

8-26-2018

The Relationship Between Core Knowledge Sequence Curriculum and Smarter Balance Assessment English Language Arts Scores

Melissa Meyer
mmeyer06@georgefox.edu

This research is a product of the Doctor of Education (EdD) program at George Fox University. [Find out more](#) about the program.

Recommended Citation

Meyer, Melissa, "The Relationship Between Core Knowledge Sequence Curriculum and Smarter Balance Assessment English Language Arts Scores" (2018). *Doctor of Education (EdD)*. 119.
<https://digitalcommons.georgefox.edu/edd/119>

This Dissertation is brought to you for free and open access by the Theses and Dissertations at Digital Commons @ George Fox University. It has been accepted for inclusion in Doctor of Education (EdD) by an authorized administrator of Digital Commons @ George Fox University. For more information, please contact arolf@georgefox.edu.

THE RELATIONSHIP BETWEEN CORE KNOWLEDGE SEQUENCE CURRICULUM AND
SMARTER BALANCE ASSESSMENT ENGLISH LANGUAGE ARTS SCORES

by
Melissa Meyer

FACULTY RESEARCH COMMITTEE:

Chair: Dane C. Joseph, Ph.D.

Member: Ginny Birky, Ph.D.

Member: Karen Buchanan, Ed.D.

Presented to the Doctoral Department
and College of Education, George Fox University
In partial fulfillment of the requirements for the degree of
Doctor of Education

August 26, 2018

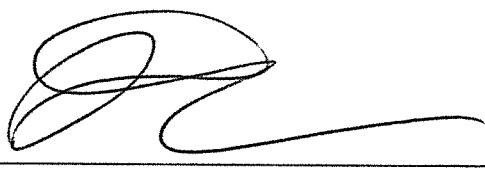


GEORGE FOX
UNIVERSITY

COLLEGE OF EDUCATION

"THE RELATIONSHIP BETWEEN CORE KNOWLEDGE SEQUENCE CURRICULUM AND SMARTER BALANCE ASSESSMENT ENGLISH LANGUAGE ARTS SCORES," a Doctoral research project prepared by MELISSA MEYER in partial fulfillment of the requirements for the Doctor of Education degree in Educational Leadership.

This dissertation has been approved and accepted by:

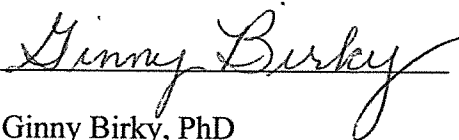
9/6/18 

Date

Dane Joseph, PhD

Committee Chair

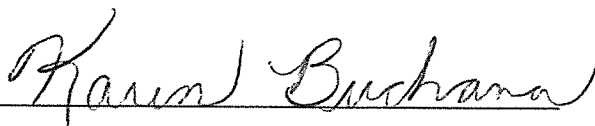
Assistant Professor of Education

9/6/18 

Date

Ginny Birky, PhD

Professor of Education

9/6/18 

Date

Karen Buchanan, EdD

Professor of Education

ABSTRACT

The purpose of this study was to examine student achievement as measured by the English language arts (ELA) Smarter Balance Assessment (SBAC) between students who were taught the Core Knowledge Sequence curriculum and students taught the Houghton Mifflin Harcourt's Journeys curriculum. The 11,493 participants were third through sixth grade students in the 2014-2015, 2015-2016, and 2016-2017 school years. The study used a multiple regression model to examine the extent to which the students' gender, ethnicity, socioeconomic status (SES), disability status, English language learner (ELL) status, and curriculum predicted SBAC ELA scores. The findings of this study suggest that curriculum did not predict SBAC ELA scores. The multiple regression model indicated that demographic variables offer more predictive information on SBAC ELA scores.

ACKNOWLEDGEMENTS

Thank you to my dissertation committee for your unwavering support. I could not have completed this effort without your grace and understanding in the face of my seemingly all-encompassing professional work. Dane, thank you for asking me if the cost of my professional work is worth it – your personal interest in why my work is so important to me makes a difference every time I think of the conversation. Karen, thank you for your interest in wanting to know how the Core Knowledge Sequence fits with the Common Core State Standards. This simple request opened up a broad avenue which I had not yet explored. Ginny, thank you for many years of ongoing encouragement and editing directions. Hopefully, your high standards have made this a stronger dissertation (if not, it is entirely my responsibility!)

Thank you to my husband, Matt, for always believing in me. Thank you for the countless early breakfasts, late dinners, and cups of coffee with sugar on the rim. Thank you for reading and editing my writing with thoughtfulness – your humor keeps me laughing even when struggling with this work. You are the love of my life without whom I would be much less than I am.

First and last, thank you God. Let any wisdom found in this work build houses and the knowledge fill the rooms with riches. May all the glory be Yours.

“Through wisdom a house is built,

And by understanding it is established.

By knowledge the rooms are filled

With all precious and pleasant riches.” Proverbs 24:3-4

TABLE OF CONTENTS

ABSTRACT.....	i
ACKNOWLEDGEMENTS.....	ii
LIST OF TABLES	vi
CHAPTER 1	1
Introduction.....	1
Rationale for the Study.....	2
Statement of the Problem	3
Purpose of the Study	5
Research Question.....	5
Significance of the Study	5
Key Terms	6
Limitations and Delimitations.....	8
CHAPTER 2	9
Literature Review.....	9
Selecting Curriculum.....	9
The Core Knowledge Sequence	11
Houghton Mifflin Harcourt’s Journeys	17
Standardized Assessment as an Accountability Measure	27
CHAPTER 3	38
Methodology	38
Design.....	38

Sampling & Participants.....	39
Smarter Balance Assessment Instrument	39
Data Collection.....	41
Timeline	42
Data Analysis	42
Independent variables.	43
Dependent variable.	45
Research Ethics	45
CHAPTER 4	46
Results.....	46
Sample Demographics.....	46
Descriptive Statistics	48
Assumptions	50
Independence of Errors.....	50
Correlations between Dependent Variables and Independent Variables.....	51
No multicollinearity.....	56
No significant outliers.	58
Research Question.....	59
2015 Data.....	59
2016 Data.....	61
2017 Data.....	63
Summary	65

CHAPTER 5	66
Conclusion	66
Discussion	67
Limitations	71
Implications	71
Recommendations for Further Research	74
Conclusion.....	75
References.....	77
APPENDICES	87
Appendix A: Common Core State Standards.....	88
Appendix B: ODE Memorandum of Understanding	89

LIST OF TABLES

Table 1: Meta-analysis ranking CSR models, Boreman et al. (2015).....	14
Table 2: State-by State Journeys Implementation 2010-11	20
Table 3: Depth of Knowledge Taxonomy and ODE SBAC Achievement Levels	41
Table 4: Variable Operationalization.....	44
Table 5: Sample Demographics	47
Table 6: Achievement Standards Summary	48
Table 7: Descriptive Statistics for Smarter Balance Assessment English Language Arts Scaled Scores...	50
Table 8: Summary of Durbin-Watson Statistics	51
Table 9: Dependent and Independent Variable Correlations for 2015.....	53
Table 10: Dependent and Independent Variable Correlations for 2016.....	54
Table 11: Dependent and Independent Variable Correlations for 2017.....	56
Table 12: Collinearity Statistics Summary for 2015.....	57
Table 13: Collinearity Statistics Summary for 2016.....	57
Table 14: Collinearity Statistics Summary for 2017.....	58
Table 15: Case Outliers.....	59
Table 16: Multiple Regression Analyses, 2015	60
Table 17: Unstandardized Coefficients for Independent Variables with the Dependent Variable 2015 SBAC ELA Score, With and Without the Independent Disability Status Variable.....	61
Table 18: Multiple Regression Analyses, 2016	62
Table 19: Unstandardized Coefficients for Independent Variables with the Dependent Variable 2016 SBAC ELA Score, with and without the Independent Disability Status Variable	63
Table 20: Multiple Regression Analyses, 2017	64

Table 21: Unstandardized Coefficients for Independent Variables with the Dependent Variable 2017

SBAC ELA Score, with and without the Independent Disability Status Variable	65
---	----

CHAPTER 1

Introduction

The contemporary scope of school accountability rests primarily on student achievement as measured by standardized test scores and is the result of provisions in President Johnson's 1965 Elementary and Secondary Education Act (ESEA) legislation converging with the 1983 National Commission for Excellence report, *A Nation at Risk*. The political dialogue surrounding *A Nation at Risk* prompted states' governors, legislatures, and education policy makers throughout the 1980s to create hundreds of education commissions and pass numerous bills to transform education standards (Massell & Fuhrman, 1994). Standards-based reform efforts and a political emphasis on decentralized control of education encouraged the growth of whole-school reform models to turn around low-performing schools. The Comprehensive School Reform (CSR) program was a primary federal funding mechanism designated to scale up whole-school reform models across the country (Borman, Hewes, Overman, & Brown, 2003).

One curriculum that qualified as a whole-school reform model under the CSR program was the Core Knowledge Sequence curriculum, developed from 1988 through 1990 by an advisory board of academics, K-12 teachers, and multicultural specialists (Core Knowledge Foundation, 2015). The Core Knowledge Sequence (CKS) met the U.S. Department of Education's criteria to qualify as a comprehensive school reform design; thus, schools implementing CKS were eligible for federal funding through CSR (Borman et al., 2003). According to Borman et al, during the decade of 1993-2003, over 800 schools nation-wide implemented the CKS curriculum.

Rationale for the Study

EAST Charter School has used the CKS as the curriculum for its kindergarten through eighth grade program since 2003. In applying for a charter with the Fairmont School District (FAIRMONT) in 2002, EAST founders conjectured that students taught the CKS curriculum would score better on the state standardized tests than students taught in schools using district-selected curricula. The Fairmont School District granted a charter, which was renewed for ten years in 2008 based on the same justification of higher standardized test scores.

The CKS was not originally intended as either a whole-school reform or developed to improve student achievement on standardized tests. The inception of cultural literacy, which is the foundation to the CKS, began in the 1970s through E.D. Hirsch, Jr.'s research at the Richmond Community College and the University of Virginia examining the factors important to a student's ability to comprehend a reading passage. His finding that background knowledge was a stronger factor in a student's ability to comprehend a reading passage than the readability of the text was published in the *Journal of Basic Writing* and *The American Scholar* between 1980 and 1983 (Core Knowledge Foundation, 2010).

In *Cultural Literacy: What Every American Needs to Know* (Hirsch, 1987) Hirsch, in collaboration with Joseph Kett and James Trefil, itemized the knowledge literate Americans tend to share. Hirsch acknowledged the itemized list was incomplete because "inappropriate omissions and inclusions are bound to occur in a first attempt" (p. 146) and hoped to open a national dialogue about the basic knowledge needed to ensure literacy. Rather than opening a dialogue, the list elicited criticism "as promoting elitist forms of knowledge arbitrarily decided by a few people" (Datnow, McHugh, Stringfield, & Hacker, 1998, p. 411). In response to the criticism, Hirsch refined his focus from what every American must know to what American

school children need to know. He reviewed international comparisons of student achievement and reported that in countries where a national core curriculum was offered, students scored better than American students (Hirsch, 1996). Hirsch also established the Core Knowledge Foundation with an advisory board of multicultural experts and independent educators and scholars, who, in 1990, were convened to develop a specific list of topics for kindergarten through the sixth grade, i.e., the Core Knowledge Sequence (Core Knowledge Foundation, 2010).

A key principle supporting the CKS is the idea that “social justice requires all citizens to share an extensive body of school-based background knowledge as a necessary foundation for communication and participation in society” (Hirsch, 1996, p. 14). According to the Core Knowledge Foundation, the CKS was based on three premises:

- There is a specific body of knowledge in history, mathematics, science, literature, art, and music that forms an educational core for a pre-school through eighth grade curriculum;
- What a student can learn is predicated upon what she already knows;
- Knowledge, language, and skills build sequentially and cumulatively (Core Knowledge Foundation, 2015).

The CKS curriculum is knowledge-based, content-rich, integrated, and intended to be taught sequentially. The scope and sequence cover language arts and English, history and geography, the visual arts, music, mathematics, and science.

Statement of the Problem

The charter agreement between EAST Charter School and the Fairmont School District rested upon the expectation that students taught the CKS would score better on the state standardized tests than students taught the district-selected curricula. Unfortunately, there was

little site-specific research to substantiate the assertion surrounding the use of the CKS, and CKS efficacy on student achievement nationally “needs several more rigorous evaluations to establish a stronger research base” (Borman et al., 2003, p. 162). The lack of either site-specific data or rigorous evaluations of CKS efficacy was not problematic at the time of the original charter since Oregon received federal approval of content-based standardized assessments in 2001 (U.S. Department of Education, 2001). While the 2002 contract between EAST and Fairmont School District did not articulate the inferences and assumptions linking standardized test performance and outcomes to renewal decisions (Kane, 2016), in retrospect, the original conjecture that students taught the CKS would perform better than other students was understandable. CKS is a content-based curriculum, and the standardized tests were specifically designed to assess how well children were “learning the material according to the State content standards” (U.S. Department of Education, 2001, p.1).

But this dearth of data became problematic. In the sixteen-year interim from the founding of EAST Charter School to a third charter renewal with the Fairmont School District in April 2018, Oregon adopted the Common Core State Standards (CCSS) and implemented the use of the Smarter Balance Assessment (SBAC) to assess students’ progress towards achieving the CCSS skill benchmarks. Student achievement was no longer measured on knowledge acquisition, but on skill acquisition, and the annual state report cards for EAST Charter School reflected a declining shift in the percentage of students who met English language arts benchmarks beginning at the time of the implementation of the SBAC assessment in 2015. The charter school was no longer meeting the contractual agreement that students taught CKS would score better on the state standardized tests than students taught the district-selected curricula.

Purpose of the Study

The purpose of this study was to investigate student English language arts achievement on the SBAC based on curriculum taught, either the CKS or the district-selected curricula. This study was a secondary-data analysis using 2015, 2016, and 2017 data from the SBAC assessments in the Fairmont School District. The sample included 11,493 student scores distributed across three years.

Research Question

This study explored the impact of the CKS curriculum on Oregon Smarter Balanced Assessment scores in language arts:

- Is there a difference in the Oregon Smarter Balance English language arts test scores by curriculum taught?

Significance of the Study

In 1999, the Oregon Legislature passed Oregon Revised Statute (ORS) 338, created charter schools

“as a legitimate avenue for parents, educators and community members to take responsible risks to create new, innovative and more flexible ways of educating children within the public-school system. The Legislative assembly seeks to create an atmosphere in Oregon’s public-school system where research and development of new learning opportunities are actively pursued... It is the intent that public charter schools may serve as models and catalysts for the improvement of other public schools and public-school system” (ORS 338.015).

Therefore, Oregon charter schools existed under a mandate to conduct research for the benefit of the State’s educational systems; but the Charter School Program Government and Legal Affairs

Analyst at the Oregon Department of Education (ODE) stated the division is not aware of any such research (Pattison, 2015). This study will provide a model for how to begin building the body of research for Oregon charter schools, as well as introduce a piece of literature to the fledgling body of research evaluating Oregon charter schools' efficacy in improving student achievement.

Charter school sponsors offer five- to ten-year contracts based on the expected student achievement on state standardized assessments. This study will add to a body of research examining differences between curricula used by charter schools and the state approved curricula used by districts. In turn, a robust body of research could provide support for clarity in decision-making and contractual relations between the charter school and sponsoring district. For example, the Fairmont School District renewed EAST Charter School's charter agreement in April 2018; a body of research examining the effectiveness of various curricula could provide contextual information about the effectiveness of Core Knowledge Sequence in comparison to other curricula when negotiating specific contractual performance expectations.

Finally, the literature examining the CKS's impact on student achievement is quite dated; this study would offer a current review of the curriculum's efficacy.

Key Terms

Charter School – Models for charter schools, an alternative way to offer public education, are distinctly different from state-to-state. A charter school in Oregon is a legally independent public school of choice created and run by parents, teachers, and community members under a contract with the local district school board. The charter school must be non-sectarian and tuition-free. Any student who wishes may attend, but school enrollment caps are

contractually binding which creates limits on access. Charter school funding, set by legislation, is 80% of the annual state school fund designation per student.

Comprehensive School Reform Program – The Comprehensive School Reform Program is the funding component of the No Child Left Behind Act granting funds to public schools and districts nation-wide to support the implementation of effective instructional practices. The legislation requires “schoolwide improvements that covers virtually all aspects of a school operations” (U.S. Dept. of Education, 2004, p. 1), and districts or schools receiving funding must partner with an external agency knowledgeable about school reform to guide the implementations. The CSR program supported reforms for low-performing schools to significantly improve student achievement.

Oregon Assessment of Knowledge and Skills (OAKS) – The Oregon Assessment of Knowledge and Skills is the statewide summative assessment system with the annual results used in policy decisions. Standardized tests in math and English language arts are administered annually for grades three through eight and eleven. Standardized tests in social sciences and science are administered annually for grades three, five, eight, and eleven. There are supplemental assessments in English language proficiency for English learners, as well as Extended Assessments for special needs students.

Smarter Balanced Assessment Consortium (SBAC) – The Smarter Balanced Assessment Consortium is a public agency hosted by UCLA’s Graduate School of Education and Information Studies and is currently governed by fifteen states and the Bureau of Indian Affairs. The SBAC is a standardized assessment system aligned to the Common Core State Standards (CCSS). The acronym is used to identify the assessment, as well. For example, the SBAC ELA is the acronym for the Smarter Balanced Assessment in English language arts.

Standards-Based Reform (SBR) – Standards-based reform, also known as standards-based education, is a national effort to set clear, measurable academic standards to effect change at the local elementary and secondary school level. The goal was to organize curricula around a specifically defined set of skills and competencies measured by assessment systems to determine “whether or not students in all districts are meeting the standards” (Gandal, 1997, p. 16). The premise for identifying specific skills and competencies that all student must possess which are assessed by standardized tests encourages students and teachers to take more responsibility for learning outcomes (Hakuta, 2000).

Limitations and Delimitations

Methodology and Time -- EAST Charter School is one of three Oregon public schools teaching the CKS, and as such it is possible to consider this work a secondary-data analysis with the Fairmont Schools District schools teaching Houghton Mifflin Harcourt’s *Journeys*. The common goal of improved student learning as measured by the SBAC ELA would suggest an analysis of the similarities, differences, and patterns between the two curricula. Unfortunately, time is a factor preventing the collection of qualitative data for the two curricula, thus the study is a secondary-data analysis examining the efficacy of the curricula on student achievement. Time is also a limiting factor. Oregon implemented the use of the SBAC ELA in 2014-2015 and there are only three years of student achievement test scores available for analysis.

Sample -- The findings of this study cannot be generalized to larger populations because a convenience sample is used for both the treatment and the control groups.

Treatment -- The fidelity of instruction for either the CKS or for Houghton Mifflin Harcourt’s *Journeys* is not accounted for in this study.

CHAPTER 2

Literature Review

This study examined the effect curriculum taught had on student achievement. The purpose of the literature review was to clarify the epistemology that drives the decision-making in selecting a curriculum, examine the relevant research concerning the Core Knowledge Sequence and Houghton Mifflin Harcourt's Journeys, and, finally, to examine the literature concerning standardized testing used as a school accountability measure.

Selecting Curriculum

Curriculum choice is driven by many factors such as private versus public school needs, fiscal resources, a developing body of research surrounding a pedagogical approach to education, and many other factors outside the scope of this study. What was important to the study was to understand the reasons EAST Charter School chose to teach the Core Knowledge Sequence (CKS), as well as the reasons the Fairmont School District selected the Houghton Mifflin Harcourt Journeys curriculum. The decision-making process in selecting a curriculum straddled state legislated mandates and educational theory.

In Oregon, the public-school instructional materials used to teach kindergarten through twelfth grade students must align with the state-adopted Common Core State Standards (CCSS) and are subject, by law, to a review process before approval by the State Board of Education for use in schools. Instructional materials are adopted by content area with specific, but limited choices available for districts' use for a six-year cycle. The Fairmont School District curriculum choice of Houghton Mifflin Harcourt's Journeys fulfills the legislated mandate. The District is required to adopt an English language arts curriculum "in which the materials must make up an organized system of instruction that align with adopted state standards" (ODE, 2013, p. 1).

Houghton Mifflin Harcourt's Journeys meets this state requirement, as well as fulfills the District's curriculum renewal goal to align "standards within our existing programs and determining if the resources we have are sufficient" (Robson, 2017, p. 2).

A charter school in Oregon, by legislative intent, is a public school run as a self-governing, separate legal entity under contract with the board of the local school district. The charter school is subject to some of the laws pertaining to traditional public schools, but free from others. In matters of curriculum, charter schools are not required to adopt teaching materials or textbooks from the State Board of Education's approved list; rather, the choice of school curricula are written into the initial application for a charter agreement. The application includes the rationale for the use of the curricula, as well as the expected learning outcomes.

The anthropological theory of education posited "all human communities are founded upon specific shared information" (Hirsch, 1987, p. xv) with acculturation, or cultural transmission from one generation to the next, as the primary purpose of education (Spindler, 1984, p. 4). Furthermore, this theoretical framework identified education as "both the deliberate inculcation of knowledge, attitudes, and values and the unconscious transmission of modes of perceiving the world" (DuBois, 1955, p. 91). In applying for a charter agreement, the EAST Charter School's Board of Directors declared a mission of acculturation by simply stating: "the Core Knowledge Sequence teaches a cultural literacy that allows students to make sense out of today's world. It provides coordinated, systematic study drawing upon such disciplines as anthropology, archeology, economics, history, geography, law, philosophy, political science, psychology, religion, and sociology, as well as appropriate content from humanities, mathematics and natural sciences in a sequential order" (Lorence, 2001, p. 5).

The Core Knowledge Sequence

It was important to the study to acknowledge the tacit assumption that any differences in student outcomes between content-based curriculum such as the Core Knowledge Sequence (CKS) and skill-based curriculum like Houghton Mifflin Harcourt's Journeys could be evaluated using the same standardized assessment tool focused on measuring skills. This was possible because a content-based curriculum implicitly included the development of domain-specific skills necessary to master the content knowledge (Feltovich, Prietula, & Ericsson, 2006).

Datnow, Borman, and Stringfield, (2000) conducted a longitudinal, mixed-methods study of the effects of the implementation of the CKS in four schools in different states over three years, 1995 to 1998. Datnow et al. (2000) prepared case studies of CKS schools with a quasi-experimental, untreated control group design (pre- and post-tests) comparing student achievement on norm-referenced standardized tests and on specifically prepared CKS subject tests. The study looked at the degree to which CKS was implemented, the conditions that supported or stalled implementation, and the effect of CKS on teachers' curricular coordination across grades and students' achievement in reading and math, as well as three CKS content subjects.

For the Datnow et al. (2000) study, the Core Knowledge Foundation identified four schools as relatively advanced in the implementation of the CKS, which the researchers then compared with four like schools that were identified with the assistance of local education agencies in Florida, Texas, Maryland, and Washington. Comparison schools were based on similarities in the percentage of students who qualified for free or reduced lunches, the school's racial and ethnic composition, and historical student achievement levels. Student achievement was matched across five measures: basic skills in math and reading, three measures in Core

Knowledge subjects of language arts, social studies, and science. Student samples were the 1995 first and third grade cohorts followed through 1998, but the analysis included only students who completed the pre- and post-tests and who had attended either the Core Knowledge or control school continuously for the three-year period of study.

Three major findings emerged from the Datnow et al. (2000) study: (a) the CKS was successfully implemented in three of the four schools as evidenced by more curricular coordination and an increase in the use of project-based teaching; (b) the Core Knowledge students' basic skills achievement scores were consistent with the achievement scores of the control group students; and (c) the Core Knowledge students scored significantly higher on the subject tests of Core Knowledge content than the control group students.

The relevant finding to this study is that there was a modest achievement gain in the basic skills assessment for the third-grade cohorts in the high-implementing schools in Florida: math ($r = .08$), reading ($r = .17$); Texas: math ($r = .37$), reading ($r = .29$); and Washington: math ($r = .44$), reading ($r = .38$). The researchers conjectured that the better relative performance in the later years could be accounted for by either a cumulative effect of the content focused curriculum on general academic skills, or a content rich curriculum is more important for mid-elementary years and could explain the better relative performance in later years.

This study examined whether the Core Knowledge was an effective comprehensive school reform model. Three of the four schools continued to implement the Core Knowledge after a five-year mark and "reliable and consistent Core Knowledge implementation fostered curricular coherence and made instruction more hands-on and content rich for students" (Datnow et al., 2000, p. 187). Yet, it is curious that the researchers would compare the student outcomes on subject tests of Core Knowledge content. If a comprehensive school reform (CSR) model is to

be judged effective in improving student achievement, then nationally normed assessments would be of more importance than learning if the control group students knew the specific content of the CKS.

Borman, Hewes, Overman, and Brown (2003) conducted a meta-analysis of twenty-nine CSR models reviewing all the known research as of 2002 on the student achievement effects of the most widely implemented, externally developed school reform programs. A program was included if it was a school-wide reform, had at least one evaluative study completed, was disseminated by external developers, and was replicated in ten or more schools. The results of the meta-analysis ranked the comprehensive school reforms with a weighted mean effect size into four categories based on the quality of evidence in studies, the quantity of evidence in terms of the number of studies, and if the studies showed statistically significant and positive results. In Table 1 the CSR models were identified according to effectiveness: three CSR models were identified with the strongest evidence of effectiveness; three CSR models were identified as showing highly promising evidence of effectiveness; six CSR models were identified as showing promising evidences of effectiveness; and seventeen CSR models were identified as in the greatest need of additional research. As seen in Table 1, the CKS ranked in the greatest need of additional research category with six studies -- too few to establish statistically reliable and generalizable results.

Table 1: Meta-analysis ranking CSR models, Boreman et al. (2015)

Model	Studies <i>N</i>	Observations <i>N</i>	<i>d</i>
<i>Strongest Evidence of Effectiveness</i>			
Direct Instruction	49	182	0.21
School Development	10	25	0.15
Success for All	42	173	0.18
<i>Highly Promising Evidence of Effectiveness</i>			
Expeditionary Learning Outward Bound	6	40	0.19
Modern Red Schoolhouse	6	23	0.26
Roots & Wings	6	14	0.38
<i>Promising Evidence of Effectiveness</i>			
Accelerated Schools	6	50	0.09
American's Choice	2	27	0.22
ATLAS Communities	3	8	0.27
Montessori	2	7	0.27
Paideia	4	5	0.3
The Learning Network	3	38	0.22
<i>In Greatest Need of Additional Research</i>			
Audrey Cohen	1	1	-0.13
Center for Effective Schools	1	26	0.13
Child Development Project	2	2	0.12
Coalition of Essential Schools	3	6	-0.09
Community for Learning	1	3	0.15
Community Learning Centers	5	17	0.03
Co-nect	5	42	0.04
Core Knowledge Sequence	6	58	0.03
Different Ways of Knowing	2	3	0
Edison	5	209	0.06
High Schools That Work	45	64	0.3
High/Scope	4	23	-0.02
Integrated Thematic Instruction	2	2	0.24
MicroSociety	3	32	0.29
Onward to Excellence II	4	13	0.25
Talent Development High School	1	2	0.14
Urban Learning Centers	3	16	-0.03

Note: Effect sizes are presented as Cohen's *d*.

Only three of the 29 CSRs chosen for inclusion in the meta-analysis had more than 40 independent studies; one CSR had ten studies, eight had five or more studies, and the remainder – 18 – had less than four studies. This is an important study of the CSR models, but conflict of interest concerns can be raised. The CSR model Success for All, one of the three CSRs receiving the “Strongest Evidence for Effectiveness” ranking, was co-founded by Robert Slavin and Nancy Madden of Johns Hopkins University (Success for All Foundation, 2018, p. 2). Another researcher, Borman, co-edited a book in 2001 with Slavin, in which the chapter written by Slavin advocates for Title I as the funding engine for school reform through the Comprehensive School Reform program (Billig, 2003).

At the request of the Core Knowledge Foundation, Wedman and Waigandt (2004) analyzed five performance data sets comparing schools who teach CKS to one another in terms of overall outcome trends over six years, outcome trends for selected content areas, school performance based on the schools’ ethnic profiles, economic profiles, and school size. The researchers discussed their analysis of data sets using ANOVA, post-hoc analysis, and t-tests to compare outcomes related to the effects of independent variables, but to review the entire data set a request must be submitted to the Core Knowledge Foundation (Wedman & Waigandt, 2004). The single reference to a national comparison was to identify a trend in the data where scores were above the national average in comparison to the National Center for Education Statistics (NCES) data for the same time.

Sonnenschein, Baker, and Garret (2005) conducted an evaluation of the CK Pre-School Sequence implemented in ten Baltimore Head Start programs in 2004-2005. The study documented “how the children receiving the Core Knowledge instruction performed relative to a nationally representative group of children matched on age” (Sonnenschein et al., 2005, p. 7) by

using two measures: The Woodcock Johnson Tests of Achievement – III and the Core Knowledge Preschool Assessment Tool (CK-PAT). Children made statistically significant moderate to moderately strong increases in performance of academic skills from pre- to post-Woodcock Johnson assessments. The students' growth was at least comparable to that of the Woodcock Johnson normed group but showed significantly more growth than the norm in Oral Comprehension ($r = 0.54$), Quantitative Concepts ($r = 0.28$), and Oral Language Cluster ($r = 0.20$) subtests. The CK-PAT assessments are aligned with the CK Pre-School Sequence to measure student growth in the eight program domains. The study documented significant individual student growth on sixteen of the seventeen measures, as well as significant growth in three composite areas: oral language, emergent literacy, and mathematics.

While the Sonnenschein et al. (2005) study involved less than one hundred pre-school children, it has bearing upon this study comparing student achievement according to curriculum taught. A supplemental analysis between the Woodcock Johnson normed scores and the CK-PAT standardized scores was conducted to evaluate the concurrent validity of the CK-PAT assessment. Statistically significant correlations were found between seven of the eight assessment tasks in oral language cluster ($r = 0.41$, $p = <.001$), mathematical reasoning cluster ($r = 0.51$, $p = <.001$), spelling and emergent literacy ($r = 0.56$, $p = <.001$), letter-word identification ($r = 0.44$, $p = <.001$) indicating that both measurement instruments documented significant student growth for children taught the CK Pre-School Sequence in “skills and knowledge that children of the age across the country are expected to master” (Sonnenschein et al., 2005, p. 3). If the CK Pre-School Sequence, as a pre-school curriculum, significantly improves student achievement as documented on the Woodcock Johnson, then similar achievement gains could be

documented on other standardized assessments for elementary and middle school students taught the CKS.

Houghton Mifflin Harcourt's Journeys

The State Board of Education is charged by legislative statute to review and adopt a list of textbooks, as well as instructional materials, for use by public school districts and is to include more than one textbook and/or instructional materials for each grade level and subject area (OregonLaws.org, 2015). Individual school district boards are required to “adopt a list of textbooks and other instructional materials” (OregonLaws.org, 2015) from the State’s list and to provide those materials to teachers and students for use the year following adoption by the State. To support the implementation of the Common Core State Standards, the State Board of Education adopted English Language Arts (ELA) instructional materials in 2013 which are contracted for use with publishers through 2020 (ODE, 2013). Thus, a school district could provide for a kindergarten student entering school in 2014 a consistent ELA curriculum through the sixth grade.

The Fairmont School District’s adopted English Language Arts curriculum is Houghton Mifflin Harcourt’s Journeys, and began a roll-out of the materials in 2014. Prior to adoption of this curriculum, Fairmont School District used Houghton Mifflin’s Reading, Grade K-5 curriculum. A 2015 What Works Clearinghouse (WWC) Beginning Reading Intervention Report describes Houghton Mifflin’s Reading, Grade K-5 curriculum as step-by-step reading instruction “developing oral language, comprehension, phonemic awareness, decoding skills (phonics, analogy, context, and word recognition), fluency, reading comprehension, writing, spelling, and grammar” (U.S. Department of Education, 2015). Houghton Mifflin Harcourt, in their online A Research-Based Approach (2016) describes the foundation of reading in Journeys as one that

“requires developing multiple skills” (p. 37) in comprehension, fluency, phonological awareness and phonics, and vocabulary. The two curricula appear to have similar instructional goals.

While there is not yet a WWC report on Journeys, the Clearinghouse identified and reviewed ten studies of the Houghton Mifflin Reading curriculum published from 1983 through 2014. Given the similarities between the Houghton Mifflin Reading and Journeys curricula, as well as an uninterrupted adoption of Journeys from Houghton Mifflin Reading by the Oregon Department of Education and the Fairmont School District, a case could be made to review the literature surrounding the Houghton Mifflin Reading program as context for more recent research of the Journeys program. Unfortunately, the WWC was “unable to draw any conclusions based on research about the effectiveness or ineffectiveness of Houghton Mifflin Reading on beginning readers in grades K-3” (U.S. Department of Education, 2015) because none of the ten studies met the WWC group design research standards.

A review of the literature for either the Houghton Mifflin Reading or Journeys curricula reveals a troubling lack of peer-reviewed research for either program. By request for this research study, the Houghton Mifflin Harcourt Research and Validation department provided seventeen case studies for the use of the Journeys curriculum in various school districts nationwide, as well as two Journeys-specific reports produced by the Educational Research Institute of America (ERIA) and three Journeys-specific reports produced by Planning Research and Evaluation Services (PRES). These were not peer-reviewed documents; nor have they been subject to the editorial review and control exerted when commercially published. As such, the merits of this grey material must be carefully assessed. Nevertheless, in 2010 Farace and Schopfel described doctoral theses and conference reports as grey literature which is useful in research “as part of the overall evidence base” (University of Michigan Library, 2016, p. 1).

While the precedence for the use of grey material is established in higher education through the practice of including unpublished dissertations in literature reviews, the Internet provides open access to a huge repository of grey literature “such as reports, preprints, internal documents (memoranda, newsletter, market surveys, etc.), theses and dissertations, conference proceedings, technical specifications and standards, trade literature, etc., not readily available through regular market channels because it was never commercially published/listed or was not widely published” (Reitz, 2014, p. 1). Even though all the literature provided by Houghton Mifflin Harcourt is proprietary, and thus subject to assertions of conflict of interest, there is a significant difference between the case studies shared by Houghton Mifflin Harcourt and the reports produced by PRES or ERIA. The case studies compared year-to-year student achievement on standardized tests following the first year of implementation of the Journeys curriculum, while the research institute reports either evaluated the implementation of Journeys as a new curriculum or assessed the efficacy of various elements of the curriculum over time.

It can be surmised from the seventeen one-page case studies that several school districts across the nation implemented Houghton Mifflin Harcourt’s Journeys in the 2010-2011 school year. Each case study follows an established format:

- A district overview;
- A description of the state’s standardized assessment;
- Quotes from teachers and principals;
- Description so student success due to Journeys.

Table 2 includes the description of each district and a final quote and an average percentage increase as evidence of improvement following the implementation of Journeys. It is quite difficult to determine the precision of each case study because there were many factors not

addressed, such as information about specific schools in the study, whether the standardized assessment changed from one year to the next, or any other of a multitude of factors which impacted student achievement. Finally, each study ended with the identical statement, in bold type: “This case study provides evidence that Journeys will be effective in improving students’ reading proficiency in other districts with similar demographics and characteristics” (Houghton Mifflin Harcourt, 2011). As a marketing strategy, it is a clear declaration of cause-and-effect, but as research there is far less clarity of evidence offered about any association between the use of Journeys and improved student achievement scores.

Table 2: State-by State Journeys Implementation 2010-11

State	District	Student			Urban		Comparison of Student Achievement Percentage Gain
		School <i>N</i>	Enrollment <i>N</i>	Employees <i>N</i>	Suburban Rural	Grades	
Arizona	Ganado Unified School District #20	3	1,600	291	R	3, 4, 5	9% +
Illinois	City of Chicago School District 299	648	400,000		U	3, 4, 5	5% +
Illinois	Diamond Lake School District 76	3	1,100	119	S	3,4,5	> 2011
Massachusetts	Canton School District	5	4,151	374	S	3,4	5% +
Missouri	Carthage R-9 School District	9	4,151	563	S	3,4,5,6	5% +
Missouri	Lawson School District R-14	3	2,150	143	R	3,4,5	10% +
Nebraska	Fremont School District	12	4,450	396		3,4,5	8% +
Ohio	Barberton City School District	7	4,000	544	S	3,4,5	> 2011
Ohio	Benjamin Logan Local School District	3	1,926	243	R	3,4,5,	4% +
Ohio	Genoa Area Local School District	3	859	81	S	3,4,5	12%
Ohio	Springfield City School District	16	8,000	845	U	3,4,5	6%
Ohio	Westfall Local School District	3	1,600	168	R	3,4,5	2%
Pennsylvania	Apollo-Ridge School District	4	1,610	197	R	3,4,5	9% +
Pennsylvania	Forest Hills School District	3	1,995	266	R	3,4,5	4%
Pennsylvania	Pleasant Valley City School District	7	5,800	456	U	3,4,5	> 2011
Pennsylvania	York City School District	10	7,000	813	U	3,4	7%

Prior to the apparent nation-wide implementation of Houghton Mifflin Harcourt’s Journeys, ERIA conducted two quasi-experimental studies in the spring of 2009 to determine the effect of the program on students’ reading skills and strategy use. In the first of these studies, ERIA’s Report Number 366 (2009), Journeys’ Unit 6 was taught for three weeks for the first

time to one hundred and five third-grade and 98 fifth-grade students in seven schools in a western state. At the end of the three weeks, teachers participating in the study also completed a survey designed to identify “the fidelity with which they used the program materials. According to their completed questionnaires, teachers at grade 3 used the program for 12 to 15 days and for an average of 60 to 90 minutes per day. At grade 5 teachers used the program for 15 or more days and for an average of 60 to 90 minutes per day” (p. 3).

This first study used a pre-test and post-test design to examine the efficacy of the program in spelling, grammar, comprehension, and vocabulary. The researchers developed the test items for each grade level to match the learning outcomes of the unit being taught, and scrambled the order of the test items from the pre-test to the post-test. The third-grade assessment included forty test items with forty-one test items on the fifth-grade test. To assess the instruments’ internal reliability, the researchers used a Kuder-Richardson 20 analysis and observed “the posttest reliabilities for both the grade 3 and the grade 5 tests were .90” (p. 7). While there was no difference in the reliability coefficient for the third-grade instrument, the reliability coefficients for the fifth-grade pre-test was .72, but it was .90 for the post-test.

To determine the program’s efficacy, the researchers compared the percent correct on the pre-test to that of the post-test for the total test, as well as the sub-sections, using t-test analyses with the effect size calculated using Cohen’s *d* statistic. The difference in third grade correct scores from the pre-test to the post-test was statistically significant at $p = <.0001$ with a medium effect size value of $d = 0.60$. The difference in the fifth-grade correct scores from pre-test to post-test was also statistically significant at $p = <.0001$ with a large effect size value of $d = 1.0$. The researchers concluded: “Based on a highly reliable test designed to measure growth on the skills taught in a single unit of instruction to students who received instruction using the Unit 6

of the Journeys program is that the program significantly increases students' reading skills and strategy use" (2009, p. 15).

In the second study, ERIA Report Number 368A (2009), a control group and experimental group study design compared Journeys effectiveness in improving students reading skills and strategy use in the control group with the program's effectiveness in improving the reading skills and strategy use of English Language Learner (ELL) students. The study included 409 students in first and fifth grades with the control group classrooms' demographics matched as closely as possible to those of the experimental group classrooms at each grade level. A single Journeys unit was taught over a three-week period with a pre-test and post-test for the experimental groups, but only a post-test for the control group. The researchers used ANOVA to determine if there was a significant difference in the first-grade post-test scores between the sixty-five control group students and the one hundred thirty-seven experimental group students and Cohen's *d* statistic to determine the effect size. The control group's average percent correct score on the post test was 67.9% while the experimental group's average was 77.4%. The difference between the two groups' scores was significant at $p = <.0001$ with an effect size value of $d = 0.52$. Similarly, there was a $p = <.0001$ significance in the difference in the fifth-grade post-test scores between the seventy-four control group students and the one hundred thirty-three experimental group students with an effect size value of $d = 0.69$. Further analysis of the differences within the experimental group classrooms was conducted comparing ELL and non-ELL students' scores at both the first and fifth grades. While both groups of students' score gains were significant, the difference in the pre- to post-test scores for the first grade ELL students was significant at $p = <.005$ with an effect size value of $d = 0.75$. The difference in the pre-to post-

test scores for the fifth-grade ELL students was significant at $p = <.001$ with an effect size value of $d = 0.61$.

The three reports prepared by the Planning, Research and Evaluation Services, Inc. (PRES) included a preliminary pilot study, an analysis of student achievement data following the progression of the Journeys curriculum implementation, and a two-year randomized control trial. The Houghton Mifflin Harcourt Journeys: Preliminary Pilot Study Report (2011) followed the implementation of the Journeys curriculum in four elementary schools in Kentucky spanning grades K-5. The study included 1,893 participants and was designed “to gather quantitative and qualitative data so as to provide a comprehensive picture of program implementation as well as obtain preliminary outcome data on a diverse set of student and teacher outcomes” (p. 2).

While the study design included teacher and student surveys, data from existing student assessments already used in the schools, classroom observations, focus groups, and teacher activity logs, PRES notes “the pilot study was not designed to produce conclusive evidence on the effectiveness of the Journeys program” (p. 9). Nevertheless, the primary key findings presented in the pilot study examined student growth across two standardized assessments. Paired sample t-tests were used to analyze second through fifth grade pre- and post-test data on the PAS Reading Benchmarks Assessment with a 5% average increase from Fall, 2010 to Spring, 2011 which were significant at $p = < .05$; an effect size value is not provided. Students in kindergarten through second grade also showed a significant positive increase in independent reading levels from the prior school year to the Spring, 2011 on the Pearson’s Developmental Reading Assessment, 2nd Edition (DRA2) which was also significant at $p = < .05$. The study also observed a positive relationship between a high level of teacher implementation and improved student scores on the PAS Reading Benchmarks, but not on the DRA2, and that both teachers

and students reported the Journeys program helped improved reading skills, but not writing skills.

The second PRES report shared by Houghton Mifflin Harcourt is a two-year randomized control trial designed to assess the efficacy of the Journeys program in two school years, 2011-12 and 2012-13, with a final sample size of 700 students who participated in both years of the study.

Criteria for school participation included

- geographical diversity across states,
- inclusion of urban and suburban schools,
- school level willingness to teacher level random assignment of curriculum,
- student mobility rate less than 20%,
- no other major reading initiative undertaken during the trial, and
- the school's reading curricula matched the comparison programs identified as a contrast to Journeys.

The study was conducted in six schools in Arizona, Rhode Island, Louisiana, and the District of Columbia. Forty-six teachers participated, with twenty-six randomly assigned the use of Journeys prior to the 2011-12 school year. The twenty control teachers continued to use the reading curricula used in their schools prior to the 2011-12 school year. Kindergarten through second grade students participated in the study in 2011-12 and these students were followed to the first through third grades in the 2012-13 school year. Treatment teachers were provided training in the use of Journeys and implementation guidelines during both school years. The study used the norm-referenced achievement test Iowa Test of Basic Skills (ITBS) – Form C to assess reading comprehension (K-3), reading words (K), word analysis (K-3), vocabulary (K-3), spelling (1-3) and language conventions (1-3). Students were surveyed to measure attitudes

about school, reading activities, perceived reading ability, as well as effort and motivation.

Teachers were surveyed to measure knowledge of and attitudes about effective teaching practices, reading, and language arts. Finally, the researchers conducted class visits to observe teacher practices in implementing key elements of the Journeys program.

The ITBS experimental group student scores were analyzed with paired t-tests for change from pre-test to post-test scores, and showed “students who were taught with Journeys exhibited significant learning gains from pre (Fall 2011) to post-testing (Spring 2013) on vocabulary, word analysis, reading comprehension, spelling and language art” (Resendez & Azin, 2013, p. 33).

The significance level for each individual subtest at the student level is $p < .001$, but the sample t-tests analysis for treatment and control groups shows statistically significant differences in favor of the Journeys program in only two of the five ITBS measures: vocabulary, $p < .002$, and reading comprehension, $p < .02$, and the “effect sizes obtained can be classified as small to moderate ($d = .15$ to $.39$)” (Resendez & Azin, 2013, p. 3.)

The final PRES report provided by Houghton Mifflin Harcourt is a statistical analysis of three-year student achievement trends on the Texas Assessment of Knowledge and Skills (TAKS) and its replacement assessment, the State of Texas Assessments of Academic Readiness (STAAR) designed to compare Journeys’ efficacy on third through fifth grade students’ literacy skills with the efficacy of two other reading curricula. The other curricula are not named, but referred to as Reading Program 1 and Reading Program 2. The student data was drawn from fifty-five schools using Journeys matched along twelve school-level characteristics to fifty comparison schools not using Journeys. The school-level characteristics matched for enrollment, percentages of students who were economically disadvantaged, had limited English proficiency, qualified for special education, identified as gifted or at-risk, as well as percentages of race or

ethnicity and the rate of student mobility. Because the TAKS data is not directly comparable to the STAAR data, longitudinal analyses were not conducted. “Instead, researchers examined whether differences existed within each grade cohort: the 2013 fifth graders (STAAR), 2012 fourth graders (STAAR), and 2011 3rd (sic) graders (TAKS) after controlling for demographics” (Resendez & Azin, 2014, p. 14). Using linear modeling, third grade data (year one of Journeys implementation) showed a statistically significant difference in Journeys students outperforming non-Journeys students on the TAKS, $p < 0.05$ with an effect size value of $d = 0.001$.

Controlling for third grade baseline performance with analysis of covariance, the fourth-grade data (year two of Journeys implementation) showed that while Journeys students tended to outperform non-Journeys students on the STAAR, it was not statistically significant, $p = > .05$. Finally, controlling for both third and fourth grade reading performance, the fifth-grade data (year three of Journeys implementation) showed Journeys students had significantly higher STAAR scores than non-Journeys students, $p = < .05$ with an effect size value of $d = 0.003$.

Taken together, the ERIA and PRES reports seem to indicate the use of the Journeys program is positively associated with improved achievement scores in comparison to other reading curricula. Nevertheless, it is important to note that none of these reports, including the randomized control trial study, have been evaluated or reviewed by the Institute of Education’s What Works Clearinghouse.

Finally, though not a study of the program’s efficacy on student outcomes, Journeys is included in a recent quantitative content analysis of fourth-grade reading texts, as well as those of Macmillan/McGraw-Hill’s Treasures and Pearson Scott Foresman’s Reading Street, to determine the percentage of informational text presented in the student basal readers (Braker-Walters, 2014). The Common Core State Standards in English and Language Arts (CCSS-ELA)

states the means to meeting the K-5 reading standards “requires a 50-50 balance between information and literary reading. Informational reading includes content rich nonfiction in history/social studies, sciences, technical students, and the arts” (CCSS, 2016). Braker-Walters’ study employed hand-coding of the text passages per genre (literary or informational), number of reading selections, number of pages devoted to type of text, and publisher. The reading selections reviewed were limited to the texts in the student-read basal readers and category frequencies were compared for significant differences between publishers using a chi-square goodness-of-fit test. Even though Macmillan/McGraw-Hill’s Treasures provided the most reading texts (103) compared to Journeys’ (50), there was little difference in terms of the percentage of literary and informational selections between all three curricula. An average of only 31% of the selections were informational reading texts. It is important to note, Journeys offered the fewest informational text selections (28) and the least pages of informational text (65) (Brakers-Walters, 2014).

Standardized Assessment as an Accountability Measure

Charter schools in Oregon participate in the annual standardized tests as the means for the Oregon Department of Education and the sponsoring school districts to equitably compare charter schools’ performance with that of traditional public schools. This is a legal requirement from which districts may not contractually exempt the sponsored charter schools (ORS 338.115 (k)). Over the duration of a charter school’s contract, the measures of student performance on the state standardized tests are, annually and cumulatively, key factors in a school district’s decision to renew the charter contract. Nevertheless, the accountability inherent in the legislation and practice of comparing charter school students’ performance to traditional public-school students’

performance on the same standardized assessment minimizes the differences in curricula, which is the fundamental reason for the charter school's contract with the district.

It is not surprising charter schools are measured with the same ruler as traditional public schools since the use of standardized assessment for the purposes of accountability has a long history in American public education. From early in the 20th century when Stanford University teaching candidates were urged to incorporate the standard scales in order “to measure the effectiveness of the work they do” (Cubberly, 1919, p. 450) to the stringent requirements in the 21st century No Child Left Behind legislation, accountability and student performance are entwined in our social discourse about school. Public education has institutionalized political ideology through the theory of standards-based reform (SBR). The hypothesis that “greater accountability means better student performance” (Carnoy, Elmore, & Siskin, 2003) rests on clearly articulated student learning outcomes charted in curriculum frameworks with specific instructional targets set for teachers and which are supported by aligned curriculum materials and assessments (Polikoff, 2012).

The body of literature examining standardized assessment is diverse and extensive, but to understand the appropriateness of holding charter schools accountable with the same measures used for traditional public schools, it is important to focus on the literature which considers the associations between standardized assessment, student achievement, and curriculum. Research investigating the associations between standardized assessment and student achievement has been of substantial focus since the initial passage of NCLB, but there is also a body of work exploring relationships between standardized assessment and curriculum.

The NCLB legislation prompted more than a decade of research exploring the impact of high-stakes standardized assessment on student achievement. Carnoy and Loeb (2002) examined

the relationship between state accountability approaches and student achievement as measured by eighth grade NAEP gains in math scores from 1996 to 2000, by ninth grade completion rates, and by high school graduation rates. Employing a recursive statistical model to estimate accountability implementation with student achievement as a function of accountability, the researchers developed an index ranking accountability-strength from zero to five based on the outcomes of high-stakes tests to penalize or reward schools. “A test is high-stakes when its results are used to make important decisions that affect students, teachers, administrators, communities, schools, and districts” (Au, 2007, p. 258). Carnoy and Loeb (2002) found states which either did not set standards or conducted statewide assessments were ranked at a zero, while states which tested in the elementary and middle grades, pressured schools to improve student achievement using severe sanctions such as principal transfer or school closure and required a minimum competency exam for high school graduation, were ranked with a five. The researchers found student achievement gains were larger in states designated “strong-accountability” than in states designated “weak-accountability.” “With a mean gain of 4.8 percentage points and a standard deviation of 3.6 in average state proportion scoring at or above basic skill levels, the increase in gain from raising the external pressure on schools by the state appears to be substantial” (Carnoy & Loeb, 2002, p. 313).

Additional literature reports positive correlations between increased use of standardized assessments as an accountability tool and student achievement gains. Hanushek and Raymond (2005) found the use of consequential-accountability, wherein the state both publicized school-level performance on standardized tests and attached consequences to the performance, evidenced a 3.24 point-gain in student achievement associated with consequential-accountability practices. Neal and Whitmore Schanzenback (2010) examined student achievement scores from

two separate high-stakes accountability systems in the Chicago Public Schools, the 1996 district-led reforms, and the 2002 No Child Left Behind reforms. Neal and Whitmore Schanzenback found a positive correlation between the implementation of accountability systems and an increase in overall average math and reading scores with a two-year fifth grade mean effect of 0.066 in math and 0.43 in reading from 1996 to 1998, and a one-year fifth grade mean effect of 0.94 in math and 0.61 in reading from 2001 to 2002. Finally, in a working paper by Ballou and Springer (2008), the effects of high-stakes accountability efforts on student achievement were compared within grade levels across low-stakes and high-stakes years using longitudinal, student test score data from seven states between 2002 and 2006. Using the Northwest Educational Assessment (NWEA) Growth Research Database to determine annual yearly progress scores, Ballou and Springer state “there has been a tendency for scores to rise across the board, accompanied by a redistribution of achievement gains from high-performing to low-performing students” (p. 23).

Conversely, there is a large body of literature suggesting a negative correlation between high-stake testing and student achievement. Using their Accountability Pressure Rating (APR) index developed in their 2006 research, Nichols, Glass, and Berliner (2012) re-examined the impact the No Child Left Behind Act (NCLB) high-stakes testing pressure had on student learning in twenty-five states. The APR was designed to measure the comparative pressure of states’ standardized testing policies enactment and implementation. The APR was derived by asking 300 graduate students to use a comparative judgements method to assign a comparative weight (on a scale of 1 to 7) to two states’ portfolios of complex qualitative information found in legislative documents, state-generated accountability reports, and newspaper articles and editorials documenting the implementation and impact of state testing policies. The researchers

applied to the students' ratings the least-squares solution for uni-dimensional scale values which resulted in a scale "ranging from .54 to 4.78" (Nichols et al., 2012, p. 5) to identify a state's level of testing pressure. The researchers examined the associations between student achievement in fourth and eighth grade math and reading as measured by the National Assessment for Education Progress (NAEP) 2002-2009 tests and the states' APR indexes through correlation and regression analyses. Interestingly, Nichols, et al. found

data suggest that test-related pressure is significantly and positively correlated with state poverty index (percent poverty in state). That is, states with greater number of individuals living in poverty also tended to employ test-related practices that exerted greater amounts of pressure. The nation's poorest children, and the teachers who teach them, tend to feel more pressure when it comes to high-stakes tests than their more privileged contemporaries. When disaggregated by SES and race, data suggest that the relationship between APR and NAEP performance is mixed. In terms of SES, high-stakes testing pressure has no connection to NAEP performance in math. By contrast, APR is more strongly and negatively connected with NAEP performance in reading, especially for low-income students. (2012, p. 24)

The correlations between APR and NAEP reading are most negatively related for low SES eighth grade students in 2003 ($r = -.370$), in 2005 ($r = -.336$), and in 2007 ($r = -.245$), but consistently negatively associated for low SES fourth grade students throughout the study: 2002 ($r = -.176$), 2003 ($r = -.282$), 2005 ($r = -.279$), 2007 ($r = -.234$), and 2009 ($r = -.214$).

Other researchers have also found negative correlations between increased accountability and student achievement gains. Sims (2012) examined the relationship between high-stakes accountability and racial sub-group rules in California which produced "future test score

decreases for the schools as a whole as well as for the minority students at those schools” (p. 264). When examining the role subgroup cutoff scores prefigure in a school’s failure to meet accountability standards, Sims found that in the simple counting of subgroups the addition of a “subgroup appears to raise the probability of failure by 12 percentage points” (p. 264). When examining these same schools where a subgroup is at least 15% of the student population, “two results stand out: each of the subgroups has a positive, statistically significant effect on the probability a school fails, but the coefficients suggest the effects are not the same magnitude for all subgroups” (p. 269). Furthermore, the examination of student achievement considering a school’s failure-performance “suggest that failure to meet a performance standard reduces future student performance by approximately 0.12-0.13 standard deviations, with similar effects for math and reading” (p. 269).

Reback (2008) examined whether accountability systems influence student achievement through a specific focus on improving the achievement of students who do not quite meet the standards; i.e., students at the test margin. Reback analyzed individual student test score data and school-level accountability data from the Texas Assessment of Academic Skills (TAAS) for 1992 through 1998 and found when the

improvement in a student’s expected math performance is associated with a .01 greater improvement in the probability that the school attains a higher rating, then this student will, on average, score .013 standard deviations higher in the math score distribution of students with similar prior year scores. To put the magnitude of this result in perspective, a one standard deviation in this incentive measure is associated with approximately a .007 standard deviation increase in a student’s place in the statewide achievement distribution. While that may not seem very large, it is important to keep in mind that this

is a within-school effect from just one year of schooling. Reading performance incentives within the school are also connected to students' reading performance: the estimated coefficient for reading incentives equals .954, which implies that a one standard deviation change in reading incentives leads to about a .009 standard deviation increase in a student's place in the statewide achievement distribution. (p. 1407)

Reback postulates the greatest impact from these short-term gains in student achievement are to the detriment of both low-performing and high-performing students through year-to-year changes in resource investment and teaching focus.

The literature supported a positive association between strong-accountability states (Carnoy & Loeb, 2002), consequential-accountability (Hanushek & Raymond, 2005) or high-stakes accountability systems (Ballou & Spring, 2009; Neal & Whitmore Schanzenback, 2010) and student achievement giving credence to the idea that holding charter schools to the same or higher accountability standards would result in improved student achievement. The literature also indicated a negative correlation between student high-stakes testing and lower student achievement in reading for lower income students (Nichols, 2012) and with the systemic higher-probability of a school failing to meet the accountability measures due to the inclusion of racial sub-groups (Sims, 2012). The literature also examined the detrimental effect of high-stakes testing on all students when teachers focus on students at the test margin (Reback, 2008). Lacking a generally agreed upon conclusion that higher accountability measures result in improved student achievement in traditional public schools undermines the idea that this would hold true for charter schools.

The body of literature examining school accountability measures' effects on student achievement is considerable, but another avenue of research which examines the relationship

between accountability measures and curriculum provides additional context for appraising the soundness of using the same assessment tools for charter schools and traditional public schools. Much of the curriculum-oriented research over the past decade has emphasized the effect of standardized assessments' influence on instructional decisions and pedagogy (Au, 2007; Diamond, 2007), as well as the alignment of curriculum to state standards (Polikoff, Porter, & Smithson, 2011; Srikantaiah, 2009). Reviewing the literature is important given that charter schools exist primarily to offer either a different curriculum, different pedagogy, or both.

Au (2007) conducted a qualitative meta-synthesis of 49 studies examining the impact of high-stakes assessment on curriculum, particularly the frequency and type of curricular changes made after the institution of high-stakes testing. Using a form of thematic meta-analysis through textual data coding, Au's findings "suggest that there is a significant relationship between the implementation of high-stakes testing and changes in the content of a curriculum, the structure of knowledge contained within the content, and the types of pedagogy associated with communication of that content" (2007, p. 262). The dominant theme of changes to the content of the curriculum is one of narrowing the content; 69.4% of the forty-one studies indicating an overall change to curriculum content reported non-test subjects were dropped from the content to focus on tested subject matter. Of those studies reporting content contraction, 49% indicated a "fragmentation of knowledge ... manifested in the teaching of small, individuated and isolated test-size pieces, as well as teaching in direct relation to the tests rather than in relation to other subject matter knowledge" (Au, 2007, p. 262). Only 20.4% of these same studies indicated an improved integration of knowledge. Finally, 77.6% of the forty-nine studies reported changes in pedagogy with a majority of those changes indicating an increase in teacher lecturing and "direct transmission of test-related facts" (Au, 2007, p. 263).

Diamond, using the first-year data collected through the multi-year Distributed Leadership Project, examined case studies to identify “how and to what extent accountability policies influence instruction” (2007, p. 286). The Distributed Leadership Project design included formal teacher and classroom observations, teacher interviews, and shadowing school leaders in eight schools for 50-70 days; 105 classes taught by 47 different teachers were observed. The observations focused on academic tasks, content, and pedagogy with analyses conducted through computer software Nu*DIST to ascertain patterns by topic codes such as who or what influenced the lesson as well as the focus of the lesson. Diamond found 31% of the teachers’ interviewed reported high-stakes tests was one of many influences on lesson content. These teachers focused on the test subjects, directly taught specific knowledge within math and language arts, covered material quickly in order to finish the content prior to the testing window, and reported spending more time on test preparation. Interestingly, Diamond found the teachers reported high-stakes testing had a greater impact on content than on pedagogy with only 19% of the teachers indicating tests influenced instructional strategies.

The Center on Education Policy (Srikantaiah, 2009) conducted case studies of 18 schools in Illinois, Rhode Island, and Washington to examine the impact of state accountability systems and found “state and federal policies are having a significant impact on curriculum and instruction” (p. 2). The case studies included interviews of district superintendents, principals, teachers, students, and parents, as well as formal observations of 105 classrooms. Educators in each state reported an increased effort to align curriculum content to state standards, but focused instruction on material covered on the standardized tests; when the test items and state standards were misaligned, teachers reported they were more likely to cover material assessed on the state tests. Educators and parents in all three states reported “the emphasis on teaching tested content

has diminished time available for other subjects and activities” (Srikantaiah, 2009, p. 2) with a concurrent narrowing of the curriculum breadth, as well as depth. Finally, educators expressed a need for accountability measures other than standardized assessments as a counter-balance to the tests’ limitations for English language learners, students with special needs, and those schools in high-poverty districts.

Polikoff, Porter, and Smithson (2011) explored the extent to which state standards are aligned with state assessments under the No Child Left Behind legislation and the nature of the alignment or misalignment. Polikoff et al. (2011) conducted a secondary analysis of data gathered by the Council of Chief State School Officers State Collaborative on Assessment and Student Standards thirty-one state database of Surveys of Enacted Curriculum (SEC). Data from nineteen states where both a standards document and an assessment document were available for at least one grade were coded for content and cognitive demands. A total of one hundred thirty-eight paired documents across English language arts (ELAR), math, and science were analyzed and coded by three to five analysts. Each analyst’s coding data was then converted to proportions of total test/standard content and each test item or standard was weighted based on the number of points it was worth on the assessment. An average of the proportions was calculated across the analysts’ matrices. The reliability for three (ELAR) coders across grades for two states, using Cohen’s *d*, was $d = .74$. with the mathematics’ three-coder reliability at $d = .78$. Finally, the average proportions were converted to an index indicating the extent to which the two documents were aligned. The researchers found the “average alignment indices for state standards and assessments were below .30 in mathematics and science and below .20 ELAR. No alignment index was greater than .50 for any state, grade, or subject analyzed” (Polikoff et al., 2011). Examining the nature of the misalignment between the standards and assessment

documents revealed disjointedness in topics, cognitive demand levels, or both that “comprised moderate to large proportions of the total content targeted. Substantial proportions of the topics specified in the standards were not tested at all, especially in ELAR. Perhaps more alarmingly, roughly a quarter of the content on typical tests was not reflective of topics that were mentioned at all in their corresponding grade-level standards” (Polikoff et al., 2011, p. 990).

Considering the literature examining high stakes testing accountability systems and curriculum or pedagogy consistently indicates that instructional content is narrowed to the test subjects (Au, 2007; Diamond, 2007; Srikantaiah, 2009), that instructional practices are changed to teach material in more fragmented, test-oriented pieces (Au, 2007; Diamond, 2007), and that there is often a substantial misalignment between the standards and assessments (Polikoff et al., 2001), it is questionable that standardized tests would be an appropriate accountability measure for charter schools. It is possible that in Oregon a charter school’s performance is evaluated with an assessment tool which is unlikely to measure the curriculum’s learning outcomes, as well as inadequately measure whether students meet state standards.

CHAPTER 3

Methodology

The purpose of this study was to investigate the impact of two curricula, the Core Knowledge Sequence (CKS) and Houghton Mifflin Harcourt's Journeys, on student achievement on the Smarter Balance Assessment (SVAC) in English language arts (ELA). The study was a secondary-data analysis examining three years of SBAC scores for students in the Fairmont School District. This chapter describes the research methods of the study, including study design, a description of the SBAC instrument, information on sampling and participants, data analyses, role of the researcher, and ethical considerations.

Design

A multiple regression analysis was used to examine student achievement as measured by the SBAC between students who were taught the CKS curriculum and students taught the Journeys curriculum. EAST Charter School was purposefully chosen to serve as the treatment school where the CKS was taught, while the ten elementary schools and three middle schools in the Fairmont School District were selected as the comparison control schools where Journeys was taught. The decision to conduct a secondary-data analysis over other possible methods rests primarily on the goal of the study which was to assess discrete results (Nardi, 2003). While a small-scale randomized control trial is the ideal design to evaluate specific outcomes when comparing treatment and control conditions in K-12 education, Dritz-Esser, Bass, and Stark (2014) point out "the sample sizes are impractical" (p. 593) due to budget and time constraints.

Sampling & Participants

The study was conducted in the Fairmont School District. The district encompasses ten kindergarten through fifth-grade elementary schools, three sixth through eighth-grade middle schools, two ninth through twelfth-grade high schools, one alternative high school, one online academy, and one kindergarten through eighth-grade charter school (Fairmont School District, 2015). The total 2015-2016 student enrollment according to the district report card was 12,615 students (Fairmont School District, 2016). The sample was third through sixth grade students who took the SBAC ELA assessment in 2014-2015, 2015-2016, or 2016-2017.

Smarter Balance Assessment Instrument

The U.S. Department of Education, through a Race to the Top grant, awarded the Smarter Balanced Assessment Consortium (SBAC) a \$175 million, multi-year grant to develop a comprehensive assessment system for mathematics and English language arts (ELA). The Consortium partnered with the Center for Research on Evaluation, Standards, and Student Testing (CRESST) in the Graduate School of Education and Information Studies at the University of California, Los Angeles (UCLA) and was governed by twenty state participants and three affiliate members. The Consortium's primary goal was to assist in the preparation of students for success in college and careers by creating an assessment aligned with the Common Core State Standards. Oregon adopted the Common Core State Standards in June 2010 and began the implementation of the SBAC assessment in 2014. Relevant to this study was the crosswalk of both the Core Knowledge Sequence and Journeys curricula with the Common Core State Standards (see Appendix A).

Given the broad nature of the Common Core State Standards and the SBAC assessments, most relevant to this study was the SBAC ELA claim that third through eighth grade students

would demonstrate progress toward college and career readiness in English language arts and literacy and the measurements associated with this claim. Progress towards the primary ELA claim is measured by inferences about student knowledge and skills through their performance on test items in four domain-specific claims:

ELA Claim 1: Students can read closely and analytically to comprehend a range of increasingly complex literary and informational texts. There are fourteen assessment targets delineated for grades 3-8 and a student's performance across the test items and tasks results in a Total Reading Score.

ELA Claim 2: Students can produce effective and well-grounded writing for a range of purposes and audiences. There are nine assessment targets delineated for grades 3-8 and a student's performance across the test items and tasks results in a Total Writing Score.

ELA Claim 3: Students can employ effective speaking and listening skills for a range of purposes and audiences. There is only one assessment target for Claim 3 and a student's performance on these test items contribute to the overall ELA Total Score.

ELA Claim 4: Students can engage in research and inquiry to investigate topics, and to analyze, integrate, and present information. There are three assessment targets delineated for grades 3-8 and a student's performance across the test items and tasks results in a Total Research/Inquiry Score.

Wixson's 1999 revision of Norm Webb's Depth of Knowledge (DOK) Levels of Cognitive Difficulty taxonomy is used to assess a student's ability to demonstrate conceptual understanding by applying their knowledge and skills to a new task in each assessment target. Table 3 includes the four levels of DOK used to assess student performance on Oregon Department of Education (ODE) ELA SBAC test items.

Table 3: Depth of Knowledge Taxonomy and ODE SBAC Achievement Levels

Depth of Knowledge Taxonomy	Oregon Department of Education SBAC Achievement Levels
<i>Level 1</i>	
Recall and Reproduction	Demonstrates limited to no mastery of knowledge and skills Does not meet proficiency standards
<i>Level 2</i>	
Skills & Concepts	Demonstrates inconsistent or partial mastery of knowledge and skills Does not meet proficiency standards
<i>Level 3</i>	
Strategic Thinking & Reasoning	Demonstrates adept knowledge and skills Meets proficiency standards
<i>Level 4</i>	
Extended Thinking	Demonstrates exceptional knowledge and skills Exceeds proficiency standards

Students performing at a level 3 or level 4 were considered on track for college and career readiness, but these levels were determined by each student's scaled score performance on the SBAC achievement standards. The achievement standards are also known as cut scores indicating the point on a scale that is the start of an achievement level. For example, a grade 3 English Language Arts/Literacy achievement standard of 2367 means that any student with a scaled score between 2367-2431 is in the level 2 achievement standard range (ODE, 2015, p. 1) and not on track for college or career readiness.

Data Collection

A secondary-data set from the ODE was used for this study through a State of Oregon Research Agreement / Data Use Agreement (see Appendix B). ODE provided data for 2015, 2016, and 2017 for third, fourth, fifth, and sixth grade students enrolled in the Fairmont School District who participated in the SBAC ELA assessment in any of the three years. The

demographic data included gender, ethnicity, English language learner status, student with disability status, and qualifying status for free or reduced lunch. The demographic data also included school attended as the means to identify whether the student was taught the CKS or Journeys curriculum. The assessment data included the SBAC ELA achievement scores as well as the SBAC ELA scaled scores.

Timeline

Proposal approved: July 23, 2017

Institutional Review Board (IRB) approved: March 21, 2018

Request for Oregon Department of Education data approved: May 9, 2018

Oregon Department of Education contract signed: June 21, 2018

Received data from Oregon Department of Education: June 27, 2018

Data analyses: June 29, 2018 and July 24, 2018

Submitted draft to dissertation committee: August 26, 2018

Defended dissertation: September 6, 2018

Data Analysis

This study used a multiple regression (MR) analysis to determine the relationships between a single dependent variable SBAC ELA achievement scores, and multiple independent variables (curriculum, gender, ethnicity, English language learner status, student with disability status, and socio-economic status). The data were analyzed by the discrete years 2015, 2016, and 2017 in which the SBAC assessment was conducted to avoid repeat data across years. The MR analysis determined how much of the variance in SBAC ELA scores was explained by instruction in either the CKS curriculum or Journeys curriculum. The MR model was determined to be a good fit for the data and met the assumptions described by Laerd (2018).

1. The independence of errors was assessed with the Durbin-Watson statistic.
2. Linearity is tested through examination of scatterplots of the dependent variable with each independent variable and with the dependent variable with the combined independent variables. The independent variables in this study were categorical; linearity could not be assessed accurately.
3. 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.
4. Outliers were detected using casewise diagnostics. A greater number of outliers than could be anticipated by chance in each of the 2015, 2016, and 2017 analyses prompted a review of the data to determine the cause (Laerd, 2018).

Independent variables.

The following independent variables were operationalized as predictor variables in the data analyses (see Table 4).

Table 4: Variable Operationalization

Independent Variable	
Operationalized as a Categorical Variable	
<i>Gender</i>	
Male	0
Female	1
<i>English Language Learner Status</i>	
No	0
Yes	1
<i>Student with Disability Status</i>	
No	1
Yes	
<i>Socio-economic Status</i>	
Eligibility for free and reduced lunch	
No	0
Yes	1
<i>Ethnicity</i>	
White	0
Asian	1
Black/African American	2
Hispanic	3
Pacific Islander	4
Multi-racial	5
Indigenous	6
<i>Curriculum</i>	
School Attended	
EAST Charter School	0
Alcott Elementary	1
Babbitt Elementary	1
Carle Elementary	1
Dahl Elementary	1
Estes Elementary	1
Fitzhugh Elementary	1
Gaiman Elementary	1
Henry Elementary	1
Ibbotson Middle School	1
Jacques Middle School	1
Kipling Middle School	1
L'Engle Elementary	1
FAIRMONT Online Academy	1
McCloskey Center	1
Norton Elementary	1

Dependent variable.

The SBAC English language arts scaled score was conceptualized as an indicator of level of proficiency in meeting the Common Core State Standards in ELA for third, fourth, fifth, and sixth grades.

Research Ethics

I am a graduate student conducting this study as the final step toward a doctoral degree (Ed.D.) at George Fox University. I am also employed at EAST Charter School as the Executive Director with many of the same responsibilities and duties as fulfilled by a traditional district administrator. I am responsible for continued improvement in student achievement on state standardized assessments, as well as for instructional fidelity to the CKS curriculum. While the George Fox University Research Ethics Institutional Review Board (IRB) approved this study, additional care regarding research ethics and personal integrity were required to insure the objectivity of the findings in the study. Student anonymity was critically important, and as personally identifiable information was not relevant to the research question, it was not provided in the data set from ODE. All data sets and data output was stored on an encrypted flash drive.

In the normal course of my duties, I access achievement data for students enrolled at EAST Charter School, but for this study, I used only the historical data provided by the ODE as an additional safeguard for students' privacy.

EAST Charter School's contract with the Fairmont School District was renewed in April 2018. If the study's findings had been available prior to April 2018, the potential would have existed for a conflict of interest during the district's charter renewal evaluation of EAST Charter School which took place from August 2016 through November 2017.

CHAPTER 4

Results

The purpose of this secondary-data analysis was to examine student achievement as measured by the Smarter Balance Assessment (SBAC) English language arts (ELA) assessment between students who were taught the Core Knowledge Sequence (CKS) curriculum and students taught using the Houghton Mifflin Harcourt's Journeys curriculum. This study used a multiple regression (MR) analysis to determine the relationships between a single dependent variable, SBAC ELA achievement scores, and multiple independent variables (curriculum, gender, ethnicity, English language learner status, student with disability status, and socio-economic status). This chapter details the results of the data analyses. It includes information on the demographics of the sample, descriptive statistics, and the results of MR test assumptions and answers the research questions.

Sample Demographics

The sample used in this study was third through sixth grade students in the Fairmont School District enrolled in a district elementary, middle, or charter school during 2015, 2016, or 2017. The total sample across the three years was 11,493 students; 5,147 (45%) students took the SBAC ELA assessment multiple years with 6,346 (55%) students taking the test only one of the three years. The 2015 sample included 3,775 students; 1,554 (41%) of students qualified for the federal free and reduced lunch program, 407 (11%) of students were identified as students with disabilities, 541 (14%) were English language learners, and 2,209 (59%) identified as white with another 949 (25%) identified as Hispanic. The 2016 sample included 3,832 students; 1,710 (45%) of students qualified for the federal free and reduced lunch program, 401 (10%) of students were identified as students with disabilities, 943 (25%) were English language learners,

and 2,193 (57%) identified as white with 992 (26%) identified as Hispanic. The 2017 sample included 3,886 students; 1,608 (42%) of students qualified for the federal free and reduced lunch program, 400 (11%) were identified as students with disabilities, 945 (24%) were English language learners, and 2,244 (58%) identified as white with 1,050 (27%) identified as Hispanic. Table 5 offers descriptive statistics for the sample. The descriptive demographics for the sample are relatively constant except for the English language learner category. There was one-year increase of 402 (43%) students identified as English language learners from 2015 to 2016.

Table 5: Sample Demographics

Demographic Characteristic	2015 N = 3,775	2016 N = 3,832	2017 N = 3,886
<i>Gender</i>			
Female	48%	49%	48%
Male	52%	51%	52%
<i>SES Status: FRL</i>			
No	59%	55%	58%
Yes	41%	45%	42%
<i>Student with Disability</i>			
No	89%	90%	89%
Yes	11%	10%	11%
<i>English Language Learner</i>			
No	86%	75%	76%
Yes	14%	25%	24%
<i>Ethnicity</i>			
White	59%	57%	58%
Ethnic Minority			
Asian	6%	5%	5%
African American	2%	2%	2%
Hispanic	25%	26%	27%
Native American/Alaska Native	0.5%	0.6%	0.4%
Multi-Ethnic	6%	7%	6%
Pacific Islander	2%	2%	2%
<i>School Attended</i>			
EAST Charter School	3%	2.4%	2.25%
All Other FAIRMONT Schools	97%	97.6%	97.75%

Descriptive Statistics

The Oregon Department of Education SBAC ELA scores were scaled according to each grade level. The scaled scores were then categorized into four levels: Level 1 – limited to no mastery of knowledge and skills/ does not meet proficiency; Level 2 – inconsistent or partial mastery of knowledge and skills/does not meet proficiency; Level 3 – adept knowledge and skills/meet proficiency; Level 4 – exceptional knowledge and skills/exceeds requirements for proficiency. Table 6 provides the cut-scores for each categorized level of proficiency for the SBAC ELA for third through sixth grades. Achievement standards for the SBAC represent the point on the scale of scores that is the beginning of a level. For instance, a grade six ELA achievement standard of 2457 means that a student scoring 2457-2530 is in the level 2 range. A sixth-grade student scoring below the score of 2457 means the student is in the level 1 range.

Table 6: Achievement Standards Summary

Grade	Test	2	3	4
3	English Language Arts/Literacy	2367	2432	2490
4	English Language Arts/Literacy	2416	2473	2533
5	English Language Arts/Literacy	2442	2502	2582
6	English Language Arts/Literacy	2457	2531	2618

The mean score for the third-grade sample's SBAC ELA 2015 assessment was 2400.68 (± 344.70), for the 2016 assessment was 2422.34 (± 238.68), and for the 2017 assessment was 2406.15 (± 266.49). These mean scores indicate that for each annual assessment, on average, third-grade students achieved a Level 2 or did not meet proficiency standards due to inconsistent or partial mastery of English language arts knowledge and skills expected of third-grade students. The mean score for the fourth-grade sample's SBAC ELA 2015 assessment was 2466.69 (± 299.31), the 2016 assessment was 2461.55 (± 277.78), and the 2017 assessment was 2461.04 (± 234.11). These mean scores indicate that on average, fourth-grade students achieved

a Level 2 or did not meet proficiency standards due to inconsistent or partial mastery of English language arts knowledge and skills expected of fourth-grade students, for each annual assessment. The mean score for the fifth-grade sample's SBAC ELA 2015 assessment was 2483.24 (± 378.93), for the 2016 assessment was 2533.00 (± 193.43), and for the 2017 assessment was 2504.10 (± 273.80). The 2015 mean score indicates that on average, fifth-grade students achieved a Level 2 or did not meet proficiency standards due to inconsistent or partial mastery of English language arts knowledge and skills expected of fifth-grade students. The mean scores for 2016 and 2017 indicate that on average fifth-grade students achieved a Level 3 or met proficiency standards due to adeptness with English language arts knowledge and skills expected of fifth-grade students. The mean score for the sixth-grade sample's SBAC ELA 2015 assessment was 2487.84 (± 358.19), for the 2016 assessment was 2527.23 (± 290.52), and for the 2017 assessment was 2538.23 (± 273.80). These mean scores indicate that on average, sixth-grade students achieved a Level 3 or met proficiency standards due to adeptness with English language arts knowledge and skills expected of sixth-grade students, for each annual assessment. Table 7 is a summary of the descriptive statistics for participants SBAC ELA scaled scores.

Table 7: Descriptive Statistics for Smarter Balance Assessment English Language Arts Scaled Scores

<i>Year/Grade¹</i>	<i>M</i>	<i>SD</i>	<i>N</i>
2015			
Grade 3	2400.68	344.70	960
Grade 4	2466.69	299.31	934
Grade 5	2483.24	378.93	926
Grade 6	2487.84	358.19	955
2016			
Grade 3	2422.34	238.68	1008
Grade 4	2461.55	277.78	973
Grade 5	2533.00	193.43	927
Grade 6	2527.23	290.52	924
2017			
Grade 3	2406.15	266.49	1001
Grade 4	2461.04	234.11	990
Grade 5	2504.10	273.80	976
Grade 6	2538.23	222.96	919

Assumptions

The following assumptions were tested in order to conduct the data analyses:

independence of errors, no multicollinearity, and no significant outliers.

Independence of Errors. The independence of errors was inspected using the Durbin-Watson statistic. The first model included all independent variables, while model two excluded the student with disability variable. There was independence of errors as assessed by Durbin-Watson statistics for the models ranging from 1.53 to 1.82.

For 2015, the first model exhibited an independence of residuals, as assessed by a Durbin-Watson statistic of 1.82; the second model also exhibited an independence of residuals, as assessed by a Durbin-Watson statistic of 1.64. For 2016, the first model exhibited an independence of residuals, as assessed by a Durbin-Watson statistic of 1.74; the second model also exhibited an independence of residuals, as assessed by a Durbin-Watson statistic of 1.59.

¹ Achievement Standard Level 3 cut scores: Grade 3, 2432; Grade 4, 2473; Grade 5, 2502; Grade 6, 2531.

For 2017, the first model exhibited an independence of residuals, as assessed by a Durbin-Watson statistic of 1.77; the second model also exhibited an independence of residuals, as assessed by a Durbin-Watson statistic of 1.53. These statistics are within the normal ranges for Durbin-Watson as described by Laerd (2018), with typical ranges between 1.5-2.5. See Table 8 for a summary of the Durbin-Watson statistics.

Table 8: Summary of Durbin-Watson Statistics

Year/Model		Durbin-Watson
2015	All Independent Variables	1.82
	Without Disability Status Variable	1.64
2016	All Independent Variables	1.74
	Without Disability Status Variable	1.59
2017	All Independent Variables	1.77
	Without Disability Status Variable	1.53

Correlations between Dependent Variables and Independent Variables. Results of the Pearson correlation for 2015 indicate that there was a positive association between SBAC ELA achievement scores and gender ($r = .08$, $p < .001$.) Results of the Pearson correlation indicate that there was a negative association between SBAC ELA achievement scores and ethnicity ($r = -.08$, $p < .001$), SBAC ELA and socioeconomic status ($r = -.21$, $p < .001$), SBAC ELA and disability status ($r = -.49$, $p < .001$), SBAC ELA and English language learner status ($r = -.22$, $p < .001$), and SBAC ELA and curriculum taught ($r = -.04$, $p = .01$.) In other words, on the 2015 SBAC ELA there was a correlation between higher achievement scores and being female, as well as a correlation between lower achievement scores and being either of an ethnicity other than white, of a lower socioeconomic status, having a disability, or being an English language learner. There was no significant correlation between higher or lower SBAC ELA achievement scores and learning the Core Knowledge Sequence.

The product moment correlation coefficient for the independent disability status variable indicated there was a stronger negative linear relationship to the SBAC ELA achievement scores than the linear association for the other independent variables; the disability status variable was also associated with a higher percentage of case outliers than would be expected by chance. This greater number of outliers than could be anticipated by chance prompted a review of the data to determine the cause; in each year the majority of outlier achievement scores were associated with students with disabilities. Removing the student with disability status variable in a second model of data analyses reduced the percentage of case outliers to what could be expected by chance.

A second 2015 Pearson correlation analysis without the disability status variable resulted in stronger linear associations between the SBAC ELA achievement scores and other independent variables. Results of the Pearson correlation indicate that there was a positive association between SBAC ELA achievement scores and gender ($r = .13$, $p < .001$). Results of the Pearson correlation for the second model indicate that there was a negative association between SBAC ELA achievement scores and ethnicity ($r = -.25$, $p < .001$), SBAC ELA and socioeconomic status ($r = -.41$, $p < .001$), SBAC ELA and English language learner status ($r = -.44$, $p < .001$), and SBAC ELA and curriculum taught ($r = -.05$, $p < .001$). In other words, without the disability status variable there was a correlation between higher achievement scores and being female, as well as a correlation between lower achievement scores and being either of a non-white ethnicity, of a lower socioeconomic status, being an English language learner, or learning the Core Knowledge Sequence. See Table 9 for a summary of the variable correlations for 2015.

Table 9: Dependent and Independent Variable Correlations for 2015

	SBAC ELA	Gender	Ethnicity	SES	Disability	ELL	Curriculum
Model 1							
SBAC ELA	-						
Gender	0.08	-					
Ethnicity	-0.08	0.01	-				
SES	-0.21	-0.03	0.37	-			
Disability	-0.49	-0.08	0.00	0.13	-		
ELL	-0.22	-0.04	0.37	0.40	0.11	-	
Curriculum	-0.04	-0.03	0.05	0.14	0.02	0.06	-
Model 2							
SBAC ELA	-						
Gender	0.13	-					
Ethnicity	-0.25	0.00	-				
SES	-0.41	-0.01	0.38	-			
Disability					-		
ELL	-0.44	-0.03	0.36	0.40		-	
Curriculum	-0.05	-0.02	0.04	0.14		0.05	-

Results of the Pearson correlation for 2016 indicate that there was a positive association between SBAC ELA achievement scores and gender ($r = .07$, $p < .001$). Results of the Pearson correlation indicate that there was a negative association between SBAC ELA achievement scores and ethnicity ($r = -.10$, $p < .001$), SBAC ELA and socioeconomic status ($r = -.19$, $p < .001$), SBAC ELA and disability status ($r = -.37$, $p < .001$), SBAC ELA and English language learner status ($r = -.15$, $p < .001$). The results of the Pearson correlation indicate no significant linear relationship between SBAC ELA achievement scores and curriculum taught ($r = -.02$, $p = .11$). In other words, for the 2016 SBAC ELA there was a correlation between higher achievement scores and being female, as well as a correlation between lower achievement scores and being either of an ethnicity other than white, of a lower socioeconomic status, having a disability, or being an English language learner. There was no significant correlation between higher or lower achievement scores and learning the Core Knowledge Sequence.

A second 2016 Pearson correlation analysis without the disability status variable results in stronger linear associations between the SBAC ELA achievement scores and other

independent variables. Results of the Pearson correlation indicate that there was a positive association between SBAC ELA achievement scores and gender ($r = .09$, $p = < .001$). Results of the Pearson correlation for the second model indicate that there was a negative association between SBAC ELA achievement scores and ethnicity ($r = -.22$, $p = < .001$), SBAC ELA and socioeconomic status ($r = -.38$, $p = < .001$), and SBAC ELA and English language learner status ($r = -.33$, $p = < .001$). The result of the Pearson correlation for the second model indicated no change is the association between SBAC ELA achievement scores and curriculum taught ($r = -.02$, $p = .06$). In other words, without the disability status independent variable, there was a stronger correlation between higher achievement scores and being female, as well as a stronger correlation between lower achievement scores and either being of an ethnicity other than white, of a lower socioeconomic status, or an English language learner. There was no significant correlation between higher or lower SBAC ELA achievement scores and learning the Core Knowledge Sequence in 2016. See Table 10 for a summary of the variable correlations for 2016.

Table 10: Dependent and Independent Variable Correlations for 2016

	SBAC ELA	Gender	Ethnicity	SES	Disability	ELL	Curriculum
Model 1							
SBAC ELA	-						
Gender	0.07						
Ethnicity	-0.10	0.01					
SES	-0.19	-0.04	0.37				
Disability	-0.37	-0.08	0.01	0.10			
ELL	-0.15	-0.01	0.45	0.47	0.03		
Curriculum	-0.02	-0.01	0.04	0.13	0.03	0.07	
Model 2							
SBAC ELA	-						
Gender	0.09	-					
Ethnicity	-0.22	0.00	-				
SES	-0.38	-0.02	0.38	-			
Disability					-		
ELL	-0.33	-0.01	0.44	0.49		-	
Curriculum	-0.02	0.00	0.04	0.13		0.07	-

Results of the Pearson correlation for 2017 indicate that there was a positive association between SBAC ELA achievement scores and gender ($r = .07$, $p < .001$). Results of the Pearson correlation indicate that there was a negative association between SBAC ELA achievement scores and ethnicity ($r = -.10$, $p < .001$), SBAC ELA and socioeconomic status ($r = -.17$, $p < .001$), SBAC ELA and disability status ($r = -.39$, $p < .001$), and SBAC ELA and English language learner status ($r = -.14$, $p < .001$). The results of the Pearson correlation indicate no significant linear relationship between SBAC ELA achievement scores and curriculum taught ($r = -.002$, $p = .46$). In other words, on the 2017 SBAC ELA there was a correlation between higher achievement scores and being female, as well as a correlation between lower achievement scores and being either of an ethnicity other than white, of a lower socioeconomic status, having a disability, or being an English language learner. There was no significant correlation between higher or lower achievement scores and learning the Core Knowledge Sequence.

A second 2017 Pearson correlation analysis without the disability status variable results in stronger linear associations between the SBAC ELA achievement scores and other independent variables. Results of the Pearson correlation indicate that there was a positive association between SBAC ELA achievement scores and gender ($r = .09$, $p < .001$). Results of the Pearson correlation for the second model indicate that there was a negative association between SBAC ELA achievement scores and ethnicity ($r = -.22$, $p < .001$), SBAC ELA and socioeconomic status ($r = -.38$, $p < .001$), and SBAC ELA and English language learner status ($r = -.33$, $p < .001$). The results of the Pearson correlation indicate no significant linear relationship between SBAC ELA achievement scores and curriculum taught ($r = -.01$, $p = .33$). In other words, without the disability status independent variable, there was a stronger correlation between higher achievement scores and being female, as well as a stronger

correlation between lower achievement scores and either being of an ethnicity other than white, of a lower socioeconomic status, or an English language learner. There was no significant correlation between higher or lower SBAC ELA achievement scores and learning the Core Knowledge Sequence in 2017. See Table 11 for a summary of the variable correlations for 2017.

Table 11: Dependent and Independent Variable Correlations for 2017

	SBAC ELA	Gender	Ethnicity	SES	Disability	ELL	Curriculum
Model 1							
SBAC ELA	-						
Gender	0.07						
Ethnicity	-0.10	0.03					
SES	-0.17	-0.02	0.37				
Disability	-0.39	-0.12	0.01	0.10			
ELL	-0.14	0.01	0.46	0.51	0.05		
Curriculum	-0.002	0.02	0.05	0.12	-0.02	0.05	
Model 2							
SBAC ELA	-						
Gender	0.09	-					
Ethnicity	-0.22	0.03	-				
SES	-0.38	-0.01	0.38	-			
Disability					-		
ELL	-0.33	0.02	0.44	0.52		-	
Curriculum	-0.01	0.03	0.06	0.12		0.05	-

No multicollinearity. Collinearity was examined through correlation coefficients and Tolerance/VIF values, as well as the correlation matrix values for independent variables with the dependent variables. The first model included all independent variables, while model two excluded the student with disability variable. Tolerance measures greater than 0.1 and VIF statistics less than 10 satisfy the assumption that independent variables do not highly correlate. No tolerance measures were less than 0.1 and no VIF statistics were greater than 10, which satisfied the assumption of no multicollinearity for either model in 2015, 2016 or 2017. See Tables 12, 13, and 14 for summaries of the collinearity statistics.

Table 12: Collinearity Statistics Summary for 2015

					Collinearity	Statistics
	β	t	p	Partial Correlations	Tolerance	VIF
Model 1						
Gender	0.03	2.35	0.02	0.04	0.99	1.01
Ethnicity	0.00	0.08	0.94	0.00	0.80	1.26
SES	-0.09	-5.83	0.00	-0.09	0.76	1.31
Disability	-0.47	-33.14	0.00	-0.48	0.97	1.03
ELL	-0.13	-8.45	0.00	-0.14	0.78	1.29
Curriculum	-0.01	-0.50	0.62	-0.01	0.98	1.02
Model 2						
Gender	0.11	7.67	0.00	0.13	1.00	1.00
Ethnicity	-0.02	-1.53	0.13	-0.03	0.80	1.24
SES	-0.28	-16.41	0.00	-0.27	0.76	1.31
ELL	-0.32	-19.66	0.00	-0.32	0.79	1.26
Curriculum	0.01	0.89	0.37	0.02	0.98	1.02

Table 13: Collinearity Statistics Summary for 2016

					Collinearity	Statistics
	β	t	p	Partial Correlations	Tolerance	VIF
Model 1						
Gender	0.04	2.39	0.02	0.04	0.99	1.01
Ethnicity	-0.02	-1.12	0.26	-0.02	0.76	1.31
SES	-0.11	-6.27	0.00	-0.10	0.73	1.37
Disability	-0.36	-23.92	0.00	-0.36	0.98	1.02
ELL	-0.08	-4.74	0.00	-0.08	0.69	1.44
Curriculum	0.01	0.53	0.60	0.01	0.98	1.02
Model 2						
Gender	0.08	5.44	0.00	0.09	1.00	1.00
Ethnicity	-0.04	-2.30	0.02	-0.04	0.77	1.30
SES	-0.28	-15.28	0.00	-0.25	0.72	1.38
ELL	-0.18	-9.61	0.00	-0.16	0.69	1.46
Curriculum	0.03	1.91	0.06	0.03	0.98	1.02

Table 14: Collinearity Statistics Summary for 2017

					Collinearity	Statistics
	β	t	p	Partial Correlations	Tolerance	VIF
Model 1						
Gender	0.03	1.94	0.05	0.03	0.98	1.02
Ethnicity	-0.04	-2.54	0.01	-0.04	0.76	1.31
SES	-0.09	-4.88	0.00	-0.08	0.70	1.44
Disability	-0.37	-25.17	0.00	-0.37	0.98	1.03
ELL	-0.06	-3.20	0.00	-0.05	0.65	1.53
Curriculum	0.01	0.48	0.63	0.01	0.98	1.02
Model 2						
Gender	0.09	5.97	0.00	0.10	1.00	1.00
Ethnicity	-0.05	-2.84	0.00	-0.05	0.77	1.29
SES	-0.28	-15.40	0.00	-0.25	0.69	1.44
ELL	-0.16	-8.66	0.00	-0.15	0.66	1.53
Curriculum	0.04	2.27	0.02	0.04	0.99	1.01

No significant outliers. Outliers were detected using casewise diagnostics. In 2015 there were 88 cases identified out of 3,772 cases, a percentage of 2.34 which is higher than would be expected by chance. In 2016 there were 39 cases identified out of 3,827, a percentage of 1.01 which is higher than would be expected by chance. In 2017 there were 41 cases identified out of 3,883, a percentage of 1.05 which is higher than would be expected by chance. A greater number of outliers than could be anticipated by chance prompted a review of the data to determine the cause; in each year the majority of outlier achievement scores were associated with students with disabilities. Removing this variable from the analyses, results in a range of outliers from 0.14% to 0.23% which is below the approximate proportion of cases expected by chance. See Table 15 for a summary of the case outliers.

Table 15: Case Outliers

		N	Cases	Percentage
2015	Model 1	3,772	88	2.34%
	Model 2	3,365	8	0.23%
2016	Model 1	3,827	39	1.01%
	Model 2	3,430	5	0.14%
2017	Model 1	3,883	41	1.05%
	Model 2	3,483	7	0.20%

Research Question

To what extent does the Core Knowledge Sequence curriculum impact SBAC ELA achievement scores for third through sixth-grade students?

2015 Data. Multiple regression analyses were run to determine whether the linear regression between the SBAC ELA achievement scores and the independent variables (gender, ethnicity, English language learner, socioeconomic status, disability status, and curriculum) was statistically significant ($p < .05$). The MR analyses were run as well to determine how much of the variation in the SBAC ELA score could be explained by the curriculum. The 2015 Pearson multiple correlation coefficient, or the R value, for the first model between the SBAC ELA score and all independent variables was 0.53 which was a moderate, positive correlation. The result $F(6, 3765) = 244.79$, $p < .001$ from the ANOVA test indicates that this correlation is statistically significant ($p < .05$). However, the adjusted r^2 value for this model was 0.28; this is a modest adjusted r^2 value indicating that 28% of the variability in SBAC ELA outcomes in the population could be accounted for by the independent variables. The 2015 Pearson multiple correlation coefficient, or the R value, for the second model between the SBAC ELA score and all independent variables, except the disability status variable, was 0.53; a moderate, positive correlation. The result $F(5, 3359) = 257.97$, $p < .001$ from the ANOVA test indicates that this

correlation is statistically significant ($p < .05$). However, the adjusted r^2 for this second model was 0.28, a modest effect size according to Cohen (1988). Table 16 summarizes the multiple regression and ANOVA analyses for 2015.

Table 16: Multiple Regression Analyses, 2015

2015	<i>R</i>	r^2	Adjusted r^2	<i>SE</i> of the Estimate
Model 1	0.53	0.28	0.28	304.15
Model 2	0.53	0.28	0.28	78.27

Independent Variables to SBAC ELA Achievement Score ANOVA, 2015

2015					
Model 1	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>p</i>
Regression	135867017.74	6	22644502.96	244.79	.00
Residual	348288004.53	3765	92506.77		
Total	484155022.26	3771			
Model 2	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>p</i>
Regression	7901295.81	5	1580259.16	257.97	.00
Residual	20576542.21	3359	6125.79		
Total	28477838.02	3364			

All of the variables, except ethnicity and curriculum, were found to have statistically significant ($p < .05$) unstandardized coefficients for 2015. The independent variable disability status was found to have a statistically significant unstandardized coefficient ($\beta = -537.02$, $p = < .001$) showing that when classified as a student with a disability, participants' SBAC ELA scores were on average 537.02 points lower than students classified with no disability. The size of the unstandardized coefficient for the disability status variable prompted a review of the data; the casewise diagnostics determined there were a greater number of outliers than could be anticipated by chance and the majority of these outlier achievement scores were associated with the disability status variable.

A second analysis for the 2015 SBAC ELA scores without the independent disability status variable determined all variables, except ethnicity and curriculum, were found to have

statistically significant ($p < .05$) unstandardized coefficients for 2015. Table 17 summarizes the 2015 unstandardized coefficients for independent variables with and without the disability status variable.

Table 17: Unstandardized Coefficients for Independent Variables with the Dependent Variable 2015 SBAC ELA Score, With and Without the Independent Disability Status Variable

Model 1	Unstandardized Coefficients		Standardized Coefficient	<i>t</i>	<i>p</i>
	β	<i>SE</i>	β		
SBAC ELA	2567.15	29.30		87.61	0.00
Gender	23.38	9.95	0.03	2.35	0.02
Ethnicity	0.26	3.30	0.00	0.08	0.94
SES Status	-67.27	11.54	-0.09	-5.83	0.00
Disability Status	-537.02	16.20	-0.47	-33.14	0.00
ELL Status	-135.44	16.02	-0.13	-8.45	0.00
Curriculum	-14.77	29.37	-0.01	-0.50	0.62

Model 2	Unstandardized Coefficients		Standardized Coefficient	<i>t</i>	<i>p</i>
	β	<i>SE</i>	β		
SBAC ELA	2537.34	7.82		324.45	0.00
Gender	20.71	2.70	0.11	7.67	0.00
Ethnicity	-1.36	0.89	-0.02	-1.53	0.13
SES Status	-51.94	3.17	-0.28	-16.41	0.00
ELL Status	-88.50	4.50	-0.32	-19.66	0.00
Curriculum	7.00	7.84	0.01	0.89	0.37

2016 Data. The 2016 Pearson multiple correlation coefficient, or the R value, for the first model between the SBAC ELA score and all independent variables was 0.41 which was a moderate, positive correlation. The result $F(6, 3820) = 130.81$, $p < .001$ from the ANOVA test indicates that this multiple correlation is statistically significant ($p < .05$.) However, the adjusted r^2 value for this model was 0.17; this is a modest adjusted r^2 value indicating that 17% of the variability in SBAC ELA outcomes in the population could be accounted for by the independent variables. The 2016 R value for the second model between the SBAC ELA score and all independent variables except the disability status variable was 0.43 which was a moderate, positive correlation. The result $F(5, 3424) = 151.00$, $p < .001$ from the ANOVA test indicates

that this correlation is statistically significant ($p < .05$). However, the adjusted r^2 for this second model was 0.18, a modest effect size for the population. Table 18 summarizes the multiple regression and ANOVA analyses for 2016.

Table 18: Multiple Regression Analyses, 2016

	<i>R</i>	r^2	Adjusted r^2	<i>SE</i> of the Estimate
Model 1	0.41	0.17	0.17	229.57
Model 2	0.43	0.18	0.18	94.49

Independent Variables to SBAC ELA Achievement Score ANOVA, 2016

Model 1	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>p</i>
Regression	41365608.46	6	6894268.08	130.81	.00
Residual	201329849.68	3820	52704.15		
Total	242695458.14	3826			
Model 2	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>p</i>
Regression	6741040.28	5	1348208.06	151.00	.00
Residual	30571099.43	3424	8928.48		
Total	37312139.71	3429			

The data analysis for 2016 SBAC ELA scores determined the unstandardized coefficients for the independent variables including gender, ethnicity, socio-economic status, disability status, English language learner status, and curriculum. All of the variables, except ethnicity and curriculum, were found to have statistically significant ($p < .05$) unstandardized coefficients for 2016. The independent variable disability status was found to have a statistically significant unstandardized coefficient ($\beta = -293.90$, $p = < .001$) showing that when classified as a student with a disability, participants' SBAC ELA scores were on average 293.90 points lower than if they were classified with no disability.

A second analysis for the 2016 SBAC ELA scores without the independent disability status variable determined all variables, except curriculum, were found to have statistically significant ($p < .05$) unstandardized coefficients. Table 19 summarizes the 2016

unstandardized coefficients for independent variables with and without the disability status variable.

Table 19: Unstandardized Coefficients for Independent Variables with the Dependent Variable 2016 SBAC ELA Score, with and without the Independent Disability Status Variable

All Independent Variables	Unstandardized	Coefficients	Standardized Coefficient		
	β	SE	β	t	p
SBAC ELA	2535.05	24.344		104.13	0.00
Gender	17.86	7.46	0.04	2.39	0.02
Ethnicity	-2.79	2.48	-0.02	-1.12	0.26
SES Status	-54.87	8.75	-0.11	-6.27	0.00
Disability Status	-293.90	12.29	-0.36	-23.92	0.00
ELL Status	-49.05	10.35	-0.08	-4.74	0.00
Curriculum	12.89	24.44	0.01	0.53	0.60

No Disability Status Variable	Unstandardized	Coefficients	Standardized Coefficient		
	β	SE	β	t	p
SBAC ELA	2528.44	10.24		246.94	0.00
Gender	17.57	3.23	0.08	5.44	0.00
Ethnicity	-2.47	1.07	-0.04	-2.30	0.02
SES Status	-58.57	3.83	-0.28	-15.28	0.00
ELL Status	-43.66	4.54	-0.18	-9.61	0.00
Curriculum	19.68	10.29	0.03	1.91	0.06

2017 Data. The 2017 Pearson multiple correlation coefficient, or the R value, for the first model between the SBAC ELA score and all independent variables was 0.42 which was a moderate, positive correlation. The result $F(6, 3876) = 135.12$, $p < .001$ from the ANOVA test indicates that this correlation is statistically significant ($p < .05$). However, the adjusted r^2 value for this model was 0.17; this is a modest adjusted r^2 value indicating that 17% of the variability in SBAC ELA outcomes in the population could be accounted for by the independent variables. The 2017 R-value for the second model between the SBAC ELA score and all independent variables except the disability status variable was 0.43 which was a moderate, positive correlation. The result $F(5, 3477) = 154.98$, $p < .001$ from the ANOVA test indicates that this correlation is statistically significant ($p < .05$). However, the adjusted r^2 for this second model

was 0.18, a modest effect size for the population. Table 20 summarizes the multiple regression and ANOVA analyses for 2017.

Table 20: Multiple Regression Analyses, 2017

	<i>R</i>	<i>r</i> ²	Adjusted <i>r</i> ²	<i>SE</i> of the Estimate
Model 1	0.42	0.17	0.17	232.51
Model 2	0.43	0.18	0.18	88.59

Independent Variables to SBAC ELA Achievement Score ANOVA, 2017

Model 1	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>p</i>
Regression	43828287.30	6	7304714.55	135.12	.00
Residual	209538991.35	3876	54060.63		
Total	253367278.65	3882			
Model 2	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>p</i>
Regression	6080764.34	5	1216152.87	154.98	.00
Residual	27285142.59	3477	7847.32		
Total	33365906.94	3482			

The data analysis for 2017 SBAC ELA scores determined the unstandardized coefficients for the independent variables including gender, ethnicity, socio-economic status, disability status, English language learner status, and curriculum. All of the variables, except curriculum, were found to have statistically significant ($p < .05$) unstandardized coefficients for 2017.

A second analysis for the 2017 SBAC ELA scores without the independent disability status variable determined all variables were found to have statistically significant ($p < .05$) unstandardized coefficients. The independent variable curriculum was found to have a statistically significant unstandardized coefficient ($\beta = 23.63$, $p = < .001$) showing that when taught the Core Knowledge Sequence curriculum, participants' SBAC ELA scores were on average 23.63 points higher than when taught the Journeys curriculum. Table 21 summarizes

the 2017 unstandardized coefficients for independent variables with and without the disability status variable.

Table 21: Unstandardized Coefficients for Independent Variables with the Dependent Variable 2017 SBAC ELA Score, with and without the Independent Disability Status Variable

All Independent Variables	Unstandardized	Coefficients	Standardized Coefficient		
	β	SE	β	t	p
SBAC ELA	2524.20	25.27		99.89	0.00
Gender	14.64	7.53	0.03	1.94	0.05
Ethnicity	-6.40	2.52	-0.04	-2.54	0.01
SES Status	-44.29	9.08	-0.09	-4.88	0.00
Disability Status	-312.87	12.43	-0.37	-25.17	0.00
ELL Status	-34.48	10.77	-0.06	-3.20	0.00
Curriculum	12.33	25.43	0.01	0.48	0.63

No Disability Status Variable	Unstandardized	Coefficients	Standardized Coefficient		
	β	SE	β	t	p
SBAC ELA	2512.58	10.32		243.47	0.00
Gender	17.94	3.01	0.09	5.97	0.00
Ethnicity	-2.86	1.00	-0.05	-2.84	0.00
SES Status	-56.69	3.68	-0.28	-15.40	0.00
ELL Status	-37.78	4.36	-0.16	-8.66	0.00
Curriculum	23.63	10.42	0.04	2.27	0.02

Summary

The following chapter discusses the study findings indicating demographic variables offer more predictive information on SBAC ELA scores than curriculum's predictive information. It also offers a broader perspective of education than the prevailing idea that education is a commodity.

CHAPTER 5

Conclusion

The path towards this study began two decades ago when I home schooled my young daughters using the Core Knowledge Sequence as one of our curricular resources. Teaching my daughters to read and write well by giving them access to a wealth of knowledge about the world was one of the more fulfilling seasons of my life. Now, as the Executive Director for EAST Charter School I am responsible and accountable for the quality of my teachers' instruction and my students' learning, and do so, again, by giving them access to teach and to learn using the same rich knowledge-based curriculum I used in teaching my children. The journey towards this study was by no means linear, but in retrospect the track never veered far from reaching a point which provided the privilege to dig deeper into the CKS.

The aim of this study was to examine the impact of two curricula, the Core Knowledge Sequence and Journeys, on student achievement in English language arts as measured by the Smarter Balance Assessment (SBAC). Identifying the rationale used in selecting curriculum highlighted the differences in the philosophy and fundamental purposes for not only the curricula, but for EAST Charter School and the Fairmont School District, as well. The literature reviewed for the Core Knowledge Sequence and Journeys afforded the discovery of a commonality between the curricula: current, peer-reviewed research is needed to more accurately understand each curriculum's effectiveness in meeting publisher-stated learning outcomes. Finally, considering the literature examining standardized assessment as an accountability tool opened the door to a number of challenges faced by practitioners regardless of the where they teach. This chapter offers a discussion of the findings, study limitations, implications for practitioners, and recommendations for further research.

Discussion

The following section discusses the study findings related to the research question. The research question was: Is there a difference in the Oregon Smarter Balance English language arts test scores by curriculum taught?

There was no statistically significant difference in the Oregon SBAC ELA student achievement scores between students taught the Core Knowledge Sequence and students taught the Houghton Mifflin Harcourt Journeys curricula in the 2015, 2016, or 2017. In 2015, the first model indicated a weak, negative correlation ($r = -0.04$) between CKS and SBAC ELA scores which was not statistically significant ($p < .001$). The second model, the analyses conducted without the disability status variable, indicated a weak, negative correlation ($r = -0.005$) between the CKS and the SBAC ELA scores which was not statistically significant ($p > .001$). In 2016, the first and second models indicated weak, negative correlations ($r = -0.02$) between CKS and SBAC ELA scores which were not statistically significant ($p < .001$). In 2017, there was no correlation between the CKS and the SBAC ELA scores ($r = 0.00$) in the first model. The second model indicated a weak, negative correlation ($r = -0.01$) between CKS and SBAC ELA scores which was not statistically significant ($p > .001$).

The independent variable curriculum provided little difference in predictive information on SBAC ELA achievement scores which mirrors prior research comparing student achievement scores based on curricula. The longitudinal study conducted by Datnow, Borman, and Stringfield (2000) comparing the effects of the implementation of the Core Knowledge Sequence in four different schools on norm-referenced standardized tests found students' basic skills in reading achievement scores were consistent with the achievement scores of students in schools not taught the CKS curriculum. Similarly, the Sonnenschein, Baker, and Garret (2005) study of pre-school

children taught the CKS Pre-School Sequence found achievement scores on the Woodcock Johnson Tests of Achievement – III indicated growth in oral language development was comparable to that of the normed group taught a different pre-school curriculum. Finally, the 2013 PRES report examining the impact of Journeys on student achievement on the Iowa Test of Basic Skills (ITBS) indicated statistically significant differences with a small effect size in favor of the Journeys’ curricula in two of the five ITBS language arts measures in the second year of the study. Overall, though, there was no statistically significant difference in student achievement on the ITBS for students taught Journeys in comparison to the six other curricula evaluated.

A number of inferences could also be made about the curriculum variable’s predictive capacity regarding student achievement by reviewing the SBAC ELA descriptive statistics in the context of previous research examining state standards, assessments, curriculum and pedagogy. More than 96% of the sample in this study received instruction in Houghton Mifflin Harcourts Journeys in 2015, 2016, and 2017. On average, third and fourth grade students in the Fairmont School District scored a level 2 or lower on the SBAC ELA in 2015, 2016, and 2017. On average fifth grade students scored a level 2 or lower on the SBAC ELA in 2015, while sixth grade students, on average, scored a level 2 or lower on the SBAC ELA in 2015 and 2016. In these descriptive statistics, the three data points indicating students on average scored a level 3, signifying they met the Common Core State Standards, occurred in the fifth and sixth grades in the second and third years of the assessment. Given that 2015 was the first year the SBAC ELA was administered statewide, it could be concluded that there was a misalignment between the Journeys curriculum and the SBAC ELA test items, a delay between alignment of instructional practices to the SBAC ELA assessment requirements, or both. Again, these findings mirror prior

research findings. Polikoff, Porter, and Smithson (2011) found a misalignment in topics and in cognitive demands between standardized assessments and curriculum in data from nineteen states while Au (2007) and Diamond (2007) found that over time instructional practices changed to teach material in direct relation to the standardized tests.

In this study, demographic variables (gender, disability status, English language learner status, socio-economic status, and ethnicity) were also examined in relation to SBAC ELA scores. These demographic variables had a stronger predictive power on the SABC ELA scores than the curriculum variable. In 2015, 2016 and 2107 the strongest predictor of SBAC ELA scores was disability status; there were strong, negative correlations in 2015 ($r = -0.49$), in 2016 ($r = -0.37$), and in 2017 ($r = -0.39$) which were each statistically significant ($p < .001$). But it was found that the disability status variable was associated with a greater number of outlier achievement scores than could be anticipated by chance.

When a second model of data analyses was conducted excluding the disability status variable, the socio-economic status (SES) variable became the strongest predictor of SBAC ELA scores. In this study SES was operationalized as qualifying for free or reduced lunch under the National School Lunch Program using Federal income poverty guidelines. There were strong, negative correlations in 2015 ($r = -51.94$), in 2016 ($r = -58.57$), and in 2017 ($r = -56.69$) between SES and SBAC ELA scores which were each statistically significant ($p < .001$). In other words, when identified as meeting the Federal poverty guidelines, participants' SBAC ELA scores were on average more than 50 points lower than participants not identified with a low SES. This finding is consistent with the Nichols, Glass, and Berliner (2012) findings of a negative correlation between standardized assessments used as school accountability tools and lower reading achievement scores for lower income students.

These findings raise concerns about the capacity of standardized assessments to measure student learning. The findings in this study indicate the curriculum has no predictive power on SBAC ELA student achievement scores: if what is taught is not measured by the assessment, then what is measured? In the case of this study, the findings indicate demographic variables have a far greater predictive power than curriculum on SBAC ELA student achievement scores. Is the assessment measuring the student rather than what the student has learned? Hopefully, this is not the case. But the juxtaposition of these two findings seems to muddle Polikoff's (2012) description of the fundamental purpose of standardized assessments as measuring specific, well-defined student learning outcomes which are mapped in curriculum frameworks with specific objectives for teachers.

Similar to other valid and reliable standardized assessments, the SBAC ELA tests students by drawing upon a pre-established bank of questions and the answers are scored so that it is possible to compare the relative performance of a single student or a group of students. Specific SBAC ELA test items are intended to measure student progress towards college and career readiness as outlined in the Common Core State Standards. In this paradigm, the capacity of the standardized assessment to meaningfully measure student learning is limited to how well students have learned the English language arts skills determined to be appropriate to their grade level. This pattern is true for standardized assessments in general: only what has been pre-defined as valuable knowledge or learning is assessed. Therefore, a single standardized assessment is limited in its capacity to measure student knowledge or learning which does not fall within its pre-defined scope of material.

Limitations

As discussed in Chapter 2, examining the impact of two curricula would suggest a review of similarities and differences between the two curricula, but collecting and analyzing the qualitative data necessary to accomplish this task was of greater magnitude than resources allowed for this study. Also, the Smarter Balance Assessment was implemented across Oregon in 2015, thus providing access to only three years of student achievement assessment data.

Two additional limitations were discovered while working with the data: curriculum path and sample size. First, there was an ineffective operationalization of the curriculum path concept. Curriculum path was operationalized to provide for the researcher the number of years a student received instruction in either the Core Knowledge Sequence or Journeys, but the data set obtained from the Oregon Department of Education gave access to school attendance for the year of the assessment. Thus, it was not possible to determine how many years an individual student received instruction in one curriculum or the other, and the data analyses accounted for curriculum on a single-year basis. Next, sample size demographics became a limitation to this study in that only 2.25% to 3% of the students in the sample received instruction in the Core Knowledge Sequence compared to 97 – 98% of the sample receiving instruction in Journeys.

Implications

Caution is wisdom when interpreting statistical analyses and generalizing the findings from any study. While it would be wildly irresponsible to use this study's findings to stand on the top of a hill and yell "Something is not working!" there is cause for educators and policy makers to take stock of concerns raised by the study.

In this study, if the Core Knowledge Sequence curriculum had no predictive power on SBAC ELA student achievement scores, then student learning through this curriculum was not

meaningfully measured by this assessment. The SBAC ELA student achievement scores do not reflect the wide knowledge of literature EAST Charter School elementary students learn in third through sixth grades. Furthermore, the SBAC does not provide any opportunity for an assessment of students' knowledge of history, geography, or the arts. Students at EAST Charter School learn about the Bayeux Tapestry and the Norman Conquest of England in the fourth grade CKS, but, currently, there is no meaningful way for the school's sponsor or the state to include this type of learning and knowledge in charter accountability measures. The more than 100 charter schools across the state of Oregon hold contracts which specify the use of curricula which most likely is not aligned to the SBAC assessment items; if curriculum has little predictive power on SBAC student achievement scores, then the balance between charter autonomy and accountability is not weighted fairly for students who attend charter schools.

Revisiting Oregon's charter school law and the specific legislative intentions and goals under ORS 338.015 includes acknowledging the unfulfilled legal mandate to “(8) establish additional forms of accountability for schools; and (9) create innovative measurement tools. [1999 c200 §1]” (ORS 338). This is a task which must be addressed collaboratively not only between the state's charter schools, but with sponsoring districts as well as the Oregon Department of Education. Charter schools exist to provide an alternative to traditional public schools; determining charter school accountability and quality of education with just the same tool used for traditional schools and districts obfuscates the purpose of charter schools' existence. Charter schools are not mini-public schools – they are alternative educational programs created as a choice for all parents. This is not an easy or simple task, but one that would require a statewide commitment of resources to create meaningful assessments useful for the charter sponsors, the charter schools, ODE, legislative policy makers, and ultimately, parents

interested in understanding the quality of the alternative education provided by a specific charter school.

Two avenues of additional assessment could be examined to create additional meaningful accountability measures for charter schools: value-added measures and individual charter measurements aligned to the curriculum in use. Value-added measures could include relevant data such as charter school students' high school graduation rates and graduation on-time data, completion of Advanced Placement or International Baccalaureate courses, over-all high school GPA, SAT or ACT scores, and college admissions. It makes sense to also create additional accountability measures tailored for the differences inherent in charter contracts. While an additional assessment and accountability tool for EAST Charter School would necessitate addressing the rich knowledge base within the curriculum, a charter school employing the Expeditionary Learning curriculum would require a completely different assessment tool to capture the students' project-based, real-world learning outcomes. A single, standardized assessment may make it easier to compare the student learning between traditional and charter schools, it does not provide an accurate or complete reflection of the learning opportunities available to students through a charter school education. Other accountability tools are necessary to balance the scale of charter school autonomy and accountability.

A second concern raised by the findings in this study goes to the heart of closing the achievement gap for underserved student populations. The juxtaposition of two findings mentioned in the discussion section, i.e., the curriculum variable holding no predictive information and demographic variables holding strong predictive information for SBAC ELA student achievement scores, should signal an alert for K-12 educators and policy makers in the Fairmont School District. If at the end of three assessment cycles students in poverty consistently

score on average 50 points lower than other students, and third and fourth grade students on average do not reach benchmarks in any assessment cycle, something is amiss. This study's findings that demographic variables have a greater predictive impact than that of the curriculum taught echoes an alarm sounded by a 2017 research review by Johns Hopkins Institute for Education Policy and Johns Hopkins Center for Research and Reform in Education. Curriculum choice, especially content-rich curriculum, is identified as the critical factor in student academic success, but

because most state standards, including the Common Core, and most state assessments, including PARCC and SmarterBalanced (sic), are largely skills focused, many curricular materials in the United States, especially in ELA, focus on skills rather than knowledge. This is unsurprising, given the fact that it has been notoriously difficult to agree upon which key text students should read or which areas of knowledge they should master (Steiner, 2018, p.8).

The Journeys curriculum taught in the Fairmont School District focuses on English language arts skill development, and therefore underserved student populations, such as students living in poverty, continue at a learning disadvantage as long as they are not also taught content knowledge.

Recommendations for Further Research

An impetus for this study grew from concerns about the use of a single measurement tool (SBAC) for holding accountable schools who are teaching very different curricula; how is the accountability process fair or effective if there is a misalignment between the curriculum and the assessment tool? Through the process of preparing this study, it has become clear that a

significant dichotomy of educational thought exists not only between EASY Charter School and Fairmount School District, but across public education.

Prevailing school accountability rests on the idea that education is a commodity which can be valued by its product: students who are prepared for success in college and career to ensure that future national economic growth can be sustained. If we assume accurate the premise for educational reform is a skills-driven global economy, as suggested by the development of the Common Core State Standards (NGA, CCSO & Achieve, 2008), then research is required to identify what makes a skills-based curriculum effective in student learning. If we assume that education is more than a commodity, but also a national resource, and that the premise for educational reform is not only an economic imperative, but also on a socio-political imperative for a widely-shared body of common knowledge to protect citizens and promote democratic principles (Bagley, 1934; Koch & Peden, 1944; Hirsch, 1998), then research is required to identify what makes a knowledge-based curriculum effective in student learning.

Conclusion

The Core Knowledge Sequence is a knowledge-based curriculum in a world where educational success is currently determined by skills acquisitions. This research project has brought home to me the significance of this mismatch. I can tell you (and even show you in the classroom) what and how my students are learning through the Core Knowledge Sequence but am unable to provide statistical data to support what I know. Which leads me full circle to the original statement of the problem in Chapter 1: How does EAST Charter School fulfill its contractual agreement that our students would score better on state standardized tests than students taught the district-selected curriculum? Beyond simply a mismatch between curriculum

and assessment tools, there is a profound difference in philosophies of education between the charter school and district.

In the process of preparing this dissertation, EAST Charter School has completed most of a charter renewal process with the Fairmont School District with school evaluation based on the previous contractual expectations. The charter has been renewed, but bridging the philosophical differences is critical to the ongoing welfare of the school and our students as a new charter contract is negotiated.

References

- Au, W. (2007). High-stakes testing and curricular control: A qualitative metasynthesis. *Educational Researcher*, 36(5), 258-267.
- Bagley, W.C. (1934). Education and the emergent man: A theory of education with particular application to public education in the United States. New York: Nelson.
- Ballou, D., & Springer, M.G. (2008). *Achievement trade-offs and No Child Left Behind*. Working paper. Peabody College of Vanderbilt University.
- Borman, G. D., Hewes, G. M., Overman, L. T., & Brown, S. (2003). Comprehensive school reform and achievement: A meta-analysis. *Review of Educational Research*, 73(2), 125-130. Retrieved from <http://www.jstor.org/stable/10.2307/3516091?ref=no>
- Braker-Walters, B. A. (2014). Informational text and the Common Core: A content analysis of three basal reading programs. *SAGE-Open*, October-December 2014: 1-8. DOI: 10.1177/2158244014555119
- Carnoy, M., Elmore, R., & Siskin, L. S. (Eds.) (2003). *The new accountability: High schools and high-stakes testing*. New York, NY: Taylor and Francis Books, Inc.
- Carnoy, M., & Loeb, S. (2002). Does external accountability affect student outcomes? A cross state analysis. *Educational Evaluation and Policy Analysis*, 24(4), 305-331. Retrieved from <http://www.jstor.org.georgefox.idm.oclc.org/stable/3594120>
- Common Core State Standards Initiative. (2016). *Key shifts in English language arts*. Retrieved from <http://www.corestandards.org/other-resources/key-shifts-in-english-language-arts/>
- Core Knowledge Foundation. (2010). *E. D. Hirsch, Jr.* Retrieved from Core Knowledge Foundation: <http://www.coreknowledge.org/ed-hirsch-jr>
- Core Knowledge Foundation. (2010). *Introduction to the 2010 Sequence*. Retrieved from Core

Knowledge Foundation: <http://www.coreknowledge.org/introduction-to-the-2010-sequence>

Core Knowledge Foundation. (2015). *The Core Knowledge Sequence: Content and skill guidelines for grades K-8*. Charlottesville, VA: Core Knowledge Foundation. Retrieved from Core Knowledge Foundation: <http://www.coreknowledge.org>

Cubberly, E.P. (1919). *Public education in the United States: A study and interpretation of American educational history*. Cambridge, MA: Riverside Press.

Datnow, A., Borman, G., and Stringfield, S. (2000). School reform through a highly specified curriculum: Implementation and effects of the Core Knowledge Sequence. *The Elementary School Journal*, 101(2), 167.

Datnow, A., McHugh, B., Stringfield, S., & Hacker, D. J. (1998). Scaling up the Core Knowledge Sequence. *Education and Urban Society*, 30(3), 409-432.

Diamond, J.B. (2007). Where the rubber meets the road: Rethinking the connection between high-stakes testing policy and classroom instruction. *Sociology of Education*, 80, 285-313. Retrieved from <http://www.jstor.org.georgefox.idm.oclc.org/stable/20452714>

Drits-Esser, D., Bass, K. M., & Stark, L. A. (2014). Using small-scale randomized controlled trials to evaluate the efficacy of new curricular materials. *CBE – Life Sciences Education*, 13(4), 593-601. DOI: 10.1187/cbe.13-08-0164

Educational Research Institute of America. (2009). *A study of the instructional effectiveness of Houghton Mifflin Harcourt Journeys: Unit 6. Report number 366*. Retrieved from <https://prod-hmhco-vmg-craftcms-private.s3.amazonaws.com/documents/Instructional Effectiveness>

Educational Research Institute of America, (2009). *A study of the instructional effectiveness of*

- Houghton Mifflin Harcourt Journeys. Report number 368A.* Retrieved from <https://prod-hmhco-vmg-craftcms-private.s3.amazonaws.com/documents>
- Farace, D. J. & Schopfel, J. (2010). *Grey literature in library and information studies*. Berlin: De Gruyter.
- Feltovitch, P. J., Prietula, M. J., & Ericsson, K. A. (2006). Studies of expertise from psychological perspectives. In K. A. Ericsson, N. Charness, P. J. Feltovich, & R. R. Hoffman (Eds.), *The Cambridge handbook of expertise and performance*, 41–67. Cambridge: Cambridge University Press.
- Gandal, M. (1997). *Making standards matter, 1997: An annual fifty-state report on efforts to raise academic standards*. AFT Educational Issues Department. Retrieved from <http://files.eric.ed.gov/fulltext/ED410661.pdf>
- Hakuta, K. (2000). *Education 388 Foundations of Academic Language: What is standards based reform?* Stanford University School of Education. Retrieved from https://web.stanford.edu/~hakuta/www/archives/syllabi/CalTex_SBR/whatissbr.html
- Hanushek, E. A., & Raymond, M. E. (2005). Does school accountability lead to improved student performance? *Journal of Policy Analysis and Management*, 24(2), 297-327. Retrieved from <http://www.jstor.org.georgefox.idm.oclc.org/stable/3326211>
- Hess, K. (2013). A guide for using Webb's depth of knowledge with common core state standards. *The Common Core Institute, Center for College and Career Readiness*. Retrieved from http://www.crecnow.info/blendedsolutions/docs/docs/Webbs_Depth_of_Knowledge.pdf
- Hirsch, E. D. (1987). *Cultural literacy: What every American needs to know*. New York: Random House.

Hirsch, E. D. (1998). *The schools we need and why we don't have them*. New York: Random House.

Hirsch, E. D. (2013). *Core Knowledge Foundation*. Retrieved from Letter from E.D. Hirsch, Jr.: <http://www.coreknowledge.org/ed-hirsch-jr>

Hirsch, E. D. (2016). *Why knowledge matters: Rescuing our children from failed educational theories*. Cambridge, MA: Harvard Education Press.

Houghton Mifflin Harcourt. (2011). *Journeys: Apollo-Ridge School District, Pennsylvania*. Case study.

Houghton Mifflin Harcourt. (2011). *Journeys: Barberton City School District, Ohio*. Case study.

Houghton Mifflin Harcourt. (2011). *Journeys: Benjaming Logan Local School District, Ohio*. Case study.

Houghton Mifflin Harcourt. (2011). *Journeys: Canton School District, Massachussettes*. Case study.

Houghton Mifflin Harcourt. (2011). *Journeys: Carthage R-9 School District, Missouri*. Case study.

Houghton Mifflin Harcourt. (2011). *Journeys: City of Chicago School District 299, Illinois*. Case study.

Houghton Mifflin Harcourt. (2011). *Journeys: Diamond Lake School District 76, Illinois*. Case study.

Houghton Mifflin Harcourt. (2011). *Journeys: Forest Hills School District, Pennsylvania*. Case study.

Houghton Mifflin Harcourt. (2011). *Journeys: Fremont Community School District, Nebraska*.

Case study.

Houghton Mifflin Harcourt. (2011). *Journeys: Ganado Unified School District #20, Arizona*.

Case study.

Houghton Mifflin Harcourt. (2011). *Journeys: Genoe Area Local School District, Ohio*. Case study.

Houghton Mifflin Harcourt. (2011). *Journeys: Lawson School District R 14, Missouri*. Case study.

Houghton Mifflin Harcourt. (2011). *Journeys: New Miamia Local School District, Ohio*. Case study.

Houghton Mifflin Harcourt. (2011). *Journeys: Pleasant Valley School District, Pennsylvania*. Case study.

Houghton Mifflin Harcourt. (2011). *Journeys: York City School District, Pennsylvania*. Case study.

Houghton Mifflin Harcourt. (2011). *Journeys: Springfiled City School District, Ohio*. Case study.

Houghton Mifflin Harcourt. (2011). *Journeys: Westfall Local School District, Ohio*. Case study.

Houghton Mifflin Harcourt. (2016). *Journeys: A research-based approach, grades k-5*.

Retreived from <http://www.hmhco.com/shop/education-curriculum/reading/core-reading/programs/journeys>

Kane, M. T.. (2016). Explicating validity. *Assessment in education: Principles, Policy & Practice*, 23(2), 198-211. Retrieved from

<https://www.tandfonline.com/doi/abs/10.1080/0969594X.2015.1060192>

Koch, A. & Peden, W. (Eds.), (1944). *The life and selected writings of Thomas Jefferson*. New York: Random House.

Laerd Statistics. (2018). *Multiple regression using SPSS statistics*. Retrieved from <https://statistics.laerd.com>

Lorence, D. (2001). Application for M.I.T.C.H. elementary public charter school: Exhibit A. Retrieved from: http://mcs.k12.or.us/uploads/8/2/7/1/82715892/exhibit_a.pdf

Massell, D., & Fuhrman, S. (1994). *Ten years of state education reform, 1983-1993: Overview with four case studies*. Washington, DC: Center for Policy Research in Education. Retrieved from <http://files.eric.ed.gov/fulltext/ED366095>

Nardi, P. M. (2003). *Doing survey research: A guide to quantitative methods*. Boston, MA: Pearson Education.

National Governors' Association. (2008). Benchmarking for success: Ensuring U.S. students receive a world-class education. Washington, D.C.: NGA. Retrieved from <https://www.edweek.org/media/benchmarking%20for%20success%20dec%202008%20final.pdf>

Neal, D., & Whitmore Schanzenback, D.W. (2010). Left behind by design: Proficiency counts and test-based accountability. *The Review of Economics and Statistics*, 92 (2), 263-283. Retrieved from <http://www.jstor.org.georgefox.idm.oclc.org/stable/pdf/27867536.pdf?refreqid=excelsior%3Ad3c7845d1ab17bfb77cd9a6f460e3f65>

Nichols, S. L., Glass, G., & Berliner, D.C. (2012). High-stakes testing and student achievement: Updated analysis with NAEP data. *Education Policy Analysis Archives*, 20 (20), 1-30.

Retrieved from

<http://web.b.ebscohost.com/geo/fox.idm.oclc.org/ehost/detail/detail?vid=0&sid=f7e6ba-d-8afa-4121>

863e946732832cf8%40sessionmgr103&bdata=JnNjb3BIPXNpdGU%3d#db=eue&AN=84357343

Oregon Department of Education. (2017). Understanding student assessment. *Student Assessment: The design, development, and implementation of statewide assessments.*

Retrieved from <http://www.oregon.gov/ode/educator-resources/assessment/Pages/default.aspx>

Oregon Department of Education. (2013). *Oregon Common Core State Standards English language arts (ELA) and literacy, instructional materials contract years 2014 to 2020.*

Retrieved from <http://www.oregon.gov/ode/educator-resources/teachingcontent/instructional-materials/Pages/default.aspx>

Oregon Department of Education. (2015). Achievement standards: 2014-15 achievement standards summary. Retrieved from <http://www.oregon.gov/ode/educator-resources/standards/Documents/asmtachstdsummary1415.pdf>

OregonLaws.Org. (2015). ORS 337.120: *District school board adoption of list and use of approved textbooks and instructional materials.* Retrieved from <https://www.oregonlaws.org/ors/337.120>

OregonLaws.org. (2015). ORS 337.150: *School board and charter school duty to provide textbooks.* Retrieved from <https://www.oregonlaws.org/ors/337.150>

OregonLaws.org. (2015). ORS 337.155: *School board and charter school duty to provide textbooks.* Retrieved from <https://www.oregonlaws.org/ors/337.150>

- OregonLaws.org (2017). ORS 338.015: *Legislative intent*. Retrieved from <https://www.oregonlaws.org/ors/338.105>
- OregonLaws.org. (2015). ORS 338.115: *Applicability of laws*. Retrieved from <https://www.oregonlaws.org/ors/338.115>
- Pattison, K. (2015). Program Analyst, Office of Learning, Oregon Department of Education. Private email (M. Meyer).
- Planning Research & Evaluation Services. (2011). *Houghton Mifflin Harcourt Journeys: Preliminary pilot study report*.
- Polikoff, M.S. (2012). The association of state policy attributes with teachers' instructional alignment. *Educational Evaluation and Policy Analysis*, 34(3), 278-294. Retrieved from: <http://www.jstor.org.georgefox.idm.oclc.org/stable/23254160>
- Polikoff, M.S., Porter, A. C., and Smithson, J. (2011). How well aligned are state assessments of student achievement with state content standards? *American Education Research Journal*, 48(4), 965-995. Retrieved from <http://aerj.aera.net>
- Reback, R. (2008). Teaching to the rating: School accountability and the distribution of student achievement. *Journal of Public Economics*, 92 (5), 1394-1415. Retrieved from <http://www.sciencedirect.com.georgefox.idm.oclc.org/science/article/pii/S0047272707000862?via%3Dihub>
- Reitz, J. (2017). *Online dictionary for library and information science*. Retrieved from http://www.abc-clio.com/ODLIS/odlis_A.aspx
- Resendez, M. & Azin, M. (2013). *A study of the effectiveness of Journeys program: Year 2 final report*. Jackson, Wyoming: PRES Associates, Inc.
- Resendez, M. & Azin, M. (2014). *A study of the effectiveness of Journeys in Texas using existing*

- data sources*. Jackson, Wyoming: PRES Associates, Inc.
- Robson, T. (2017). Curriculum renewal communication. Retrieved from <https://tinyurl.com/yax9u4nv>.
- Sims, D. P. (2012). Can failure succeed? Using racial subgroup rules to analyze the effect of school accountability failure on student performance. *Economics of Education Review*, 32, 262-274. Retrieved from <https://doi.org/10.1016/j.econedurev.2012.12.003>
- Srikantaiah, D. (2009). *How state and federal accountability policies have influenced curriculum and instruction in three states: Common findings from Rhode Island, Illinois, and Washington*. Washington, DC: Center for Education Policy. Retrieved from www.cep-dc.org
- Smarter Balanced Assessment Consortium. (2017). Smarter Assessments. Retrieved from <http://www.smarterbalanced.org/assessments/>
- Sonnenschein, S. B., Baker, L., & Garret, A. (2005). *An analysis of academic progress of children participating in the Core Knowledge preschool program in Baltimore county Head Start Centers*. Charlottesville, VA: Core Knowledge Foundation. Retrieved from <http://www.coreknowledge.org>
- Springer, M.G. (2017). Researcher. Vanderbilt University. Private email (M. Meyer).
- Steiner, D. (2017). Curriculum research: What we know and where we need to go. StandardsWork. Retrieved from <https://standardswork.org/wp-content/uploads/2017/03/sw-curriculum-research-report-fnl.pdf>
- Fairmont School District. (2015). *Our schools*. Retrieved from <https://tinyurl.com/qe9krb7>
- Fairmont School District. (2016). *Oregon report card 2015-16: Fairmont SD 23J*. Retrieved from <https://tinyurl.com/yccyjt4>

U.S. Department of Education. (2015). Beginning reading intervention report: Houghton Mifflin Reading. *Institute of Education Sciences, What Works Clear Clearinghouse*. Retrieved from <http://whatworks.ed.gov>

U.S. Department of Education. (2004). Comprehensive School Reform Program: About CSR (Archived information). *Programs*. Retrieved from <https://www2.ed.gov/programs/compreform/2pager.html>

U.S. Department of Education. (2001). Guidance on standards, assessments & accountability (Archived information). *Laws & guidance/Elementary & secondary education*. Retrieved from: <http://www2.ed.gov/policy/elsec/guid/standardsassessment/guidance.html>

U.S. Department of Education. (2001). Oregon Assessment Letter 3 (Arcived information). *Laws & Education / Elementary & secondary education*. Retrieved from http://www2.ed.gov/admins/lead/account/finalassess/or_3.html

University of Michigan Library, (2017). Research guides: grey literature. Retrieved from <http://guides.lib.umich.edu/c.php?g=283076&p=1886023>

Wedman, J. & Waigandt, E. (2004). *Core knowledge curriculum and school performance: A national study*. Charlottesville, VA: Core Knowledge Foundation. Retrieved from <http://www.coreknowledge.org>

APPENDICES

Appendix A: Common Core State Standards

Due to the large size of the files for the *Journeys* and Common Core documents, these additional appendices have not been added. However, readers interested in viewing these documents can contact the dissertation author at mmeyer@educerelibrarian.org or view the documents directly at ODE Common Core Standards: <https://www.oregon.gov/ode/educator-resources/standards/ELA/Documents/oregon-common-core-state-standards.pdf> and Houghton Mifflin Harcourt:

Journeys CCSS Correlation 3rd Grade: <http://hnhco-v1.prod.webpr.hnhco.com/~media/sites/home/education/global/pdf/correlations/reading/journeys-common-core/journeys-commoncore-grade3-2011-12-ela.pdf?la=en>

Journeys CCSS Correlation 4th Grade:
https://forms.hnhco.com/assets/pdf/journeys/Journeys_CC-Correlations_grade4.pdf

Journeys CCSS Correlation 5th Grade: <http://hnhco-v1.prod.webpr.hnhco.com/~media/sites/home/education/global/pdf/correlations/reading/journeys-common-core/journeys-commoncore-grade5-2011-12-ela.pdf?la=en>

Journeys CCSS Correlation 6th Grade:
https://forms.hnhco.com/assets/pdf/journeys/Journeys_CC-Correlations_grade6.pdf

Appendix B: ODE Memorandum of Understanding
STATE OF OREGON

RESEARCH AGREEMENT/ DATA USE AGREEMENT

This Agreement is between the State of Oregon, acting by and through its **Department of Education** hereafter called "**Agency**" and **George Fox University**, hereafter called "**Researcher**".
Agency's Administrator for this Agreement is:

Administrator: Brian Reeder, Office of Research and Data Assistant Superintendent	Oregon Department of Education, Analysis	Public Service Building 255 Capitol Street NE Salem, Oregon 97310-4285
Email address: Brian.reeder1@state.or.us	Oregon Department of Education website: http://www.oregon.gov/ode	

Researcher's Administrator for this Agreement is:

Administrator: Melissa Meyer	Agency: George Fox University	Address:
Administrator phone number: (503) 939-6725	Fax number:	
Administrator email address: Mmeyer061@georgefox.edu		

Agreement Period

This Agreement shall become effective on the date this Agreement has been fully executed by every party. Unless extended or terminated earlier in accordance with its terms, this Agreement shall terminate on **August 31, 2018** unless extended by a written amendment.

Amendments

This Agreement may be amended. No changes to or waivers of provisions of this Agreement will be valid until they have been reduced to writing, approved and signed by all parties.

Agreement Documents

This Agreement consists of the following documents, which are listed in descending order of precedence: this Agreement less all exhibits, attached Exhibit A (Research Proposal) and Exhibit B (the Research Project Confidentiality Agreement) and Exhibit C (Individual Acknowledgement of Confidentiality Agreement). Exhibits A - C are attached and are hereby incorporated by reference.

Research Work

The Researcher shall perform the Work (the 'Work') as set forth in the Research Proposal and is attached hereto as Exhibit A. - Researcher shall perform the Work in accordance with the terms and conditions of this Agreement.

Termination

This Agreement may be terminated by mutual written consent of both parties, or by either party with thirty (30) day written notice. Agency may terminate this Agreement effective upon delivery of written notice to the Researcher, or at such later date as may be established by Agency. Any termination under paragraph A or B above of this Section shall be without prejudice to any obligations or liabilities of either party already accrued prior to such termination. Agency by written notice of default to Researcher may terminate the whole or any part of this Agreement. If Researcher fails to provide services, if any, called for by this Agreement within the time specified herein or any extension thereof; or if Researcher fails to perform any of the other provisions of this Agreement, or so fails to pursue the work as to endanger performance of this Agreement in accordance with its terms, and after receipt of written notice from Agency, fails to correct such failures within ten (10) days or such longer period as Agency may authorize.

Subcontractors

The Researcher shall not enter into any subcontracts for any of the Work scheduled under this Agreement without obtaining prior written approval from Agency.

No Third Party Beneficiaries. Agency and Researcher are the only parties to this Agreement and are the only parties entitled to enforce the terms of this Agreement. Nothing in this Agreement gives, is intended to give, or shall be construed to give or provide any benefit or right not held by or made generally available to the public, whether directly, indirectly or otherwise, to third persons unless such third persons are individually identified by name herein and expressly described as intended beneficiaries of this Agreement.

Representations and Warranties

Researcher's Representation and Warranties. Researcher represents and warrants to Agency (1) Researcher has the power and authority to enter into and perform this Agreement, (2) this Agreement, when executed and delivered, shall be a valid and binding obligation of Researcher enforceable in accordance with its terms, (3) Researcher has the skill and knowledge possessed by well-informed members of its profession and Researcher will apply that skill and knowledge with care and diligence to perform the Work in a professional manner and in accordance with standards prevalent in Researcher's profession, (4) Researcher shall, at all times during the term of this Agreement, be qualified, professionally competent and duly licensed to perform the Work, and (5) Researcher prepared its proposal related to this Agreement, if any, independently from all other proposers, and without collusion, fraud, or other dishonesty.

Compliance with Applicable Law.

Applicable Law. Researcher shall comply with all federal, state and local laws, regulations, executive orders and ordinances applicable to this Agreement Without limiting the generality of the foregoing, Researcher expressly agrees to comply with the following laws, regulations and executive orders to the extent they are applicable to the Agreement: (i) Titles VI and VII of the Civil Rights Act of 1964, as amended; (ii) Sections 503 and 504 of the Rehabilitation Act of 1973, as amended; (iii) Americans with Disabilities Act of 1990, as amended; (iv) Executive Order 11246, as amended; (v) Health Insurance Portability and Accountability Act of 1996; (vi) Age Discrimination in Employment Act of 1967, as amended, and the Age Discrimination Act of 1975, as amended; (vii) Vietnam Era Veterans' Readjustment Assistance Act of 1974, as amended; (viii) ORS Chapter 659 and 659A as amended; (ix) All regulations and administrative rules established pursuant to the foregoing laws; and (x) All other applicable requirements of federal and state civil rights and rehabilitation statutes, rules and regulations.

These laws, regulations and executive orders are incorporated by reference herein to the extent they are applicable to the Agreement and required by law to be so incorporated. Agency's performance under the Agreement Is conditioned upon Researcher's compliance with the provisions of ORS 2798.220, 2798.225, 279B.230, 2798.235 and 2798.270 which are incorporated by reference herein. Researcher shall, to the maximum extent economically feasible In the performance of this Agreement, use recycled paper (as defined in ORS 279A.010(1)(gg)), recycled PETE products (as defined in ORS 279A.010(1)(hh)), and other recycled products (as "recycled product is defined in ORS 279A.010(1)(ii)).

FERPA. The Family Educational Rights and Privacy Act (FERPA), 20 USC §1232g, applies to education records of individual students held by the Agency. If Researcher has access to personally identifiable education records, it shall not disclose them to anyone and upon completion of the Work it shall destroy the records. Researcher shall comply with all applicable statutes and rules related to FERPA and education records.

Force Majeure. Neither Agency nor Researcher shall be held responsible for delay or default caused by fire, riot, acts of God, terrorist acts, or other acts of political sabotage, or war where such cause was beyond the reasonable control of Agency or Researcher, respectively. Researcher shall, however, make all reasonable efforts to remove or eliminate such a cause of delay or default and shall, upon the cessation of the cause, diligently pursue performance of its obligations under this Agreement.

Survival. All rights and obligations shall cease upon termination or expiration of this Agreement, except for the rights and obligations set forth in Sections 1, 10, 13, and 17.

Notice. Except as otherwise expressly provided in this Agreement, any communications between the parties hereto or notices to be given hereunder shall be given in writing by email,

personal delivery, facsimile, or mailing the same, postage prepaid, to Researcher or Agency at the address, number or email address set forth in this Agreement, or to such other addresses or numbers as either party may indicate pursuant to this Section 14. Any communication or notice so addressed and mailed shall be effective five (5) days after mailing. Any communication or notice delivered by facsimile shall be effective on the day the transmitting machine generates a receipt of the successful transmission, if transmission was during normal business hours, or on the next business day, if transmission was outside normal business hours of the recipient. To be effective against Agency, any notice transmitted by facsimile or e-mail must be confirmed by telephone notice to Agency's Agreement Administrator. Any communication or notice given by personal delivery shall be effective when actually delivered. Any communication or notice given by email shall be effective upon the sender's receipt of confirmation generated by the recipient's email system that the notice has been received by the recipient's email system.

Severability. The parties agree if any term or provision of this Agreement is declared by a court of competent jurisdiction to be illegal or in conflict with any law, the validity of the remaining terms and provisions shall not be affected, and the rights and obligations of the parties shall be construed and enforced as if the Agreement did not contain the particular term or provision held to be invalid.

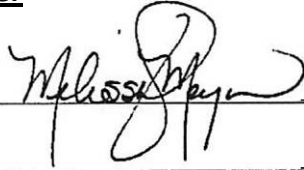
Counterparts. This Agreement may be executed in several counterparts, all of which when taken together shall constitute one agreement binding on all parties, notwithstanding that all parties are not signatories to the same counterpart. Each copy of the Agreement so executed shall

Governing Law; Venue; Consent to Jurisdiction. This Contract shall be governed by and construed in accordance with the laws of the State of Oregon without regard to principle conflicts of law. Any claim, action, suit or proceeding (collectively, "Claim") between Agency (and/or any other Agency of the State of Oregon) and Contractor that arises from or relates to this Contract shall be brought and conducted solely and exclusively within the Circuit Court of Marion County for the State of Oregon; provided, however, if a Claim must be brought in a federal forum, then it shall be brought and conducted solely and exclusively within the United States District Court for the District of Oregon. CONTRACTOR, BY EXECUTION OF THIS CONTRACT, HEREBY CONSENTS TO THE IN PERSON JURISDICTION OF SAID COURTS.

Merger Clause; Waiver. This Agreement and attached exhibits constitute the entire agreement between the parties on the subject matter hereof. There are no understandings, agreements, or representations, oral or written, not specified herein regarding this Agreement. No waiver, consent, modification or change of terms of this Agreement shall bind either party unless in writing and signed by both parties and all necessary State approvals have been obtained. Such waiver, consent, modification or change, if made, shall be effective only in the specific instance and for the specific purpose given. The failure of Agency to enforce any provision of this Agreement shall not constitute a waiver by Agency of that or any other provision.

BOTH PARTIES, BY THE. SIGNATURE BELOW OF ITS AUTHORIZED REPRESENTATIVE, HEREBY ACKNOWLEDGES THAT SHE OR HE HAS READ THIS AGREEMENT, UNDERSTANDS IT AND AGREES TO BE BOUND BY ITS TERMS AND CONDITIONS.

Researcher

Authorized Signature:  Title: _____ Date: 06.2018
Print _____
Signature _____

Oregon Department of Education

Lisa A. Kennedy-Reid

Authorized Assistant
Signature: 