of the study, and recommendations for further research.

Chapter 5

Conclusions

The purpose of this study was to examine the predictive validity of the 2013 Oregon Kindergarten Assessment (OKA) early literacy and early math scores on 2016-2017 third-grade Smarter Balanced Assessment (SBAC) English language arts/literacy (ELA) and math scores for a single cohort of students from Shepherd Public Schools (pseudonym). This chapter offers a discussion of the findings of the study. It includes a discussion of the findings, implications of these findings, limitations of the study, and recommendations for further research.

Discussion

This study explored four research questions to examine the predictive validity of the OKA on SBAC scores.

1. To what extent do OKA early literacy scores predict third-grade SBAC ELA achievement?

There was a moderate, positive correlation (r = 0.550) between OKA early literacy scores and SBAC ELA scores that was statistically significant ($p < .001$) and a modest adjusted r^2 value (0.302). The adjusted R^2 value for the first model of the ELA hierarchical regression, which only included demographic variables as independent variables, was 0.349. This adjusted R^2 value was higher than the adjusted r^2 value when looking at OKA early literacy scores as the only independent variable (r^2 value of .302). This comparison demonstrates that OKA early literacy scores alone had a weaker predictive power on SBAC ELA scores than demographic variables. These findings indicate that OKA early literacy scores were only a modest predictor of SBAC ELA scores.

2. To what extent do OKA early math scores predict third-grade SBAC math achievement?

There was a moderate, positive correlation (r = 0.472) between OKA early math scores and SBAC math scores that was statistically significant ($p < .001$) as well as a modest adjusted r^2 value (0.223). The adjusted R^2 value for the first model of the math hierarchical regression, which only included demographic variables as independent variables, was 0.276. This adjusted R^2 value was higher than the adjusted r^2 value when looking at OKA early math scores as the only independent variable (r^2 value of .223). This comparison suggests that OKA early math scores had a weaker predictive relationship on SBAC math scores than demographic variables. These findings suggest that OKA early math scores were predictive of third-grade SBAC math achievement to only a modest extent.

3. To what extent do OKA early literacy scores interact with students' ethnicity, socioeconomic status (SES), English learner (EL) status, classification for special education services, and gender to predict their SBAC ELA achievement scores?

With the addition of OKA early literacy scores as an independent variable in the hierarchical regression, the R value increased to 0.672, a moderately strong, positive correlation. The standard error decreased slightly in the second model of the hierarchical regression to 65.371; this was still a large standard error. With the addition of OKA early literacy scores as an independent variable, the adjusted R^2 value increased to 0.450, which is still a moderate R^2 value. The change in R^2 was 0.101, $p < .001$, which was statistically significant ($p < .05$). Though statistically significant, the change in R^2 was low, indicating that OKA early literacy scores moderately added to the strength of the model. The large standard error, the moderate R^2 value, and the moderately low change in R^2 in the hierarchical regression indicate that, when interacting with students' ethnicity, SES, EL status, classification for special education services, and gender, OKA early literacy scores were only a moderate predictor of third-grade SBAC ELA outcomes.

4. To what extent do OKA early math scores interact with students' ethnicity, SES, EL status, classification for special education services, and gender to predict their SBAC math achievement scores?

When OKA early math scores were added as an independent variable in the hierarchical regression, the R value increased to 0.613, a moderately strong, positive correlation. The standard error decreased slightly in the second model of the hierarchical regression to 68.149. When adding OKA early math scores as an independent variable in the hierarchical regression, the adjusted R^2 value increased to 0.374, still a modest R^2 value. The change in R^2 was 0.098, $p < .001$, which was statistically significant ($p < .05$). Though statistically significant, the change in R^2 was low, indicating that OKA early math scores moderately added to the strength of the model. The large standard error, the modest R^2 value, and the moderately low change in R^2 in the hierarchical regression indicate that, even when interacting with students' ethnicity, SES, EL status, classification for special education services, and gender, OKA early math scores were only a modest predictor of third-grade SBAC math outcomes.

On each of the four research questions, it was found that OKA scores predicted SBAC scores to only a modest or, at best, moderate extent. Several studies have found incoming kindergarten academic skills to be predictive of later reading and math achievement (Claessens & Engel, 2013; Claessens, Duncan, & Engel, 2009; Duncan et al, 2007). The findings of modest to moderate predictive validity of OKA scores on SBAC scores are in line with other findings in the literature. What is surprising is that the findings from the hierarchical regressions demonstrated that a model that only utilizes the demographic variables of ethnicity, SES, EL status, classification for special education services, and gender has stronger prediction of SBAC scores than a model that only utilizes OKA scores. According to the findings of this study, data on students' demographic information offer more predictive information on third-grade scores than OKA scores do.

The findings of this study raise concerns. Is an assessment that only has modest or, at best, moderate predictive power good enough? Is an assessment valuable if it offers less predictive information on students' future outcomes than demographic information can? Are the resources and time that are devoted to this assessment, which is administered in every kindergarten classroom across the state, well-spent? When I think back to why I chose to conduct this study – concern for students' future outcomes and the need to identify students in need of early intervention – I fear that moderate prediction is not good enough. The OKA was, in part, implemented with the intention of being predictive of third-grade scores, and it appears the assessment is not fulfilling that intention.

Implications

Recognizing that it is imprudent to interpret statistical results out of context and generalize beyond the limits of an initial study, I am cautious about making policy recommendations. Yet even before embarking on this study, I recognized that the results of this study could be useful; this study was the first of its kind, examining the relationships between these two prevalent assessment tools. Research resulting from small-scale studies have the potential to offer results that invite additional research and heighten awareness for educational leaders. The findings of this study point to several possible implications for the Shepherd School District, education policymakers, and classroom teachers.

Implications for Shepherd School District. The original impetus for this study was the Shepherd School District leadership's desire to find a data source that would assist with early identification of students who were not on a trajectory to read proficiently by the end of third grade. I wondered if the OKA could serve as a predictive tool for third-grade academic performance in Shepherd Public Schools. The findings of this study suggest that the OKA is a modest predictor of third-grade SBAC achievement for the Shepherd School District population. For this reason, the OKA likely should not serve as a source of data when aiming to identify students who are not on a trajectory to read proficiently by the end of third grade, as identified by third-grade SBAC performance. It is my recommendation that the Shepherd Public Schools should not rely on the OKA as a key tool when identifying these students earlier than the third grade and should seek other sources of information that have stronger predictive power. I recommend that the Shepherd Public Schools examine other existing sources of data, such as students' progress report scores and district reading assessment scores from the end of kindergarten, first, and second grade, to determine if those sources of data may have stronger relationships with SBAC outcomes than the OKA appears to have.

Implications for policymakers. As noted in Chapter 1, the leaders of the work group who developed the composite assessment that came to be known as the OKA remarked on the importance of a kindergarten assessment for predicting future outcomes, stating, "predictive validity is also important because a central aim of school readiness assessments is to assess skills at kindergarten entry that significantly predict third-grade reading and math skills" (McClelland and Squires, 2012, p. 22). The findings from this study suggest that, for the Shepherd Public Schools and possibly for school districts with populations similar to that of Shepherd, the OKA only modestly predicts third-grade reading and math skills when third-grade reading and math skills are defined as SBAC achievement; in fact, demographic variables were a stronger predictor. These findings suggest that educators may be administering the OKA under a false assumption that the OKA has stronger predictive validity than the instrument actually has.

The Oregon Kindergarten Assessment Specifications document from the Office of Teaching, Learning, and Assessment (2017) states that the approaches to learning component of the OKA has been found to be "strongly predictive of reading and math achievement in elementary grades" (p. 5); there is no information on the predictive validity of the early math or early literacy segments in the assessment specifications document. This claim of predictive validity potentially leads school districts and teachers to believe that the OKA measures have strong predictive validity for the reading and math outcomes that are important accountability measures in Oregon, specifically the SBAC. For that reason, it is my recommendation that more research be done on the predictive validity of the OKA. Upon the completion of this recommended research, documents about the OKA that are disseminated to school districts should have language that clearly states whether or not the OKA measures have been found to have low, moderate, or strong predictive validity with SBAC results. The OKA is a mandated assessment for every kindergarten teacher in the state; it is ODE's responsibility to provide research-based findings specific to OKA predictive validity, and to make those findings clear in OKA documentation.

Moreover, this study has revealed the need to research the reliability of the OKA and for those research-based findings to be clearly stated in OKA documents. There is no information on the reliability of the early math or early literacy segments in the OKA assessment specifications document. This is a topic that policymakers have the responsibility to explore. Assessment reliability is particularly relevant when considering several factors related to the test takers, including a) the wide variability in kindergarten students' developmental stages, b) students' trust in the test administrator, and c) students' moods and physical needs. These factors represent only a few of the many factors beyond student knowledge that can affect test results. Assessment outcomes can be affected by students' trust in the administrator, by students' mood at the time of administration, and by their physical needs at the time of test administration, such as hunger or tiredness. Assessment outcomes are also affected by students' relationship with the person who is administering the assessment, whether it be novice educators or experienced educators. The OKA is typically administered in the first week of kindergarten, at a time when some children are having their first school experience. The test can be administered any time within the first six weeks of the start of kindergarten, meaning outcomes could also be affected by the date within the testing window when the assessment is administered.

For all of these reasons, it is critical that the reliability of the OKA assessment be explored to determine if the assessment is, in fact, a reliable snapshot of students' incoming skills. Findings on the reliability of the OKA are a missing piece of essential information in determining the value of the OKA. If the OKA is found to be neither a strong predictor of future performance nor reliable in and of itself, its continued administration should be reevaluated.

Implications for classroom teachers. The findings of this study have implications for the teachers in the Shepherd School District and in districts with populations similar to those of Shepherd School District. Knowing that there is a moderate correlation between OKA scores and SBAC achievement, classroom teachers could utilize OKA results as a potential source of information when considering student intervention, but teachers should keep in mind the modest predictive power of the OKA. Though OKA results can give some information to teachers, teachers should trust their own professional judgment over OKA results. Ultimately, this means classroom teachers need to seek additional predictors and utilize additional sources of information when considering student intervention.

Limitations

One limitation of this study is that the sample only included one academic year's cohort of students. At the time of this writing, these were the only students who had taken both the OKA and the SBAC in the Shepherd School District. A second limitation is that this study utilized data from the first iteration of the OKA. The assessment has undergone changes since 2013 that may affect the predictive validity of the later iterations of the assessment. Changes to the early literacy measure have been limited; a Spanish early literacy segment for native Spanish speakers was added in the 2016-2017 iteration but later removed in 2017-2018 iteration. The only notable change for the early math segment of the OKA was the addition of geometry, measurement and data, and counting and cardinality questions in the 2017-2018 iteration. While it was only possible to conduct this study using the data from the first iteration of the OKA, in future years it will be possible to compare the predictive validity of the different iterations of the OKA on SBAC scores as more cohorts of students move through the grades taking both the OKA and SBAC. Another limitation of this study is that it focused on one specific school district. The findings of this study may not be meaningful for school districts with populations that differ greatly from the Shepherd Public Schools' population. As noted in Chapter 3, the study did not have information on student grade retention within the sample. Some participants may have repeated kindergarten, which could have affected assessment results somewhat. Without information on grade retention, it was not possible to exclude these participants, which could affect results. One final limitation was the violation of the assumption of homoscedasticity of residuals for SBAC ELA, which could affect the regression analyses making the coefficient estimates less precise.

Recommendations for Further Research

This research can and should be replicated in other districts with populations that differ from the Shepherd Public Schools; it would be worthwhile to examine whether the OKA has stronger predictive validity on third-grade SBAC scores in other contexts or districts. This research can also be replicated using statewide data to determine if the OKA has predictive validity on third-grade SBAC scores for the state as a whole.

Further research could use larger sample sizes that include students from multiple school districts. It will also be possible to replicate the research looking at the predictive validity of each iteration of the OKA. In future years, this research could be extended using SBAC results from fourth- through eighth-grade students, as children who initially took the OKA continue to move through their elementary and middle school education. In Chapter 3, I noted that data analysis for this study utilized participants' SBAC scores rather than SBAC achievement levels (1–4). Further predictive validity studies could utilize SBAC achievement levels as the criterion variable by employing ordinal logistic regression. Finally, while this study examined the predictive validity of the early literacy and early math measures of the OKA, it would be possible and advisable to examine the predictive validity of the "approaches to learning" segment of the OKA.

Future research of the predictive validity of the OKA could look at other academic outcomes, beyond SBAC results, to determine whether the OKA might be a predictor of those outcomes. The findings of this study also point to the need to research other data sources that potentially have stronger predictive validity on SBAC scores. This study suggests the OKA may be only a moderate predictor of SBAC scores in some school districts, including Shepherd School District. Therefore, school districts need additional tools that are a strong predictor of SBAC scores in their context.

As noted above, there is no information available on the reliability of the OKA; this is a topic that should be addressed in future research. There are many factors beyond student knowledge that can affect young children's test results, including trust in the administrator, mood at the time of administration, physical needs at the time of administration such as hunger or tiredness, and many more. Test outcomes could also be affected by who is administering the assessment, whether it be novice educators or experienced educators. There are administration protocols that are described in the assessment specifications document (Office of Teaching, Learning, and Assessment, 2017). However, these preparations are limited to an electronic slide presentation and a testing script, meaning there is a good possibility of differences in administration.

Beyond this, the OKA is administered within the first few weeks of the school year; in some cases, students take the OKA on their first day of school. This is often a time of difficult transition, and the tumultuous nature of that time could affect test outcomes. Outcomes could also be affected by the date within the OKA testing window when the assessment is administered, with scores possibly differing between students who were assessed on the first day of school and students assessed even a week later. If the OKA continues to be used, future research should explore the reliability of the OKA and the administration of the OKA, determining whether differences in administration correlate with significant differences in test outcomes.

An analysis of sample demographics found that the sample for this study had large proportions of vulnerable populations, including English learners (ELs), students classified as receiving special education services, and students living in poverty. These proportions were larger than proportions for the district as a whole and larger than proportions for the state as a whole. These large proportions point to a trend that vulnerable populations are increasing in the Shepherd School District. The large proportions also raise questions about what academic outcomes may arise as the result of these increasing proportions of vulnerable populations. Future research could explore this topic of increasing proportions of vulnerable populations and the associated outcomes.

Concluding Remarks

As noted in Chapter 1, a motivation for studying the predictive validity of the OKA was that it is presently the only universal assessment administered in the primary grades in Oregon. While this study had limitations, it provided evidence that, for students in the Shepherd Public Schools, the OKA early literacy and early math measures were only moderate predictors of third-grade SBAC achievement, indicating that it should not serve as a significant tool to predict students' future outcomes on SBAC assessments.

Educators often worry that valuable instructional time is being devoted to assessments that have little value. For educators to support the administration of the OKA it is important for ODE to assure educators of the assessment's value. Knowing that the OKA has only moderate predictive validity, many educators, myself included, may facetiously ask if the OKA is a better predictor of SBAC outcomes than just flipping a coin. The OKA may have other merits for the educators and school districts, such as assessing the outcomes of districts' preschool programs to improve preschool outcomes. If used in such a way, the potential merits of the OKA in those contexts should be further studied and communicated, rather than perpetuating the belief that OKA scores can strongly predict SBAC scores.

The workgroup that created the OKA noted that the assessments would "evolve and improve over the coming years, and . . . should be regularly reviewed to ensure long-term success" (McClelland and Squires, 2012, p. 4). If further research confirms this study's findings that the OKA has only moderate predictive validity, it is my opinion that the assessment should be revisited for a possible restructuring or replacement. In their report to the Early Learning Council, McClelland and Squires (2012) noted, "Nationally, there is work underway to develop and test new Kindergarten Readiness Assessment instruments. The instruments that are currently available are likely to be superseded by superior instruments and technologies for administration over the next five years" (p. 4). It is my opinion that the instruments that have emerged in the intervening years should be explored. There will never be a perfect predictor of human outcomes, but it is my hope that a strong predictor is achievable.

As an early childhood educator, I have had the opportunity to administer the OKA. This research leads me to personally question the appropriateness of administering a standardized assessment upon kindergarten entry. When administering the assessment, I have noted that student performance was greatly affected by many factors beyond academic skill. Trust in the administrator was a factor; naturally-reserved students did not perform for a teacher who they were meeting for the first time. The student's mood at the time of administration was also a factor; grumpy kindergarteners were not interested in the assessment. Physical needs at the time of administration such as hunger or tiredness were a factor. I noted differences between students who were testing in the morning compared to students testing in the late afternoon. These experiences, along with this research, have made me question the reliability of the OKA and question the benefit of its administration.

The fact that the only universal assessment in the primary grades appears to have only modest predictive validity on SBAC results for the Shepherd Public Schools raises a new question: What sources of available data could have stronger predictive validity? I ask this question with an important caveat: I am not supportive of more testing. As an early childhood educator, I believe students are tested more than enough and adding another assessment to the primary years is likely not the answer to identifying students not on the trajectory to read proficiently by the end of third grade.

From my perspective as an early childhood educator, the only modestly-predictive power of the OKA points to two important dispositions of a classroom teacher – to be vigilant and be optimistic. Be vigilant of the student with high kindergarten entry scores; their path to achievement is not guaranteed. Without vigilance and continued support they may have disappointing outcomes on later assessments. Be optimistic, believing that with support, the student with low kindergarten entry scores may achieve highly in later assessments.

References

Alonzo J., Tindal G., Ulmer K., & Glasgow A. (2006). easyCBM® online progress monitoring assessment system. Eugene, OR: Center for Educational Assessment Accountability.

Barry, M., & Reschly, A. L. (2012). Longitudinal predictors of high school completion. *School Psychology Quarterly*, *27*(2), 74–84. http://doi.org/10.1037/a0029189

Bechtoldt, F. (2017). Oregon's 2016 Net Migration Was the Highest Since 1993. State of Oregon Employment Department. https://www.qualityinfo.org/-/oregon-s-2016-net-migration-was-the-highest-since-1993

Berkeley, S., Bender, W. N., Gregg Peaster, L., & Saunders, L. (2009). Implementation of response to intervention: A snapshot of progress. *Journal of Learning Disabilities*, *42*(1), 85–96.

Bronson, M. B., Goodson, B. D., Layzer, J. I., & Love, J. M. (1990). Child behavior rating scale. Cambridge, MA: ABT Associates.

Center for Public Education. (2015). Learning to read, reading to learn: Why third-grade is a pivotal year for mastering literacy. http://www.centerforpubliceducation.org/Main-Menu/Policies/Learning-to-Read-Reading-to-Learn-At-a-Glance/Learning-to-Read-Reading-to-Learn-Full-Report-PDF.pdf

Chall, J.S. (1983). *Stages of reading development.* New York: McGraw-Hill.

The Children's Hospital of Philadelphia PolicyLab. (n.d.) Promoting positive educational outcomes for vulnerable populations. http://policylab.chop.edu/project/promoting-positive-educational-outcomes-vulnerable-populations

Claessens, A., Duncan, G., & Engel, M. (2009). Kindergarten skills and fifth-grade achievement: Evidence from the ECLS-K. *Economics of Education Review*, *28*, 415–427. http://doi.org/10.1016/j.econedurev.2008.09.003

Claessens, A., & Engel, M. (2013). How important is where you start? Early mathematics knowledge and later school success. *Teachers College Record*, *115*(6), 1–29.

Coleman, J. (1966). *Equality of educational opportunity*. Chicago: University of Chicago Press.

Cooper, D. H., & Farran, D. C. (1991). The Cooper-Farran behavioral rating scales. Brandon, VT: Clinical Psychology Publishing Co., Inc.

Donlon, T. E, & Angoff, W. R (1971). The scholastic aptitude test. In W. H. Angoff (Ed.) *The College Board admissions testing program* (pp. 15-47). New York: College Entrance Examination Board.

Duncan, G. J., Claessens, A., Magnuson, K., Klebanov, P., Pagani, L. S., Feinstein, L., Engel, M., Broos-Gunn, J., Sexton, H., Duckworth, K., & Japel, C. (2007). School readiness and later achievement. *Developmental Psychology, 43*(6), 1428–1446. http://doi.org/10.1037/0012-1649.43.6.1428

Forgasz, H. J. (2010). *International Perspectives on Gender and Mathematics Education.* Charlotte, NC: Information Age Publishing.

Goldstein, J., McCoach, D. B., & Yu, H (2017). The predictive validity of kindergarten readiness judgments: Lessons from one state. *The Journal of Educational Research*, *110*(1), 50–60. http://doi.org/10.1080/00220671.2015.1039111

Good, R. H., & Kaminski, R. A. (Eds.). (2002). Dynamic indicators of basic early literacy skills (6th ed.). Eugene, OR: Institute for the Development of Educational Achievement.

Herbers, J., Cutuli, J., Supkoff, L., Heistad, D., Chan, C., Hinz, E., & Masten, A. (2012). Early reading skills and academic achievement trajectories of students facing poverty, homelessness, and high residential mobility. *Educational Researcher, 41*(9), 366–374. http://www.jstor.org.georgefox.idm.oclc.org/stable/23360361

Hernandez, D. (2012). *Double jeopardy: How third-grade reading skills and poverty influence high school graduation*. Baltimore, MD: The Annie E. Casey Foundation.

House Bill 4165. (2012). https://olis.leg.state.or.us/liz/2012R1/Downloads/MeasureDocument/HB4165/Enrolled

Jordan, N. C., Glutting, J., & Watkins, M. W. (2010). Validating a number sense screening tool for use in kindergarten and first grade: Prediction of mathematics proficiency in third grade. *School Psychology Review, 39*(2), 181–195.

Judge, S. (2017). Longitudinal predictors of reading achievement among at-risk children. *Journal of Children in Poverty*, *19*(1), 0–19. http://doi.org/10.1080/10796126.2013.765629

Kang, H. (2013). The prevention and handling of the missing data. *Korean Journal of Anesthesiology*, *64*(5), 402–406. http://doi.org/10.4097/kjae.2013.64.5.402

Kjeldsen, A., Kärnä, A., Niemi, P., & Olofsson, Å. (2014). Gains from training in phonological awareness in kindergarten predict reading comprehension in grade 9. *Scientific Studies of Reading*, *18*(6), 452–467. http://doi.org/10.1080/10888438.2014.940080

LAERD (n.d.) Hierarchical multiple regression SPSS statistics. https://statistics.laerd.com/premium/spss/hmr/hierarchical-multiple-regression-in-spss.php

Lai, C. (2013). Technical Report # 1309 easyCBM ® Reading criterion related validity evidence: Grades K-1. *Behavioral Research and Teaching*. Eugene, OR: University of Oregon.

Little, R.J.A. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association, 83*(404), 1198–1202.

Lloyd, J. E. V, & Hertzman, C. (2009). From kindergarten readiness to fourth-grade assessment: Longitudinal analysis with linked population data. *Social Science & Medicine*, *68*(1), 111–123. http://doi.org/10.1016/j.socscimed.2008.09.063

Lombardo, M. (2015). How do new NAEP scores compare with Smarter Balanced and PARCC? *Education Reform Now*. https://edreformnow.org/how-do-new-naep-scores-compare-with-smarter-balanced-and-parcc/

McClelland, M. M., Acock, A. C., & Morrison, F. J. (2006). The impact of kindergarten learning-related skills on academic trajectories at the end of elementary school. *Early Childhood Research Quarterly*, *21*, 471–490. http://doi.org/10.1016/j.ecresq.2006.09.003

McClelland, M., & Squires, J. (2012). Report from kindergarten readiness assessment workgroup to the early learning council. Salem, OR: Early Learning Council. http://library.state.or.us/repository/2012/201209240948241/

Mulligan, G. M., Hastedt, S., & Carroll, J. C. (2012). First-time kindergartners in 2010–11: First findings from the kindergarten rounds of the early childhood longitudinal study, kindergarten class of 2010–11 (ECLS-K:2011) (NCES 2012-049). Washington, DC: National Center for Education Statistics.

Multnomah County Kindergarten Counts. (2017). Register by June 1: Why it matters for children. https://multco.us/sites/default/files/sun/documents/why_earlyregisterforkgmatters.pdf

National Center for Education Statistics. (2013). The nation's report card, a first look: 2013 mathematics and reading. Institute of Education Sciences. Washington, D.C.: U.S. Department of Education. https://www.nationsreportcard.gov/reading_math_2013/#/

National Center for Education Statistics. (n.d.) NAEP overview. https://nces.ed.gov/nationsreportcard/about/

No Child Left Behind Act of 2001, P.L. 107-110, 20 U.S.C. § 6319 (2002).

Office of Teaching, Learning, and Assessment. (2016). Oregon Kindergarten Assessment Specifications 2016-2017.

Office of Teaching, Learning, and Assessment. (2017). Oregon Kindergarten Assessment Specifications 2017-2018.

Oregon Department of Education. (2016a). State SBAC results. http://www.oregon.gov/ode/educator-resources/assessment/TestResults2016/state_media_1516.xlsx

Oregon Department of Education. (2016b). Statewide report card 2015–2016: An annual report to the legislature on Oregon public schools. http://www.oregon.gov/ode/schools-and-districts/reportcards/Documents/rptcard2016.pdf

Oregon Department of Education. (2017). Oregon's consolidated state plan under the Every Student Succeeds Act. http://www.oregon.gov/ode/rules-and-policies/Documents/OR_consolidatedStateplan.pdf

Oregon Deparment of Education. (n.d.) Statewide Alternate (Extended) Assessment. http://www.oregon.gov/ode/educator-resources/assessment/AltAssessment/Pages/default.aspx

Organisation for Economic Co-operation and Development. (n.d.) What is PISA? http://www.oecd.org/pisa/aboutpisa/

Organisation for Economic Co-operation and Development. (2016), *PISA 2015 Results (Volume I): Excellence and Equity in Education* (Summary), OECD Publishing, Paris. http://dx.doi.org/10.1787/bc6256e2-en

Reynolds, R., Taylor, M., Steffensen, M., Shirey, L., & Anderson, R. (1982). Cultural schemata and reading comprehension. *Reading Research Quarterly, 17*(3), 353–366.

Rimm-Kaufman, S. E., Pianta, R. C., Cox, M. J., Carolina, N., & Hill, C. (2000). Teachers' judgments of problems in the transition to kindergarten. *Early Childhood Education Journal*, *15*(2), 147–166.

Salvador, S. K., Schoeneberger, J., & Tingle, L. (2017). Relationship between second grade oral reading fluency and third grade reading. *Assessment in Education: Principles, Policy & Practice*, *19*(3), 341–356. http://doi.org/10.1080/0969594X.2011.613368

Sireci, S. G. (2012). Smarter Balanced Assessment Consortium: comprehensive research agenda. https://portal.smarterbalanced.org/library/en/comprehensive-research-agenda.pdf

Sireci, S. G., Thissen, D., & Wainer, H. (1991) On the reliability of testlet-based tests. *Journal of Educational Measurement, 28*: 237–247.

Sirin, S. R. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of Educational Research, 75*(3), 417–453.

Smarter Balanced Assessment Consortium. (2016). Smarter Balanced Assessment Consortium: 2014–2015 technical report. https://portal.smarterbalanced.org/library/en/2014-15-technical-report.pdf

Smarter Balanced Assessment Consortium (n.d.a) Estimated testing times for Smarter Balanced summative assessments. https://portal.smarterbalanced.org/library/en/estimated-testing-times.pdf

Smarter Balanced Assessment Consortium (n.d.b) History. http://www.smarterbalanced.org/about/history/

Smarter Balanced Assessment Consortium (n.d.c) Q&A on initial achievement levels. https://portal.smarterbalanced.org/library/en/achievement-level-setting-qa.pdf

Sparks, S.D. (2011, April 8). Study: Third grade reading predicts later high school graduation. [Blog post.] *Education Week: Inside School Research.* http://blogs.edweek.org/edweek/inside-school-research/2011/04/the_disquieting_side_effect_of.html

Statistics Solutions. (2013). Homoscedasticity. http://www.statisticssolutions.com/homoscedasticity/

Stockburger, D. W. (2016) Multiple regression with many predictor variables. *Multivariate statistics: Concepts, models, and applications.* http://www.psychstat.missouristate.edu/multibook/mlt07m.html

Thurlow, M. L., Wu, Y., Lazarus, S. S., Ysseldyke, J. E., (2016). Special education – non-special education achievement gap in math: Effects of reporting methods, analytical techniques, and reclassification. *Exceptionality*, *24*(1), 32–44. http://doi.org/10.1080/09362835.2014.986614

Torgesen J. K., Wagner R. K., Rashotte C. A. (2012). Test of word reading efficiency–Second Edition. Austin, TX: Pro-Ed.

Wagner, Richard K. (1999). CTOPP: Comprehensive test of phonological processing. Austin, TX: Pro-Ed.

Wanzek, J., Petscher, Y., Otaiba, S. A., Rivas, B. K., Jones, F. G., Kent, S. C., Schatschneider, C., & Mehta, P. (2017). Effects of a year-long supplemental reading intervention for students with reading difficulties in fourth grade. *Journal of Educational Psychology*. (Advance online publication). http://dx.doi.org/10.1037/edu0000184

Appendix A

Research Proposal Form

Name of Researcher: **Kelsey Weber**
Agency and Position of Researcher: **George Fox University Doctoral Candidate**
Email: **kweber14@georgefox.edu**

Title of the study: The Predictive Validity of the Oregon Kindergarten Assessment on Smarter Balanced Assessment Mathematics and English Language Arts Scores

Scope and Significance of the Study:

Research questions:

1. To what extent do Oregon Kindergarten Assessment (OKA) early literacy scores predict third-grade Smarter Balanced Assessment (SBAC) English language arts/literacy (ELA) achievement?
2. To what extent do OKA early math scores predict third-grade SBAC math achievement?
3. To what extent do OKA early literacy scores interact with students' ethnicity, socioeconomic status (SES), English learner (EL) status, classification for special education services, and gender to predict their SBAC ELA achievement scores?
4. To what extent do OKA early math scores interact with students' ethnicity, SES, EL status, classification for special education services, and gender to predict their SBAC math achievement scores?

The practical significance of this study is its functionality for informing early literacy intervention decisions for the Shepherd School District (pseudonym). It is the aim of the district leadership to identify students who are not on a trajectory to pass the third-grade SBAC with the intention of supporting those students with academic intervention. This study could offer new understandings of how to identify students in need of early intervention upon entering kindergarten. It is also in the interest of the school district to ensure that students who enter school on a trajectory to pass the third-grade SBAC continue on that trajectory; this study could be a useful tool in monitoring student progress. This research on the predictive validity of the OKA on third-grade SBAC results can benefit Oregon teachers, principals, and school districts seeking a predictive tool for third-grade assessment outcomes.

The Shepherd Public Schools is a notable population to study as the school district has a high proportion of vulnerable populations of students. Students in poverty, ELs, and students receiving special education services are vulnerable populations. According to the district website, Shepherd Public Schools is one of the largest school districts in the state with over 40,00 students; 18% of students are ELs, well above the state EL proportion of 10%; 15% of Shepherd students receive special education services, above the state proportion of 13%, and 61% of Shepherd students are considered to live in poverty, well above the state proportion of 51%. Because Shepherd has high proportions of these vulnerable populations, it is an important population to study. This study could offer new understandings of what variables to use to identify students from these vulnerable populations in need of early intervention. In addition, this study can give impetus for further study and replication in other districts.

The scholarly significance of this study is its addition to the volume of research on early childhood and elementary level assessments. Both OKA and SBAC are new assessment tools; much research remains to be done on these assessments.

Research Design:

This study will be an *ex post facto* repeated measure, predictive validity study looking at Oregon Kinder Assessment (OKA) scores as a predictor of Smarter Balanced Assessment (SBAC) scores in the Shepherd Public Schools.

The pseudonym of Shepherd Public Schools will be used in place of the district's name.

I will analyze data using Statistical Packages for Social Sciences (SPSS). Analyses will include multiple regression and hierarchical regression.

I will store data on an encrypted external drive at my personal residence. I will analyze the data using a password protected personal laptop computer. All data will be stored on the encrypted external drive, which I will keep in a locked cabinet when not in use. I will keep these files on the encrypted external drive in a locked cabinet for five years from completion of the study; retaining data for five years after the completion of research is a policy of George Fox University. At that point in time, the files will be deleted from the encrypted external drive.

Logistics of the Study:

A. **Selection and Definitions of the Study Population**

Population: Students who were third graders in 2016–2017 school year

B. **Data to be Collected**

For students* enrolled in third grade in Shepherd in the 2016–2017 school year –
-Ethnicity
-Free and reduced lunch status
-English Learner status
-Classification for special education services
-Gender
-Oregon Kindergarten Assessment early literacy score
-Oregon Kindergarten Assessment early math score
-SBAC ELA score (Score, NOT achievement level)
-SBAC math score (Score, NOT achievement level)
-Grade repetition/retention
*Data de-identified of names and student identification numbers

C. **Source of the Data**

Student information systems

D. **Quality of the Data**

E. **Survey Distribution Methodology** (if applicable)

None

F. **Proposed Start and End Dates**

Timeline for Research

Task	**Time Period**
Obtain data from Shepherd School district	Feb 2018
Data analysis using Statistical Packages for Social Sciences (SPSS)	Feb 2018
Write results	Feb 2018
Write discussion and results in chapter 5	March 2018
Send draft to dissertation chair for comments	March 2018
Send draft to dissertation committee	March 2018
Defend dissertation	April 2018
Destroy data	April 2023

<u>Informed Consent:</u>
Include any pertinent forms, and indicate how they will be distributed, copied and filed.

N/A

Appendix B

Assumptions Tables and Figures

Figure 1
Missing Values Analysis Pie Charts

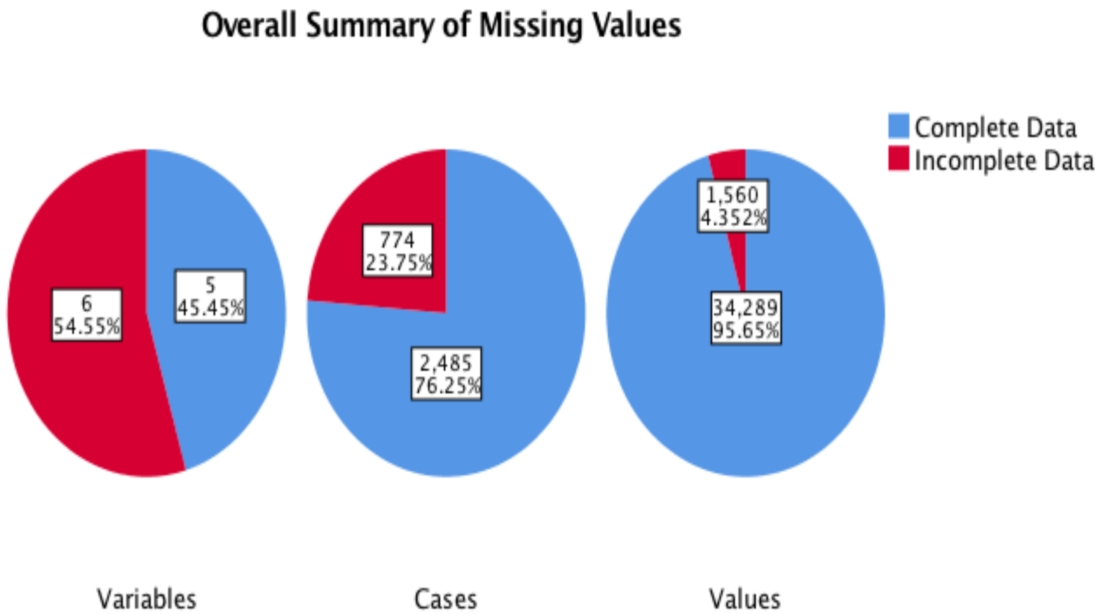


Figure 2
Missing Value Patterns

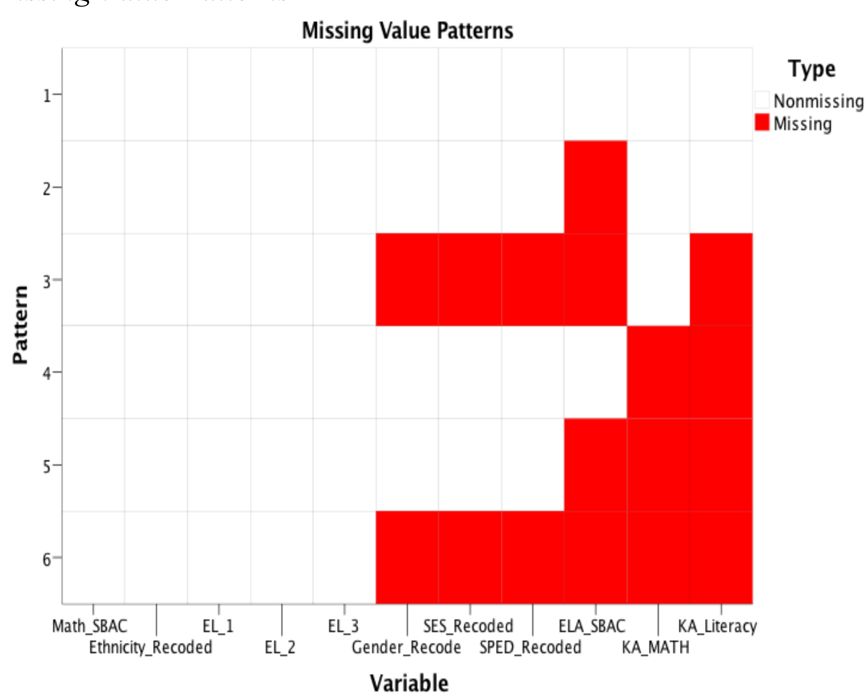


Figure 3

Partial Regression Plot of OKA Early Literacy Scores and SBAC ELA Scores

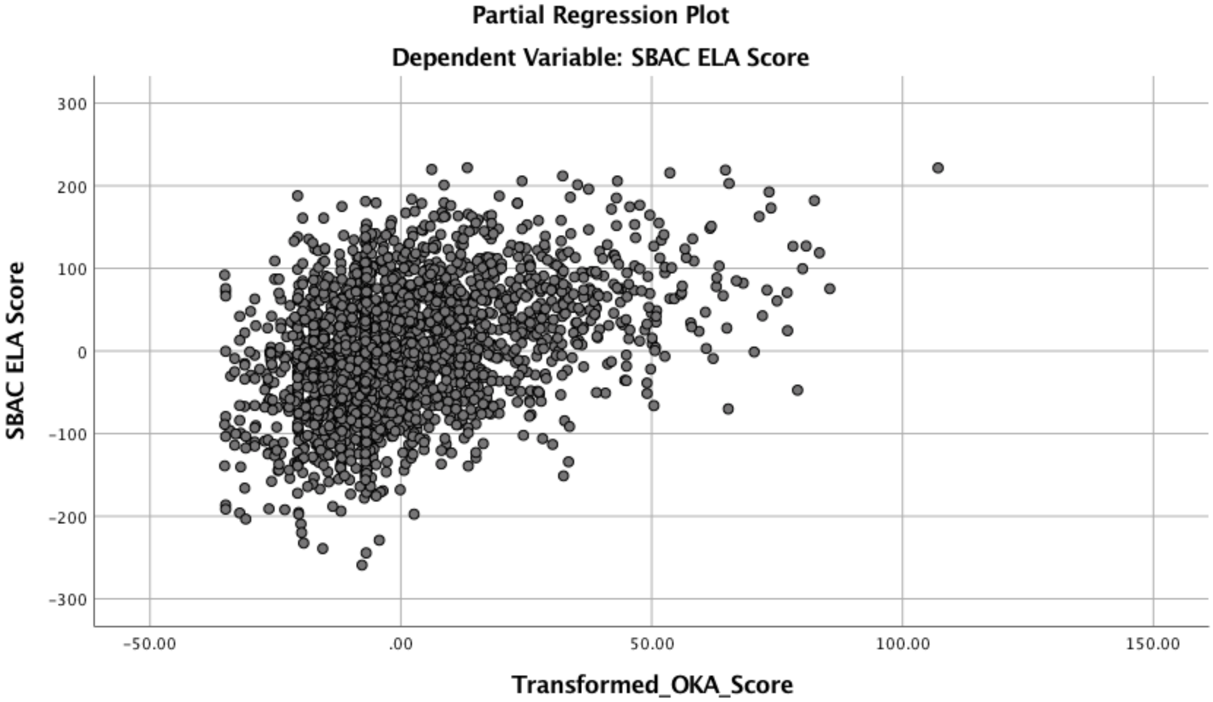


Figure 4

Partial Regression Plot of OKA Early Math Scores and SBAC Math Scores

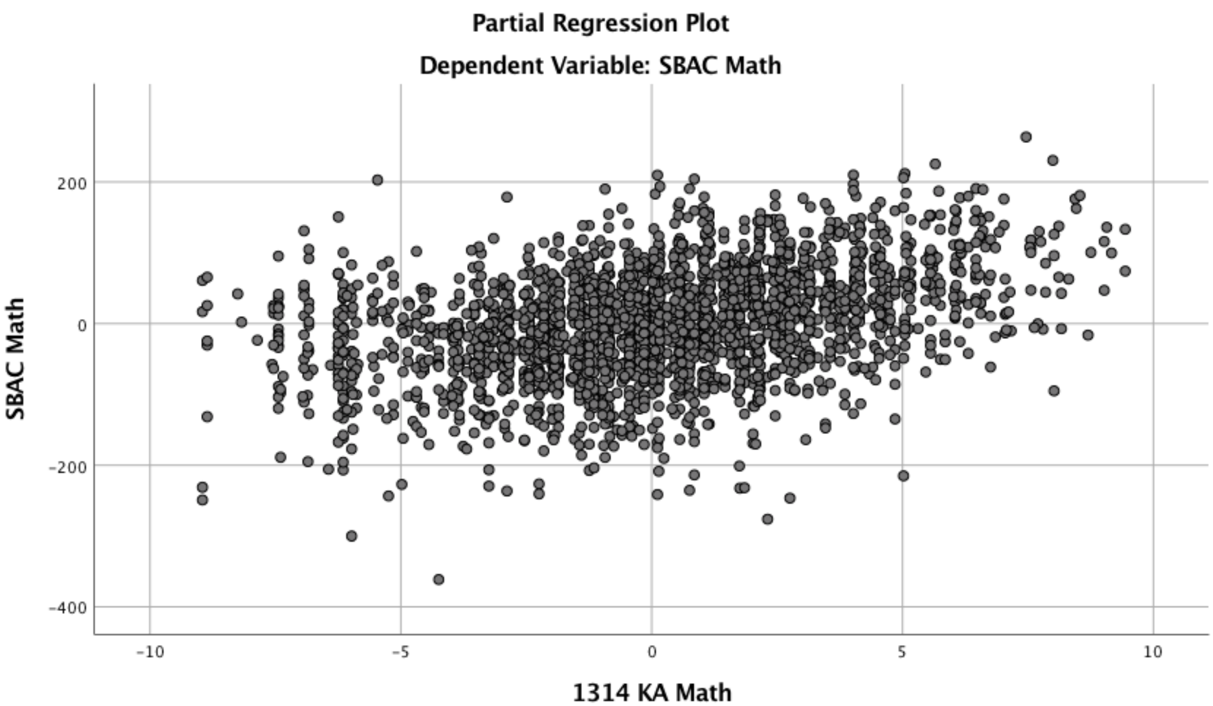


Figure 5
Plot of Studentized Residuals Against Unstandardized Predicted Values for SBAC ELA

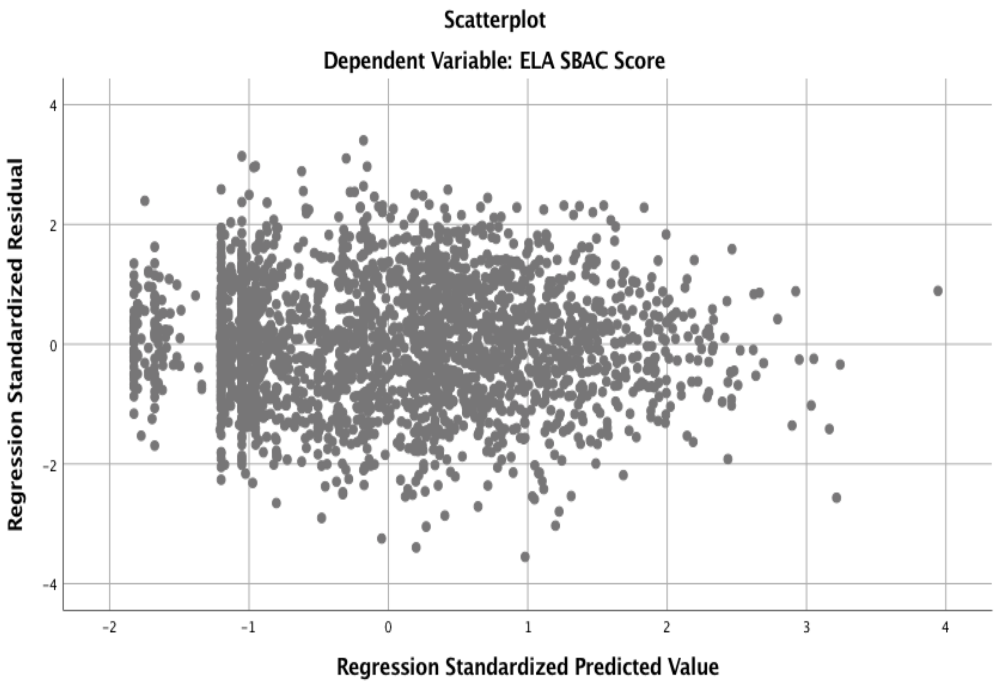


Figure 6
Plot of Studentized Residuals Against Unstandardized Predicted Values for SBAC Math

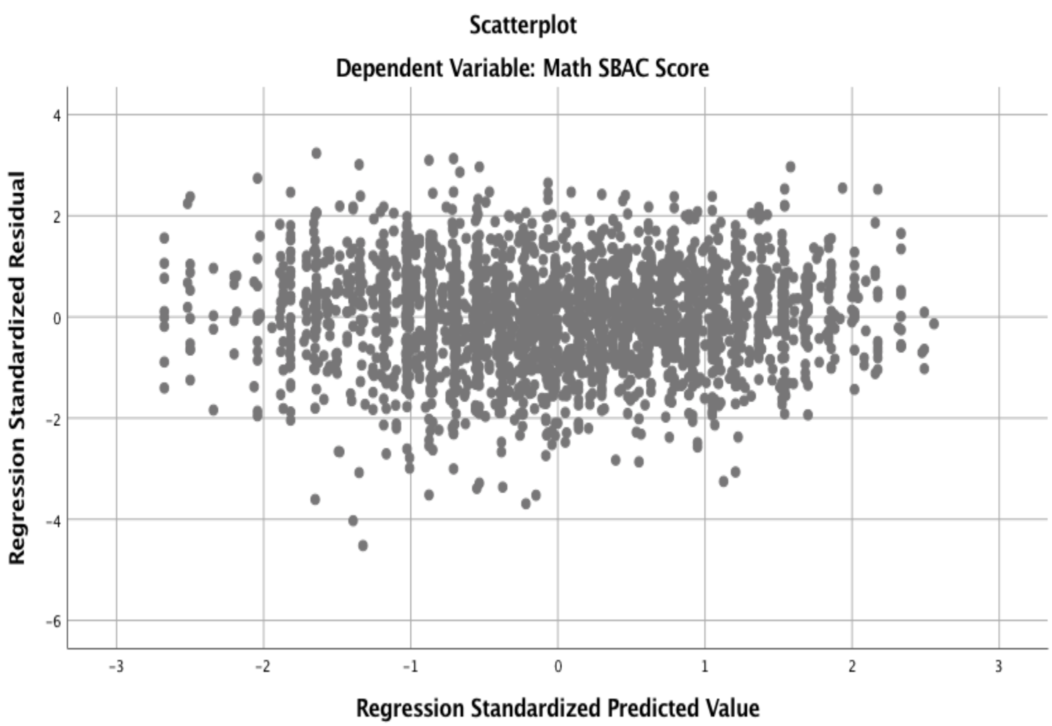


Table 18
Collinearity Statistics Summary for SBAC ELA

	Standardized β	*t*	*p*	Partial Correlation	Collinearity Statistics Tolerance	Collinearity Statistics VIF
Gender	.092	3.321	.006	.067	.981	1.019
Ethnicity	.047	-3.074	.000	-.062	.588	1.701
SES	-.043	-9.477	.000	-.187	.750	1.334
SPED	-.303	-10.759	.000	-.221	.912	1.096
EL Status	.079	-10.388	.000	.172	.546	1.833

Table 19
Collinearity Statistics Summary for SBAC Math

	Standardized β	*t*	*p*	Partial Correlation	Collinearity Statistics Tolerance	Collinearity Statistics VIF
Gender	-.054	-3.361	.001	-.067	.980	1.020
Ethnicity	-.073	-3.540	.000	-.071	.589	1.698
SES	-.192	-10.729	.000	-.211	.784	1.276
SPED	-.210	-12.733	.000	-.248	.927	1.078
EL Status	-.168	-7.976	.000	-.158	.570	1.756

Figure 7
Residuals of SBAC ELA

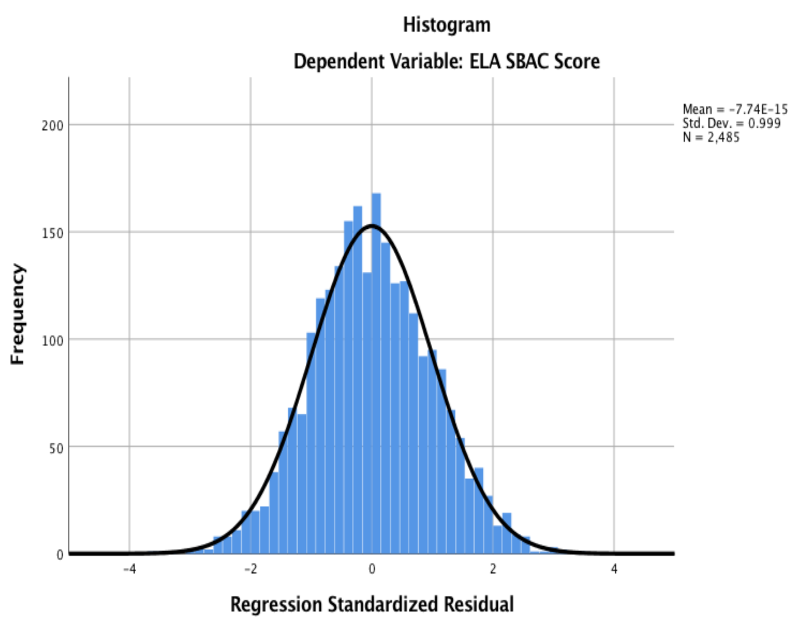
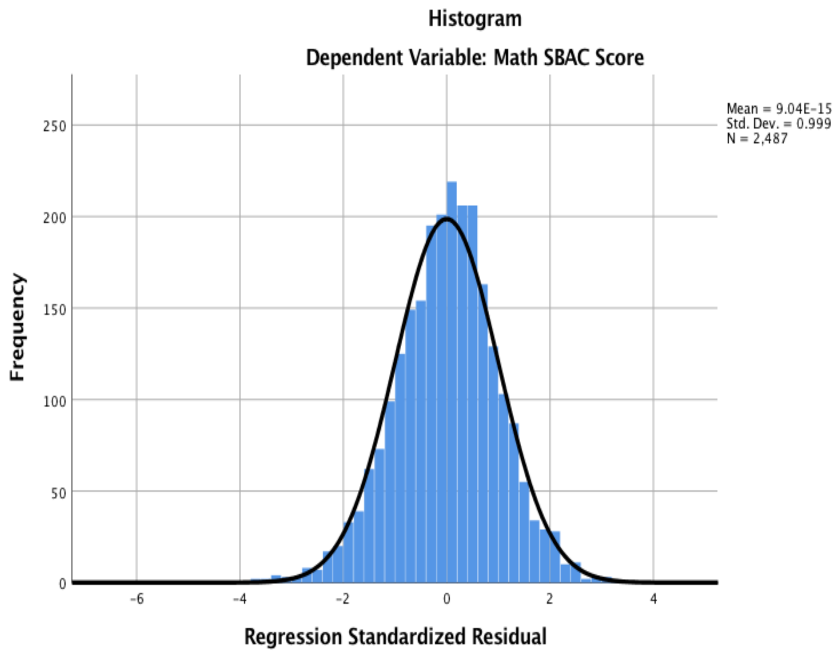


Figure 8
Residuals of SBAC Math



Appendix C:

Descriptive Statistics for Response Variables by Predictors

Table 20
Descriptive Statistics for Response Variables by Descriptors

		SBAC ELA	SBAC Math	OKA Literacy	OKA Math
Total	*N*	3252	3259	2487	2490
	Mdn	2390.00	2411.00	8.00	7.00
	Mean	2395.52	2409.22	17.25	7.07
	SD	89.30	86.68	21.34	3.44
	SEM	1.56	1.51	0.428	0.069
Gender					
Male	*N*	1190	1190	1190	1190
	Mdn	2376.00	2412.00	7.00	7.00
	Mean	2386.05	2411.04	16.25	7.09
	SD	88.659	89.656	22.026	3.542
	SEM	2.570	2.599	0.638	0.103
Female	*N*	1295	1295	1295	1295
	Mdn	2397.00	2411.00	9.00	7.00
	Mean	2401.55	2407.46	17.57	6.85
	SD	87/084	82.816	20.718	3.357
	SEM	2.420	2.301	0.576	0.093
Ethnicity					
White	*N*	1130	1130	1130	1130
	Mdn	2426.00	2439.00	20.00	8.00
	M	2422.83	2432.76	25.04	7.81
	SD	90.004	87.771	23.169	3.477
	SEM	2.677	2.611	0.689	0.103
Ethnic minority	*N*	1355	1355	1355	1355
	Mdn	2362.00	2389.51	2.00	6.00
	M	2370.20	2392.00	10.77	6.45
	SD	79.026	79.652	17.210	3.297
	SEM	2.147	2.164	0.468	0.090
SES					
Not FRL	*N*	559	559	559	559
	Mdn	2467.00	2471.00	29.00	9.00

	M	2457.76	2465.38	33.22	8.67
	SD	81.458	77.030	24.837	3.319
	SEM	3.445	3.258	1.051	0.140
FRL	*N*	1926	1926	1926	1926
	Mdn	2369.00	2395.00	4.00	7.00
	M	2375.66	2392.86	12.63	6.60
	SD	81.179	81.705	17.722	3.342
	SEM	1.850	1.862	0.404	0.076
Classification for SpEd					
Not SpEd	*N*	1993	1993	1993	1993
	Mdn	2402.00	2421.00	10.00	7.00
	M	2404.74	2420.09	18.70	7.26
	SD	86.255	81.595	22.078	3.417
	SEM	1.932	1.828	0.495	0.077
SpEd	*N*	492	492	492	492
	Mdn	2339.00	2363.50	4.00	6.00
	M	2351.15	2364.96	11.43	5.95
	SD	82.661	90.019	16.929	3.450
	SEM	3.727	4.058	0.763	0.156
EL Status					
Yes	*N*	690	690	690	690
	Mdn	2323.00	2362.00	0.00	6.00
	M	2328.02	2357.17	3.393	5.50
	SD	56.810	74.317	8.134	3.032
	SEM	2.163	2.829	0.310	0.115
No	*N*	1490	1490	1490	1490
	Mdn	2420.00	2431.00	19.00	8.00
	M	2417.31	2426.75	24.46	7.74
	SD	89.094	85.521	23.085	3.404
	SEM	2.308	2.216	0.598	0.088
Transitioned	*N*	305	305	305	305
	Mdn	2432.00	2442.00	4.00	7.00
	M	2430.41	2440.99	12.18	7.30
	SD	58.711	60.420	16.770	3.405
	SEM	3.362	3.460	0.957	0.195