PROGRAM EVALUATION: EFFECTIVE BEHAVIORAL AND INSTRUCTIONAL

SUPPORT SYSTEMS AND STUDENT READING OUTCOMES

by

KATHLEEN RYAN JACKSON

A DISSERTATION

Presented to the Department of Educational Methodology, Policy, and Leadership
and the Graduate School of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Doctor of Education

March 2014

DISSERTATION APPROVAL PAGE

Student: Kathleen Ryan Jackson

Title: Program Evaluation: Effective Behavioral and Instructional Support Systems and Student Reading Outcomes

This dissertation has been accepted and approved in partial fulfillment of the requirements for the Doctor of Education degree in the Department of Educational Methodology, Policy, and Leadership by:

Akihito Kamata	Chairperson
Keith Hollenbeck	Core Member
Scott Baker	Core Member
Brigid Flannery	Institutional Representative

and

Kimberly Andrews Espy	Vice President for Research and Innovation; Dean of the Graduate School

Original approval signatures are on file with the University of Oregon Graduate School.

Degree awarded March 2014

DISSERTATION ABSTRACT

Kathleen Ryan Jackson

Doctor of Education

Department of Educational Methodology, Policy, and Leadership

March 2014

Title: Program Evaluation: Effective Behavioral and Instructional Support Systems and Student Reading Outcomes

This program evaluation studied Effective Behavioral and Instructional Support Systems (EBISS), a Response to Intervention (RTI) initiative focused at changing district leadership behaviors to close the policy-research-practice gap and improve students' reading outcomes at third grade. A pre/post quasi-experimental comparison using a matched group design evaluated the four-year initiative. EBISS districts ($n = 25$) were matched to non-EBISS districts ($n = 25$) by important district variables of: (a) size (small, medium, and large) and (b) region (coast, central, east, south, and valley), (c) students receiving Free and Reduced Meals (FARM), (d) students who were white and non-white, (e) students who were male and female, (f) years of teacher's experience, and (g) third grade Oregon Assessment of Knowledge and Skills (OAKS) scores. The aim of this research was to answer two questions. The first asked whether the performance of students in non-EBISS and EBISS districts were significantly different when measured by their percent of students passing the third grade reading OAKS in 2006-07 (EBISS pre-treatment year) and 2010-11 (EBISS post-treatment year). The second question analyzed whether EBISS districts with high implementation scores made statistically significantly different gains from EBISS districts with low implementation scores on

their percent of students who passed the third grade OAKS reading assessment from 2006-07 to 2012-11. The results of the first research question indicated no significant differences for group ($p = .312$) or time ($p = .488$) between EBISS and non-EBISS districts scores on the OAKS reading test at third grade. Similarly, the second research question results indicated no significant differences ($p = .452$) between EBISS districts with high and low implementation scores on the OAKS reading scores at third grade. Findings are discussed in relation to: (a) the challenge when measuring administrator effectiveness based on distal outcomes, (b) confounding variables that affected internal and external validity, and (c) how this study informs the future design of evaluation research, in the pre-implementation year of an initiative, so variables that are know to be effective in improving student outcomes can be replicated.

CURRICULUM VITAE

NAME OF AUTHOR: Kathleen Ryan Jackson

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon, Eugene

Villa Maria College, Cheektowaga, New York

DEGREES AWARDED:

Doctor of Education, 2014, University of Oregon

Master of Education, 1998, University of Oregon

Bachelor of Science, Psychology, 1997, University of Oregon

Associate of Science, Education, 1972, Villa Maria College

AREAS OF SPECIAL INTEREST:

Educational Leadership and Policy

Student Engagement

PROFESSIONAL EXPERIENCE:

Trainer, University of Oregon, 2009 to present

Principal, Greater Albany Public Schools, 2007-2009

Assistant Principal, Springfield Public Schools, 2005-2007

Alternative Education Program Coordinator, Corvallis Public Schools, 2000-2005

Special Education Teacher, Corvallis Public Schools, 1998-2000

GRANTS, AWARDS, AND HONORS:

Service Learning Citizenship Award, SOLV, 2008

PUBLICATIONS:

Chaparro, E. A., Smolkowski, K., Baker, S. K., Hanson, N. & Ryan Jackson, K. M. (2012). A Model for System-Wide Collaboration to Support Integrated Social Behavior and Literacy Evidence-Based Practices. *Psychology in the Schools. 49*(5), 465-482.

Chaparro, E. A., Ryan Jackson, K.M., Baker, S. K., & Smolkowski, K. (2012). Effective Behavioral and Instructional Support Systems: An Integrated Approach to Behavior and Academic Support at the District Level. *School Based Mental Health Interventions 5,* 161-176.

Chaparro, E. A., Park, Y., Baker, S. K., & Ryan Jackson, K. M. (2011). *District system support plan: A district level self-assessment tool*. (Report No. 1104). Eugene, OR: Center on Teaching and Learning, University of Oregon.

ACKNOWLEDGMENTS

DEDICATION

This manuscript is dedicated to Carol Sadler who had a vision of how a district's leadership could support its schools successful implementation of a blended model of academics and behavior to support the needs of every single student. Carol your legacy continues to live and grow.

Dedication is also offered to my husband, Randy A. Jackson, who encouraged me to return to school, for without his steadfast encouragement across the years I would have never been touched by or had the opportunity to touch the lives of so many amazing students.

Finally, to my daughter, Lilly Rose Jackson who continues the legacy that education is every student's civil right and the belief that every student can reach their optimal potential through equitable academic and behavioral support systems in inclusive environmental settings.

TABLE OF CONTENTS

LIST OF FIGURES

xvi

LIST OF TABLES

CHAPTER I

INTRODUCTION

Purpose of the Study

The purpose of this study was to evaluate the Effective Behavioral and
Instructional Support Systems (EBISS) initiative and determine if the professional
development (PD) and technical assistance (TA) provided to 25 EBISS districts in the
State of Oregon resulted in an increase in the reading outcomes of third grade students,
over and above a matched group of 25 non-EBISS districts from time one (2006-07
EBISS pre-treatment year) to time two (2010-11 EBISS post-treatment year).
Performance of third-grade students on the Oregon Assessment of Knowledge and Skills
(OAKS) reading subtests for non-EBISS and EBISS districts were compared using a pre
post quasi-experimental comparison on a matched group design to evaluate the four-year
EBISS initiative. EBISS and non-EBISS districts were matched by district, student, and
teacher variables: (a) district size (small, medium, and large) and (b) district region
(coast, central, east, south, and valley), student's (c) receiving Free and Reduced Meals
(FARM), (d) white and non-white students, (e) male and female, (f) teacher's experience,
and (g) 3rd grade OAKS scores. The matched comparison group was chosen to minimize
selection bias to the extent possible within this quasi-experimental study while testing
differences between non-EBISS and EBISS districts (Shadish, Cook, & Campbell, 2002).

In addition, a secondary evaluation was conducted *within* EBISS districts to
determine if there were differences in EBISS districts with high and low implementation
scores on the District Systems Support Plan (DSSP) a district level self-assessment that
measured changes in the strength of a district's system over time. The secondary

1

evaluation was designed to answer the question did EBISS districts with high implementation scores make statistically significantly greater gains than EBISS districts with low implementation scores on the percent of students who passed the third grade OAKS reading assessment from time one to time two.

Comparison of student outcomes in non-EBISS and EBISS districts would shed light on the effectiveness of implementing EBISS to improve early literacy outcomes through PD and TA focused at the district leadership level using a common Response to Intervention (RTI) district framework (Chaparro, Smolkowski, Baker, Hansen, & Ryan Jackson, 2012; Chaparro, Ryan Jackson, Baker, & Smolkowski, 2012). The framework integrated policy, research, and practice through instructional leadership that was distributed across all levels of the organization (Carnine, 1999; Fixsen, Blase, Metz, & Van Dyke, 2013; Honig, Copeland, Rainey, Lorton, & Newton, 2010; Lane, 2009; Leithwood & Jantzi, 2008 Leithwood & Jantzi, 2008; Rorrer, Skrla, & Scheurich, 2008; Spillane, Halverson, & Diamond, 2001). The EBISS initiative blended two evidence-based practices (EBP), PBIS and RTI for early literacy. This study will only focus on evaluating early literacy outcomes at the third grade level.

The aim of this research was to provide evidence for the hypothesis that district leadership and policy matters when it ensures schools receive the support they require to effectively implement early literacy in every school through the use of EBP in a common RTI framework (Algozzine et al., 2010, Baker et al., 2010; Baker et al., 2011; Carnine, 1999; Fixsen et al., 2013; Greenwood & Abbott, 2001; Honig et al., 2010; Lane 2009). This study contributes to the literature that suggests district leadership (defined as including board, district and school administrators) knowledge and skills are a

prerequisite for implementing RTI for early literacy district-wide, as the district supports

the development of sustainable EBPs practices in early literacy to close the policy-

research-practice gap (Algozzine et al., 2010; Chaparro, Smolkowski et al., 2012; Sadler

& Sugai, 2009).

Research Questions

This program evaluation aims to answer the following questions. First, was there

a difference between non-EBISS and EBISS district's students' gains in the reading sub-

tests of the OAKS at third grade from time one to time two? Second, were students' gains

in the reading sub-tests of the OAKS at third grade from time one to time two different

for EBISS districts with high DSSP scores versus EBISS districts with low DSSP scores?

Examining whether the PD and TA delivered to EBISS districts resulted in gains in the

percent of students who passed the OAKS, over and above non-EBISS districts and

within EBISS districts, will inform future efforts to implement EBISS for early literacy in

a common RTI framework where: (a) districts goals and priorities are clearly defined, (b)

organizational supports are aligned (c) schools have the differentiated resources required

(d) staff competency is increased (d) adequate funding is allocated, and (e) fidelity to RTI

practice is guided by policy. Answering the second question will inform if the DSSP

fidelity measure of district leadership behaviors that are know to increase high quality

implementation of early literacy resulted in improved reading outcomes for students at

third grade.

Search Strategies

To gather peer reviewed research articles for my dissertation I searched

combinations of the following terms: (a) district leadership, (b) early literacy, and (c)

school district policy. Table 1 identifies the combination of terms that were applied under each of the three categories. I searched electronic databases as well as the worldwide web to gather articles from researchers at universities, school districts, Oregon Department of Education (ODE), think tanks, non-profits, and policy centers. Articles were selected for review if they were peer reviewed in a journal. Next, references within all articles were examined to identify additional studies, that had to be peer reviewed, and were relevant to my topic of interest. The articles generated from the literature search provided a rich collection of articles on district leadership models in general, district leadership models in early literacy, research to practice, and district policy.

Table 1

Electronic Search Strategy

Terms	Data Bases	Result
District leadership: district leadership and student outcomes, district leadership models and frameworks, and district leadership and reading	ERIC, Google Scholar, University of Oregon Library	75 articles and books selected for review
Early literacy: RTI, student achievement third grade, fidelity, best practice, evidence based practice, reading and early literacy models	ERIC, Google Scholar, University of Oregon Library	86 articles and books selected for review
School district policy: data based decision-making, collaboration, district policy and reading, research to practice, scaling up educational practice	ERIC, Google Scholar, University of Oregon Library, Worldwide Web, National Association of School Psychologists, Council for Exceptional Children	63 articles and books for review

What the Research Says About Early Literacy Outcomes

In a longitudinal study of reading scores from 1971 to 2009, the National Center for Educational Statistics (NCES) reported reading scores remained flat for fourth to eighth grade students across the United States (NCES, 2010) and the reading scores were unchanged in 2011 (NCES, 2011). Sadly, The Nations Report Card: Reading 2011 (NCES, 2011) reported that the gap widened for students in Colorado, the District of Columbia, Oregon, Vermont, and Washington from 2003 to 2011 with no statistically significant changes in the reading scores for African American and Hispanic students from 2009 to 2011. At the same time, the NCES (2011) reported that female students scored higher than male students by an average of seven points on the fourth grade reading assessment. Furthermore, a longitudinal study on students referred for special education services found that by the late elementary grades high levels of social behavior becomes chronic, and is associated with negative educational outcomes (Walker, Colvin, & Ramsey, 1995). Unfortunately, these negative school outcomes may be related to the fact that approximately 75% of the prison populations in the United States are poor readers (Kamps, Wills, Greenwood, & Thorne, 2003).

The national reading data (NCES, 2011) indicated that districts would not meet state and federal requirements of the Elementary and Secondary Education Act (ESEA, 2010) and the No Child Left Behind Act (NCLB, 2001) that called for improved educational outcomes for all students, closing achievement gaps, increasing equity, and improving the quality of instruction. Fiester's (2010) executive summary of early reading, *Early Warning! Why Reading by the End of Third Grade Matters,* predicted that if the current trend in the data continued low-income students would be at increased risk

and fail to graduate from high school on time because they would not meet proficient levels of reading on the National Assessment of Educational Progress (NAEP) by the end of third grade. Subsequently, attention was drawn to the relationship between district behaviors (board, district and school administrators), district policy, and equitable support to the schools (Lane, 2009; Rorrer et al., 2008).

What We Know About District Leading Change

Recent research on the relationship between district leadership practices and improved student outcomes gained the attention of policy makers, researchers, and educators at the federal, state, and local level (Baker et al., 2011; 2006; Fixsen et al., 2013; Lane, 2009). For example, Oregon developed standards to guide effective educational leadership and administrator practices (Oregon Educational Leadership and Administrator Standards, 2008). District leadership was defined as an *organized collective* (Rorrer et al., 2008) comprised of the board, district administrators and school principals who develop a partnership (Honig et al., 2010) and a critical link that unifies schools and the district to identify problems and develop viable solutions to change (Fixsen et al., 2013). This critical link between district and its schools creates clear communication loops (district-to-schools-to-district) and improvement cycles (on-going use of data to identify and solve problems at the district and school level) to provide equitable services to *all* of a district's schools (Fixsen et al., 2013; Rorrer et al., 2008). This new model of partnerships and links between a district and its schools is a major shift in thinking from the more traditional model of district serving a single school or what Togneri and Anderson (2003) termed *islands of excellence.*

Historically, a single school (or a few schools) in a medium to large district volunteered to participate in a new initiative. This created an *island of excellence*, because other schools would be given permission to remain resistant to change, because policy was not aligned and goals were not clear or enforced. These islands of excellence created inequitable services and resulted in a disproportionality of outcomes for the most vulnerable students (NCES, 2011), because the quality of instruction a student received was dependent on the teacher or school a student attended. To reverse this long-held trend the federal government called for states, districts, and school leaders to provide professional development to all staff using EBPs to ensure that all students regardless of their income, race, ethnic and language background, or disability would be prepared to graduate and realize their post secondary goals (ESEA, 2010). For example, Oregon developed a coordinated system of professional development for educational leaders and administrators (ODE Educator Effectiveness, 2013), as well as Educational Leadership and Administrator Standards (Oregon Educational Leadership and Administrator Standards, 2008) to be included in all evaluations of administrators in a district (ODE: SB 290). The standards clearly defined six domains: (a) visionary leadership, (b) instructional improvement, (c) effective management, (d) inclusive practice, (e) ethical leadership, and (f) a coordinated framework for evaluation. The standards of practice are intended to improve professional practice and teaching and learning in schools.

District models. Departments of Education, districts, and researchers collaborated to identify viable district models to improve the teaching and learning in schools and respond to the rigorous accountability measures put in place at the federal and state level (ESEA, 2010). The literature (e.g., case study and theory of district

leadership behaviors) recommended that districts demonstrate a deep knowledge of EBPs to meet the needs of *all* students through the rigorous analysis of data (Slavin et al., 2013) and have the skills to effectively implement high quality practices over time (Lane, 2009; Fixsen et al., 2013; Rorrer et al., 2008). This process provided direct support to principals (Honig et al., 2010) who in turn support teachers sustained, high quality implementation of reading programs with strong evidence of effectiveness in the classroom (Baker et al., 2011; Robinson, Lloyd, & Rowe, 2008). When the focus is on *all* schools in a district, versus one or two schools, district-wide implementation efforts are *scalable* across all schools (Fuchs & Fuchs, 2006) and rapid district improvement can be expected (Lane, 2009).

Current literature on scalable district models for change share common elements that are driven by clearly articulated policy to cultivate a collective commitment to change through equitable resource allocation and the development of staff skill at all levels of the system (Fixsen et al., 2013). The studies conducted over the past ten years provide converging evidence that proposes district leadership practices have a positive impact on improved early literacy outcomes through the implementation of a multi-tiered prevention model (Baker et al., 2011; Fixsen et al., 2013; Fuchs & Fuchs, 2006; Honig et al., 2010; Lane, 2009; Rorrer et al., 2008; Slavin et al., 2013).

Multi-tiered prevention models. Fuchs and Fuchs (2006) described RTI as *scalable* multi-tiered prevention model that makes use of EBPs that are implemented with fidelity by skilled practitioners who understand both assessment and interventions. Odom et al. (2005) identified an EBP as a practice that has been proven to be effective through research and scientific evidence, and Gersten et al. (2008) extended the definition to

include the importance of implementing a practice exactly the way it was developed in research in order to consider it evidence based in the classroom (often referred to as fidelity of implementation). In order to implement RTI with fidelity the National Center for Response to Intervention (NCRTI, 2013) identified four critical components that must be in place: (a) multi-level prevention system where data analysis and decision-making occur at all levels of implementation and all levels of instruction (i.e., board, district, school, classroom), (b) screening of all students to identify students who may be at risk of poor learning outcomes (e.g., academic and behavioral), (c) progress monitoring frequently to assess student's academic and behavioral performance to quantify a student's rate of improvement or responsiveness to instruction, and (d) data based decision-making that occurs at all levels of implementation and instruction. When the four critical components are in place, "Everyone knows what to implement because there is but one protocol, which makes training easier to accomplish and fidelity of implementation easier to assess and ensure; in turn, this makes it more likely that it can be *scaled-up* in a district or a building (Fuchs & Fuchs, 2006, p. 96). Although it is true that large-scale studies of scalable district models are limited (where district is the impetus for change) peer reviewed case studies and meta-analysis contribute to a body of valuable information regarding how district can reorganize to effectively support elementary schools to implement high quality reading programs to improve student outcomes. To this end data based decision-making models are emerging.

Data based decision-making. Central to the development of district as the impetus for change is a deep knowledge of how to use data to identify the needs of students and respond with the appropriate practice (Slavin et al., 2013). Slavin et al.

provided consultation on the effective use of data to 59 districts in seven states over four years to implement a district reform model. District leaders were taught to use data to identify weakness, recognize strengths, and select EBPs that met the identified student need. They found educationally significant effects for districts' elementary schools that used data to adopt EBPs for reading (42% adopted a proven reading program by year four) (Slavin & Madden, 2011). However, Slavin et al. advised that district examination of student outcome data is not enough. Schools must be supported by the district to take action, based on the data, and implement practices known to be effective through the research (EBPs) to produce educationally significant gains in reading at the elementary level. This means that staff at all levels of a district (district, school leaders, teacher leaders, individual teachers) must receive on-going training and coaching to effectively use data and improve student's reading outcomes.

Data based decision-making and problem solving. Todd et al. (2011) tested the effectiveness of a team-based problem-solving model within the context of Positive Behavioral Intervention & Support (SWPBIS) using a quasi-experimental single subject group design in three elementary schools. The School-Wide Information System (SWIS) was the data source teams used to monitor office discipline referrals and engage in team-based problem solving using the Team Initiated Problem Solving (TIPS) model. The study examined if there was a functional relationship between TIPS training and teams problem solving behaviors.

The Decision, Observation, Recording and Analysis (DORA) instrument measured (1) *meeting foundation* practices, and (2) *thoroughness* of a team's use of data to define problems, build solutions and organize action plans. Analysis of DORA data

during and after baseline demonstrated a functional relationship for meeting foundations and thoroughness of problem solving in two of the three schools under investigation. Team members also completed a Problem Solving Feedback form pre post to assess their use of data, effectiveness of solutions, and satisfaction with meeting procedures and results. Teams reported low to modest ratings before the TIPS training (19.5%) and modest to high after the TIPS training (90.5%). The findings suggested a need for continued research to replicate the finding of Todd et al. (2011) and assess the generalizability of the model to other schools in other states, other levels (middle and high), and the relationship between improved behavioral outcomes and improved academic outcomes (Newton, Horner, Algozzine, Todd, & Algozzine, 2012). Improving the behavioral and academic outcomes of all students in a district is the goal of every district. Therefore, state and national elementary reading models (Oregon Reading First, Reading First) emerged over the past ten years to support districts high quality implementation of early literacy to ensure that all students were ready to read by the end of third grade.

Large Scale Elementary Reading Models

Scientifically validated early literacy research was gathered at the national and state level in order to develop programs that would support districts efforts to improve elementary reading achievement for all students (e.g., Reading First, Oregon Reading First, and Florida Problem Solving Response to Intervention Project). The research was clear, high quality implementation of reading programs with scientific evidence of effectiveness relies on the on-going use of data to improve teacher practice and student outcomes (Baker et al., 2010; Fuchs & Fuchs, 2006). Important in this effort was the

knowledge that district-wide policy must communicate the district vision and goals for early literacy to guarantee that every school in the system was implementing the same program components and providing equitable services to all students (Fuchs & Fuchs, 2006). Oregon Reading First (Baker et al., 2011) and Florida Problem Solving Response to Intervention Project (PS/RtI: Stockslager, Castillo, Hines, Batsche, & Curtis, 2013) are two examples of statewide reading initiatives aimed at improving the reading outcomes of students at the elementary level.

Oregon Reading First. Baker et al. (2011) evaluated the Reading First initiative in the State of Oregon and its impact on student's reading outcomes. Over 3000 K-3 students in 51 schools were studied while comparing two cohorts (Cohort A began in 2003 and Cohort B began in 2005) to examine their pattern of student performance. Extensive PD and TA were provided to district and school leaders to implement the *School-wide Reading Model* (SWRM: Kame'enui & Simmons, 1998). SWRM identified seven critical features of a comprehensive elementary reading program: (a) practices focused on essential beginning reading content, (b) valid and reliable data informed instruction, (c) protected and sufficient time was allocated, (d) research based practices were applied, (e) multi-tiered differentiated instruction based on student need was implemented, (f) school leaders used data to support and sustain high quality implementation, and (g) the provision of on-going PD and TA was provided.

Baker et al. (2011) found schools that implemented SWRM were similar on eligibility criteria in the first year of implementation (e.g., demographics and mean reading performance). After two years of implementation the performance gap between Cohort A and Cohort B diminished suggesting that high quality implementation improves

12

over time and leads to corresponding improvements in student reading achievement. It is important to note that the researchers brought attention to the fact that many districts implementing Reading First established district-wide policy to implement the Reading First instructional practices in schools (e.g., common core curriculum, reading coaches, and a focus on the five elements of early literacy). The establishment of policy can (a) ensure schools have the curriculum that meets the needs of its students (Slavin et al., 2013), (b) guarantee protected and sufficient time to provide equitable instruction to all students, and (c) provide the professional development district and school staff require to implement the EBPs known to be effective through scientific research (ESEA, 2010). Providing the professional development district requires to meet ESEA goals and implement RTI models with fidelity in schools was also the focus of the statewide initiative in Florida.

  Florida Problem Solving Response to Intervention Project. The PS/RtI Project (Stockslager et al., 2013) is a statewide RTI initiative that began in 2006 in partnership with the University of South Florida and the Florida Department of Education. At the end of the 2011 school year, 27 pilot schools in six districts were evaluated for their progress in three areas: consensus, infrastructure and implementation (improvement in student outcomes was not evaluated). First, consensus was identified as district and school participation in and support of the PS/RtI model. Second, infrastructure was defined as the development of school structures and staff skill. Lastly, implementation was described as the extent to which schools actually implemented the components of the PS/RtI model (NCRTI, 2013).

Stockslager et al. (2013) developed two surveys: (a) Perceptions of RtI Skills Survey, and (b) RtI Beliefs Survey. The skills survey (known to some as a needs assessment) was administered to all staff, at all levels of the district system, to identify their perception of the skills they possessed to effectively implement the critical components of RTI. Then, a beliefs survey was administered to determine staff belief in the ability of students to achieve grade level benchmarks. At the same time project staff delivered six days of PD to districts, as well as five days of PD to district teams in the first three years of implementation. In addition, on-going TA was provided and differentiated based on the results of the district skills and beliefs surveys. In year four the same five days of training was delivered to district leadership teams to develop district capacity to train all staff (training of trainers model) to ensure fidelity to RTI was sustained past the life of the initiative. For example in 2011 over 600 site visits were made to 26 districts, an average of 23 site visits per district in one year.

In the final evaluation the researchers found that indicators of beliefs and consensus were evident in the first three years of implementation. However, once TA and PD were no longer available decreases in schools beliefs regarding the abilities of students with disabilities decreased. Staff perception of their ability to use data and align instruction and interventions also decreased. Given the findings the researchers suggest that additional PD and capacity building activities may be required to sustain implementation efforts of the PS/RtI model. The authors also reported that inconsistent district policies and procedures for implementation of the PS/RtI model posed a barrier to districts sustained implementation efforts.

Reforming District Office Practice

Recent scholarship on district office practices (Honig et al., 2010; Lane, 2009; Sadler & Sugai, 2009) and the NAEP results from 2009 that revealed 83% of the children from low-income families failed to meet the proficient level in reading at fourth grade create *clarity and moral urgency* and *a call to action* for school districts to develop *early warning systems* to improve reading outcomes by fourth grade (Fiester, 2010). The research converges on the finding that district leadership must be transformed (e.g., coherent and aligned policy, equitable funding and resource allocation, district and school partnerships) and practices must be interconnected to support high quality implementation of early literacy in schools (Baker et al., 2011; Fuchs & Fuchs, 2006; Lane, 2009; Sadler & Sugai, 2009).

District office transformation. Honig et al. (2010) found that a *policy imperative* was required to provide specific and concrete examples of how the work of all staff must be reoriented to teach *all students at ever-higher levels*. In order to provide specific and concrete examples, Honig et al. examined district leadership practices in three large urban school districts that were demonstrating higher than expected gains in student achievement (i.e., Atlanta School District showed an average increase of approximately 35% on the state reading/literacy assessment district wide from 2000 to 2003). The aim of the study was to identify the daily work practices and activities of district and school administrators in three large urban school districts to understand how they worked together to build capacity for instructional leadership that resulted in the gains in student achievement they realized. They found common themes across all three districts. First, districts applied their new way of practice systemically across all programs and initiatives

(it was not focused on one initiative, but any best practice). Second, district

administrator's main work increased principal's instructional leadership skills. They

were responsible for developing a *learning-focused* partnership with every school

principal to strengthen their instructional leadership skills through differentiated support.

Third, time was allocated to sustain on-going district-principal partnerships. Fourth,

district administrators modeled how to use and understand data. Modeling how to use

data for problem solving and action planning helped principals and school teams

recognize how instructional practices might be contributing to student test scores. Finally,

on-going assessment and evaluation of principal's skills was used to determine the

individualized supports principals required to build principal capacity district-wide. As

might be expected, the authors suggested that requiring this level of collaboration and

accountability between a district and its schools required a policy imperative that all

districts will face when engaging in this level of system-wide reform (Honig et al., 2010).

      District capacity building and improvement. Lane (2009) also found that

changes in policy were necessary to institute an integrated approach to organizational

change and district-wide improvement. Lane developed a *framework for district*

improvement following a detailed literature review and case study of two districts in the

State of Kansas (one rural, one suburban). His aim was to generate a theory of district

improvement. Two inquiry questions guided his work: (a) how has district (school board,

superintendent, district administrators) promoted and supported improvement efforts, and

(b) what are the key functions (systems, policies, processes) that district's developed to

sustain improvement efforts. Districts were chosen for the case study based on their

dramatic and rapid improvement in student achievement (like Honig et al., 2010) and for

their unique context (urban and rural) to allow for increased generalizability. The rural district of 275 students increased student performance in reading district wide (from 50% in 2004 to 91.7% in 2009). The urban district of 19,000 students increased student performance in reading district wide as well (from 11% in 1996 to 58% in 2008). Lane gathered a variety of data from each district to identify the conditions (practices and behaviors of adults) that could account for this level of improved student improvement.

Interviews and focus groups were conducted with a range of stakeholders following an extensive review of relevant district data to better understand the conditions in each district that resulted in improved student outcomes. Lane attributed the increase in student's reading scores to a new model of accountability (non-negotiable expectations for the organization of instruction and development of staff skill) and a particular set of leadership skills (policy alignment and resource allocation directed at student learning). Equally important was the development of a deeper understanding of the conditions that must be fostered (collective ownership) through collaborative problem solving. With this information, Lane developed a framework for District Capacity Building and Improvement (see Figure 1).

Lane (2009) identified two important steps that were required to *catalyze conditions for rapid improvement*. First, a district must define and communicate its vision and intent to support district-wide improvement. Once the vision was clearly communicated the second step was for a district to reorganize its systems and develop a collective and shared responsibility to build leadership and instructional capacity. This process results in *opportunities* (policy, funding), *incentives* (a sense of awareness and

urgency), and the development of *capacity* to realize district goals. Lane's model

provides districts with a theory or pathway to rapid district improvement.

Framework for District Capacity Building and Improvement - Rapid District Improvement Pathway

Catalyzing Conditions for Rapid Improvement

Capacity
A minimum threshold of capacity to initiate work

Incentives
A catalyzing event that heightens awareness and increases urgency

Opportunity
A window of opportunity - e.g., policies, funding or other mandates

Defining Districtwide Improvement Efforts

The District **Defines and Communicates** its vision and intent to support a **Districtwide Improvement Effort**

Becoming an Improvement-Oriented Learning Organization

Intensive development of district's improvement infrastructure (systems, structures, and culture that support student learning)

District Reorganizes to fully support Improvement Efforts

Reorienting District culture towards shared responsibility and accountability

Collective Problem Solving Providing dedicated time, space, and autonomy

Building Leadership and Instructional Capacity
Focus on improving instruction - Focus on improving relationships

Rapid and Sustained Improvement

Figure 1

Framework for district capacity building and improvement

District leadership theory. Like Lane, Rorrer et al. (2008) investigated district as

the unit of study to develop an understanding of how district could successfully

contribute to successful reform to *improve achievement and advance equity*. Rorrer et al.

developed a theory of district leadership as *institutional actors* and contended that 20

years of research on reform has overlooked, ignored, and even dismissed district potential

to institutionalize a coherent model (or framework) for increasing the achievement of all students. Rorrer et al. reviewed 81 peer-referred articles from 1946 to 2006 for the purpose of developing a narrative synthesis and research based understanding of district as the unit of study in district reform efforts. They identified themes of successful systemic reform (all schools in a district engaging in reform, not just one or two) as they aimed to answer three questions: (a) what is the role of district in reform, (b) what roles could district play to improve achievement and advance reform, and (c) what nature of change would be required to systematically improve achievement and advance equity?

Four essential responsibilities of districts in successful reform were identified: (a) provide instructional leadership (e.g., collective responsibility of board, superintendent and administrators to generate will and build capacity through vision and personal engagement), (b) reorient the organization (e.g., structures and processes are aligned with beliefs, expectations, and norms to unite the district and its schools through a shared commitment), (c) establish policy coherence (e.g., reconcile policy to reveal organizations purpose and align adequate resources with commitment, and (d) maintain an equity focus. The process of maintaining an equity focus was described by Rorrer et al. as a district identifying how past practices have not provided the same services to all students through equitable access (e.g., inability of students to attend an after school supplementary reading program because they do not have transportation home is a situation that can be termed *owning past inequity*). Rorrer et al. also advocated for a district moving from rhetoric to operationalizing an equity focused plan (e.g., a districts vision is clearly operationalized in the mission and policy rules with specific examples of the programs and practices in place will support all student's needs). In sum, Rorrer et al.

(2008) presented a theory of district leadership and an associated framework that district could adopt to advance equity (all students have access to programs and practices) and achievement for all students through strong district and school partnerships (Honig et al., 2010; Lane, 2009).

District leadership behaviors. Taken together the research implied that district partnerships with *all* schools in a district can have a positive effect on sustainable reform efforts that benefit every student in every school (Lane, 2009; Honig et al., 2010; Rorrer et al., 2008). Leithwood and Jantzi (2008) also hypothesized that strong partnerships between a district and its schools would result in improved student outcomes. They randomly selected 180 schools in 45 districts within nine states to investigate how district contributed to principal efficacy. The authors define efficacy as, "a belief about one's own ability (self-efficacy), or the ability of one's colleagues collectively (collective efficacy), to perform a task or achieve a goal" (p. 497). Leithwood and Jantzi (2008) administered two surveys to measure principal and teacher efficacy in relation to the district context (e.g., clearly defined district goals and priorities, access to data and the skills to use data, adequate professional development, voice in decision-making, positive district school partnerships). State reading achievement data was also collected from 2003 to 2005 for each school that participated to determine the percentages of students who *meet* or *exceeded* state benchmarks in reading from time one to time two. Using standard regression coefficients the most direct effects of district leadership were noted on the creation of specific district conditions believed to be effective in *producing student learning* (.77) and accounted for 60% of the variance in district conditions. However, leader efficacy did not explain significant variation in the annual achievement scores of

20

students. This finding aligns with the research of Grissom, Kalogrides, and Loeb (2012) and Cook (2000) who submitted that variation in implementation quality was related to variation in distal outcomes.

While gains in student achievement was not noted the aim of Leithwood and Jantzi (2008) was to identify *the links in the chain* that joined strong leadership practices with student learning. The authors concluded that district's with strong leadership behaviors were more likely to build the confidence and collective efficacy of principals. Districts that emphasized their priority to achievement and instruction through clear priorities and targeted, differentiated support for school improvement developed supportive relationships with schools. Like Honig et al. (2010), Lane (2009) and Rorrer et al. (2008), Leithwood and Jantzi (2008) found that clearly defined district goals and priorities, development of staff skill, and positive district and school partnerships can result in the a collective commitment to a common goal.

What made the Leithwood and Jantzi (2008) study unique was the call to identify antecedents to strong and weak district leadership behaviors. Antecedents are identified as the conditions associated with: (a) redesigning the organization, (b) developing people, (c) managing instructional programs, and (d) setting direction. Antecedents were measured through surveys and interviews in order to describe the size and nature of leadership's impact on the fidelity of instruction and its relationship to student outcomes. Hence, the authors suggest that a district can measure the size and nature of change in adult behavior that is related to improved student outcomes at the school level, so effective district behaviors can be replicated.

Similarly, in a four year study of four elementary schools implementing a school-wide, prevention model of academic (early literacy) and behavioral support systems, Ervin et al. (2006) found schools that received direct support from the district to implement instructional practices demonstrated common practices across schools. They found that a common implementation protocol (goals and priorities) was related to the alignment of the organizational systems and the policies within the system (Ervin et al., 2006). Direct support to schools from the district included working in collaboration with building teams to develop data systems for problem solving and solution development, and crafting actionable goals in response to school's data. In the Ervin et al. study, one principal reported that district supported the organizational structure and helped the school plan for student success based on their school's unique context (e.g., size, student demographics, staff skill level, administrator skill level). Robinson et al. (2008) also found that schools who were documenting improvement credited district leadership with considerable responsibility for positive changes in teachers effectiveness.

In 2008, Robinson el al. developed a strong argument for district to clearly articulate its goals and priorities and apply the instructional leadership skills that guide the work of principals who support teachers in the classroom. In a meta-analysis of 27 studies on the relationship between leadership and student outcomes, Robinson et al. (2008) found that school leaders who promoted and participated in teacher development had an average effect size of 0.84 standard deviations from the mean. Promotion and participation was defined as principals directly participating in PD and data team meetings as a collaborative decision-making partner. Robinson et al. also found that the outcome; higher achievement gains in elementary reading were a property of leaders

active participation as instructional leaders (achievement is the top goal and embedded into routines and procedures) and the organization of schools (objectives or guidelines are the focal point of reading instruction). In high preforming schools teachers reported that principal leaders where highly involved in instructional matters that impact student achievement: (a) setting clear performance standards, (b) guiding teachers use of data to evaluate student progress and adjust teaching, and (c) providing feedback on teacher performance. Principals who were instructional leaders were seen as a source for instructional advice and more respected by the teaching staff.

Systems can change and initiatives can be implemented when the district vision, goals, and expectations are clearly communicated and accompanied with direct support for schools through aligned policy. The importance of district policy supporting a district's goals and priorities through clear expectations was identified as a critical step to high quality implementation in all of the system wide reform models presented. Of equal importance was the need for organizational systems to be aligned to support effective district and school practices. To this end adequate funding and equitable resource allocation must be available to develop the skills of all staff at all levels of the system.

Five Practices of a Highly Effective District

The research to date on effective district behaviors can be summarized into five practices: (a) district policy guides high quality implementation of school's practices, (b) district's vision and mission guide the development of goals and priorities that are clearly defined, (c) organizational systems are aligned, (d) funding and resources are adequate, equitable and differentiated based on schools needs, and (e) on-going development of staff skill is a priority.

Policy. Fixsen et al. (2013) referred to strategic planning and accountability as a new way of working to close the policy to practice loop: a system where policy and funding structures support high quality implementation of EBPs through the active support of implementation teams at the district and school level. Carnine (1997) brought attention to the need for policy and research to work congruently in an attempt to close the implementation gap between research and practice. Greenwood and Abbott (2001) discuss implications and solutions for closing the research-to-practice gap when implementing early literacy. They suggest increasing opportunities for researchers and practitioners to work together, so researchers have first hand knowledge of the challenges schools and teachers face. This collaboration can result in the identification and co-construction of research that is a contextual fit for the problems schools face (i.e., the academic and socio-behavioral needs of students and the skills staff require to implement EBPs). Bridging the policy-research-practice gap will require a paradigm shift that brings together policy makers, researchers, and practitioners who are all focused on viable solutions to solve the real problems a district's schools' face as they strive to ensure all students learn to read by fourth grade (Carnine, 1999).

Carnine (1999) claimed that there was a need for a coalition of influential decision makers (top-down national and state policies and bottom-up teachers, teacher unions, etc.) to: (a) remedy deficiencies in accountability systems, (b) support implementation of EBPs as they are designed, and (c) provide the required funding, PD, and management of personnel. However, Cuban (2004) advised that there was a lack of coordination between policy mandates and what was actually implemented in the

classroom due to a lack of knowledge, skills, and commitment to enact policy by district administrators.

Fixsen et al. (2013) concurs that effective implementation of policy goals at the state, district, and school level must be coordinated through top-down (district) and bottom-up (teachers) collaboration. This will support the development of staff skill and capacity at the district and school level, and ensure every student benefits from a district's high quality implementation of EBPs. Fixsen et al. (2013) recommended a systematic approach to facilitate a *policy to practice communication loop* at the state, district, and school level where district policy is clearly articulated to all stakeholders (administrators, teachers, district staff, students, families, and the community). To this end we must consider the importance of a district clearly articulating it's mission, vision, goals, and priorities to all stakeholders.

Vision, mission, goals, and priorities. Development of a districts' vision and mission and the subsequent alignment of a district's goals and priorities to district policy are critical to high quality implementation across a district's schools (Fixsen et al., 2013; Lane, 2009; Rorrer et al., 2008). A districts mission operationalizes how the vision will become a reality for every student the district serves (Lane, 2009; Rorrer et al., 2008). Lane suggested that in order to have rapid improvement it will take the non-negotiable expectations for the organization of instruction, the development of staff skill and a deeper understanding of the conditions (e.g., collective ownership, time for collaborative problem solving) required for achieving rapid district improvement. Robinson et al. (2008) documented some of these district conditions using teacher's responses to a survey on district's goals and priorities. Teachers in some districts reported that a lack of

25

clarity regarding the district's agreed-upon priorities and subsequent resource alignment failed to support staff ability to implement EBPs. The researchers concluded, "Without clear goals, staff effort, and initiatives can be dissipated in multiple agendas and conflicting priorities, which over time, can produce burn-out, cynicism, and disengagement" (Robinson et al., 2008, p. 666). Thus, teachers reported that the policy-research-practice gap widened in some districts under study in Robinson.

Organizational systems. District policy must clearly communicate district goals and priorities to provide guidance for how the organizational systems across a district will be aligned to support personnel (routines and procedures for what will be used and when it will be used) and required funding and resources (materials, training, coaching) for implementation of effective practice in schools (Carnine, 1999; Fixsen et al., 2013). Organizational systems can be defined as the integration of outcomes, data, systems, and practices. Sugai and Horner (2002) identified these as the four critical elements required for successful implementation of RTI for PBIS. Early literacy and PBIS require the same organizational supports to ensure effective implementation: (a) policy to clearly define desired outcomes (b) on-going use of collaborative data based decision-making, (c) active administrative support to intervene in systems to break down barriers, and (d) fidelity of implementation for the chosen EBPs.

Sugai and Horner (2002) stressed the importance of systems and practices being research validated and evaluated for fidelity of implementation (Gersten et al., 2008); data guides all decisions through strong administrative support (Slavin et al., 2013; Fixsen et al., 2013). Administrative support can be identified as district and school leadership working together to support school staff (Baker et al., 2011; Fixsen et al.,

2013; Honig et al., 2010; Lane, 2009). Like PBIS, Reading First identified administrative support at the district and school level as a key variable to high quality implementation of a district's EBPs to ensure adequate funding and resource allocation (Baker et al., 2011).

Funding and resources. Policy must identify how the district will commit to adequate funding and resource allocation to ensure schools have the resources they require to meet district goals (Lane, 2009; Rorrer et al., 2008). This process requires a sustained commitment to adequate resource allocation and a declaration that the district will do things differently with district resources to develop staff skill through extensive PD and coaching (Baker et al., 2010). Sadler and Sugai (2009) highlighted the importance of district leadership demonstrating sustained fiscal (funding) and personnel support (resources and skill development). Lane agreed that rapid improvement will take a new model of accountability and a particular set of leadership skills (e.g., policy for system-wide improvement, district office resources are directed to student learning) to achieve rapid district improvement through adequate resources (e.g., curriculum and PD) to develop staff skill and high quality implementation of RTI for early literacy in every school in a district.

Staff skill. Fixsen and Blaze (2009) identified the need for a facilitative administration at the district and school level to engage in continuous evaluation of district and school data to identify barriers to implementation and deliver the resources schools require to develop staff skill. Honig et al. (2010) described a system where district leadership was responsible for developing a learning-focused partnership with every school principal to strengthen their instructional leadership skills through differentiated support. Time was allocated to sustain the on-going district-principal

partnerships. Then, on-going assessment and evaluation of principal's skills was used to determine the individualized supports principals required to build principal capacity district wide, so principals in turn could support teachers in the classroom (Honig et al., 2010).

This paradigm shift toward the development of district's skills to develop the skills of principals was timely. Historically, teachers were provided one day of PD and expected to return to the classroom and implement PD without follow-up support (Knight 2009) in the form of on-going coaching (Joyce & Showers, 2002) or principal support (Robinson et al., 2008). This process of training and hoping the training was implemented by teachers resulted in the failure of numerous comprehensive school reform initiatives. Aladjem and Borman, (2006) and Vernez, Karam, Mariano, and DeMartini (2006) found that in the first three years of implementing a new initiative fewer than half of the teachers received training, and only half of the teachers trained received any follow-up coaching. Furthermore, in the fourth and fifth year of implementation, fewer than 10% of the schools were making use of the initiative as it was designed. Consequently, most students never benefited from the initiative. Similar results were found in a study of PBIS implementation. Following staff development and a year of implementing PBIS, teachers surveyed reported, "Implementation was incomplete, short-lived, inaccurate and inconsistent if implemented at all" (Sugai & Horner, 2006, p. 249). Given the anecdotal response from school staff, one might question if district and school leaders had knowledge of the EBPs they were implementing and the skills to implement the EBPs within the unique context of every school. Together this ineffective implementation indicated that there was a need for the development of policy that aligns goals,

organizational systems and funding to focus all efforts on improving the instructional leadership skills of district (board, district and school administrators) to support principals and teachers who are responsible for delivering a district's EBPs with fidelity to benefit every student. The five district practices: (a) district policy, (b) district's vision goals and priorities, (c) organizational systems, (d) funding and resources, and (e) on-going skill development have been shown to be effective in developing district wide capacity and sustaining high quality implementation of early literacy programs district wide (Baker et al., 2011; Fixsen et al., 2013; Honig et al., 2010; Lane, 2009; Leithwood & Jantzi, 2008; Rorrer et al., 2008; Slavin et al., 2013).

History of EBISS and Early Literacy

The State of Oregon responded to the need to improve early literacy outcomes using the lessons learned from Oregon Reading First (Baker et al., 2011) and a district-wide RTI pilot in one Oregon district. The pilot, Effective Behavioral and Instructional Support (EBIS; Sadler & Sugai, 2009) was a district-wide RTI model that blended two EBPs: PBIS (Sugai & Horner, 2009) and early literacy using SWRM (Kame'enui & Simmons, 1998) that began in 1996. Note that the original pilot in 1996 was titled EBIS (one 'S') and the state wide initiative that began in 2007 added an additional 'S' to the end of the acronym, distinguishing itself from the original pilot.

In 1996, a district of approximately 12,000 students with nine elementary schools adopted PBIS in every elementary school. Initial implementation included the establishment of a continuum of behavioral systems, practices, and data where all elementary staff taught and encouraged pro-social behaviors. School leadership teams led school staff with the full support of district (school board members, superintendent

29

administrators) and policy to ensure equitable resource allocation (training, coaching) to systematically use data for on-going evaluation. In the fourth year of the pilot the school board established policy to adopt the PBIS model kindergarten through grade twelve. District resources were allocated to ensure accurate and sustainable implementation at all levels of the system (i.e., a district coordinator was hired, regular leadership meetings at the district and school level were held to analyze data for decision-making and action planning). In 2001, the district received a federal grant to integrate RTI for early literacy into the well-established RTI PBIS model.

Critical features of the pilot. In order to add a prevention and intervention system for early literacy a district task force developed guidelines for grouping, monitoring and adjusting interventions for students who were not responding to high quality reading instruction in the general education setting. Sadler and Sugai (2009) highlighted how a district should clearly communicate that EBIS was a priority by using common language and processes in all strategic planning. The district demonstrated fiscal and personnel support. They funded an early literacy coordinator, a K-5 literacy specialist, a literacy leadership team, and collaborative planning time at the school-level. They also developed decision rules for a standard reading protocol (e.g., 90 minutes of daily reading in K-5). Principals received training and on-going support for district to lead teams of teachers and specialists to develop and evaluate quality Tier 1 differentiated instruction and intervention activities (Tier 2 and 3). Fidelity standards were also developed.

Most important, was the fact that the board approved fidelity of implementation standards (Sadler & Sugai, 2009) that demonstrated alignment, coordination and

coherence of district policy, goals and priorities. The district was committed to implementation fidelity. Policy clearly articulated expectations to all staff followed by the resources and support staff required to meet district goals. As a result, school leaders and staff reported a sense of meaning and confidence in their work as they led and influenced the work in their schools.

Outcomes achieved. School leaders led the work of EBIS through careful analysis of data for problem solving. Sadler (2000) identified the importance of on-going summative and formative review of fidelity and outcome data to effectively implement EBIS. They found that students who met the Oral Reading Fluency (ORF) benchmark in the spring of first grade and remained in the district's reading program until the end of third grade, had a 98% to 99% chance of meeting the state reading benchmark at third grade (ORF is an individually administered and standardized test of accuracy and fluency with connected text). In fact, 75% of students who met the ORF benchmark at the end of first grade in 2002 exceeded the state reading benchmark at the end of third grade in 2004 (Sadler & Sugai, 2009).

Given the improved student outcomes that resulted from policy that supported high quality district wide implementation of RTI in one Oregon district, ODE questioned if this system could be scaled-up in more Oregon districts. To test this hypothesis, ODE gathered policymakers at the state level to work with experts (district staff and school leaders), university researchers and local practitioners (literacy experts, educational coaches, and specialists).

Scaling-up EBISS for Early Literacy

Building on the knowledge gained from implementation of the EBIS pilot in one Oregon district, a coalition of influential decision-makers (Carnine, 1999) came together in 2006 to design the five-year statewide EBISS initiative. Experts in the field developed a coordinated set of PD activities (see Table 10) to teach district leadership teams how to align systems and practices at all levels of the district in order to blend PBIS and early literacy.

Understanding how an initiative was developed is an important component of program evaluation when determining whether a social intervention is producing the intended result (Babbie, 2007). According to Babbie (2007), "we must be able to operationalize, observe and recognize the presence or absence of what is under study (p. 352). Identifying the implementation components of EBISS was an important component of this evaluation research, as the cost of the EBISS initiative was weighed against what it returned in benefits (gains in reading scores on the OAKS at third grade). The implementation components of EBISS are outlined below and give the reader a deeper understanding of why and how they were developed in an attempt for ODE to meet its implementation goal to increase student outcomes in EBISS districts.

EBISS development. Given the history of failed change (Aladjem & Borman 2006) it was clear to educational leaders in the State of Oregon that district and school leadership must be re-conceptualized and shift from leaders as managers to leaders as instructional change agents (Fixsen et al., 2013; Honig et al., 2010; Lane, 2009; Sadler & Sugai, 2009). In order to accomplish this goal a coordinated system of PD and TA was established. It is important to note that the initial focus of EBISS was on universal Tier 1

systems and practices (Gersten et al., 2008). EBISS was intentional in its focus on high quality Tier 1 instruction for all students in an effort to simultaneously decrease the number of students requiring targeted and individualized supports. This position was supported by research conducted in a medium sized district in Oregon (McIntosh et al., 2006). McIntosh et al. found that behavior and literacy interventions provided to all students at Tier 1, may reduce the need for resource dependent supplemental supports (Tier 2) and individualized supports (Tier 3). Armed with current research ODE proceeded with plans to scale-up the pilot led by Sadler and Sugai (2009).

In 2006, ODE partnered with university researchers and district level practitioners to build upon the current research that suggested implementing early literacy through the use of a common RTI framework can increase the likelihood of scaling-up high quality implementation across a district's schools (Fuchs & Fuchs, 2006; Sadler & Sugai, 2009). Equipped with lessons learned, ODE and its partners planned the development of a district wide framework to implement the RTI model. Fuchs and Fuchs (2006) provided evidence to suggest that a coordinated set of district protocols could be scaled-up district wide, because as Tucker, Edmondson, and Nembhard (2005) suggested, everyone would know what (knowledge of the RTI for literacy) and how (knowledge of how to implement the RTI for literacy) making fidelity easier to assess.

In preparation, ODE contracted with the Center on Teaching and Learning (CTL) at the University of Oregon to coordinate the EBISS initiative. CTL hired a team of experts (known as EBISS State Coordinators) in the field of literacy to provide PD and TA to districts in Oregon. The goal of ODE was to improve student outcomes through on-going PD and TA focused at the district leadership level. Districts in Oregon were

33

invited to apply. Twenty-seven districts joined the EBISS initiative in 2007. Across the five-year initiative, only two districts withdrew, resulting in 25 districts participating in the current study.

Critical Features of EBISS. As we know, Sadler and Sugai (2009) found that after ten years of implementing RTI for early literacy schools were experiencing a sustained increase in the number of students meeting early reading benchmarks at third grade (as well as a reduction in office discipline referrals). Consequently, the critical components of the Sadler and Sugai's model were conceptualized into the EBISS Teaming Framework.

The EBISS teaming framework. The EBISS teaming framework (see Figure 2) identified the key participants who actively participated on the district and school leadership teams, as well as teacher teams at the grade, content, and individual student level. The teams, distributed across a district's system, were developed to ensure effective communication and the development of a continuum of supports for improving student's reading skills and state assessment scores (EBISS Systems Coaching Manual: Chaparro, Ryan Jackson et al., 2012). Spillane et al. (2001) identified this as a system of distributed leadership where administrators and teacher leaders share the management of schools through strong instructional leadership. Recent research proposed that there was a need for a teaming framework that demonstrated a strong visual of communication loops and improvement cycles between district and its schools (Fixsen et al., 2013; Honig et al., 2010; Lane, 2009).
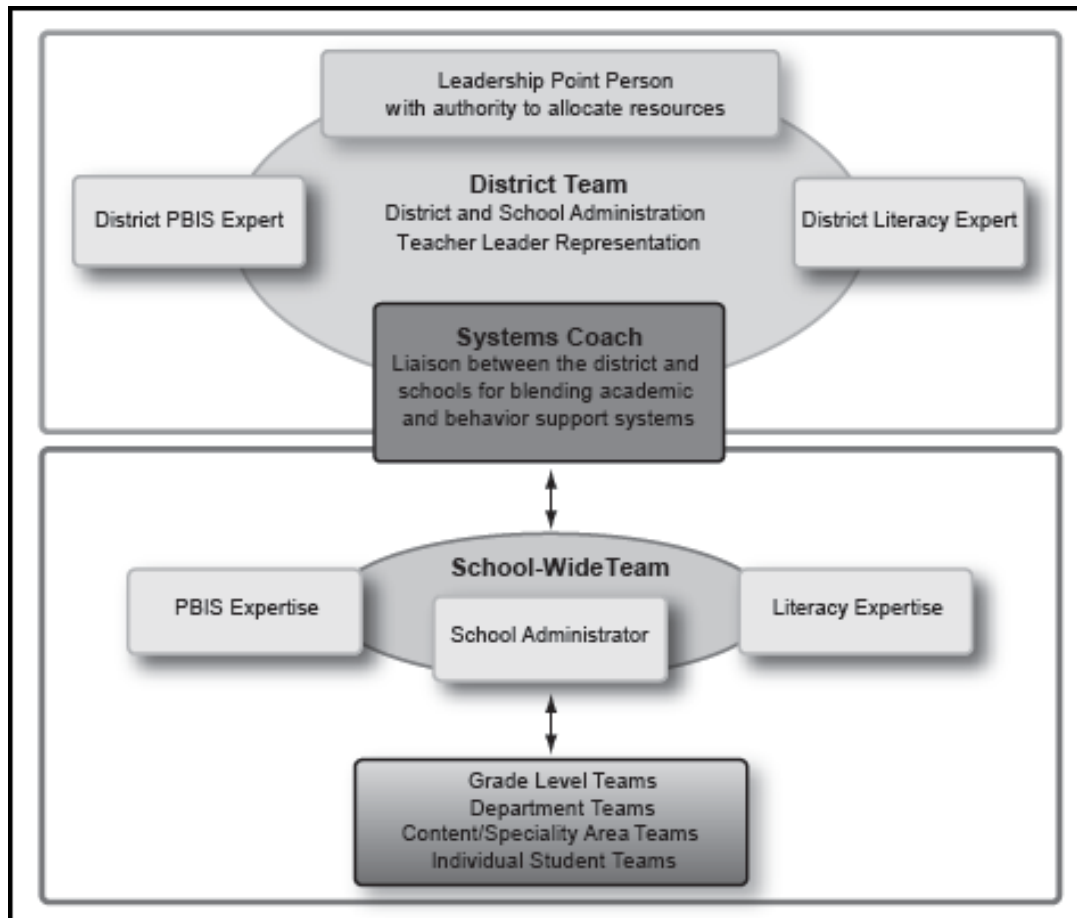
Figure 2

EBISS teaming framework

   The district leadership team. Development of the district leadership team was

found to be the first step in the advancement of high quality implementation of any EBP

(Fixsen et al., 2013). The team ensures that district policy, goals, and priorities are clear

and that policy will support implementation efforts through adequate funding and direct

support to schools. This guarantees training, coaching, and instructional materials are

available to increase staff competency and build district capacity (Fixsen et al., 2013;

Sadler & Sugai, 2009). Research suggested that district administrators who simply fill the

role of managers and purveyors of supplies fail to support the instructional needs of

35

principals and teachers who are required to implement EBPs and improve student

outcomes (Honig et al., 2010).

The school-wide leadership team. Of equal importance was the development of a

strong leadership team at each school. Sugai and Sadler (2009) found data based

decision-making teams at the school level were critical to a school's high quality

implementation of a prevention model. There was also evidence to suggest that this

process developed teachers as leaders (Spillane et al., 2001). Spillane et al. found that as

district support was increased and sustained, teacher's trust in district was increased, and

teachers showed willingness to improve practice as they learned from one another

(Robinson et al., 2008; Slavin et al., 2013). EBISS also stressed the importance of a

school principal actively participating in school and grade level team meetings. Robinson

et al. (2008) and Putnam (2008) concur as they found this to be critical to effective

problem solving team meetings at the school level.

The systems coach. At the center of the framework was the pivotal position of the

systems coach. The role of the systems coach was to facilitate communication between

the district and its schools. Systems coaches supported school leadership teams in the

development of school's goals and action plans. The systems coach also informed the

district of the successes and challenges schools faced, so the district could respond and

provide the supports schools required to meet district goals (through effective policy).

The role and function of the systems coach draws a strong comparison to the district

instructional coach, Swinnerton (2007) found instrumental to the development of

communication loops and improvement cycles between a district and its schools. In order

to support the district and school's development and on-going evaluation of their action

plans, there must be a way to measure a district's implementation efforts so everyone can use data to continually improve practice.

>*District self-assessment.* EBISS participants needed a way to test the strength of their district systems. So, expert practitioners and researchers came together to develop a district level self-assessment know as the District System Support Plan (DSSP: Chaparro, Park, Baker, & Ryan Jackson, 2011). See the Methods section for a complete description of the DSSP. District leadership (e.g., superintendent, district and school leadership, specialists) completed the DSSP each year and developed action plans based on district and school capacity and readiness (will and commitment) for change (Rorrer et al., 2008). When district's implementation data moved in a positive direction it was an indicator that the PD and TA that was focused at the district level made a difference in schools' implementation efforts. In other words, if district put the organizational systems in place (e.g., data system, time, materials, training, coaching) to implement the critical components of RTI in each school (teams, tiered model, data system, action plan) student's reading outcomes would improve by third grade. This can be conceptualized as *the links in the chain* that join strong leadership practices with student learning (Leithwood & Jantzi, 2008).

>Figure 3 conceptualizes how EBISS PD and TA, focused at the district level, had to flow through all levels of the district system to *link* district practices with student learning. The figure identifies: (a) who was responsible for PD and leadership, (b) what their purpose was, (c) how they implemented systematic reform, (d) the measures used to assess fidelity, and (e) the assumptions that underlie how EBISS PD and TA must be replicated and transferred to staff at each level of the system through guidance and direct

37

Figure 3. District Model: Effective Behavioral And Instructional Support Systems

WHO	WHAT	HOW	MEASURES	ASSUMPTIONS	
Department of Education and EBISS State Coordinators	EBISS Coordinators Two days of training per year to *District Leadership Teams (DLT)*. • Systems and practices at the district level for schools to implement RTI for Reading and Behavior One Technical assistance visit per year and ongoing telephone consultation • Training Webinars in year 3,4,5 for use beyond the grant	*Develop DLT Capacity* • Assesse systems and practices district-wide • Develop DLT skills to use data to develop solutions and action plans for on-going evaluation • Engage stakeholders, distribute leadership • Partner with schools, understand barriers, successes • Provide resources, e.g., training, coaching • Establish policy	*District System Support Plan (DSSP). DLT* Assesses • Leadership & Commitment • Plan for each schools • Assessment Systems • Coaching • Training • Stakeholder Support • Funding	*DLT...* • Attends Trainings • Communicates commitment • Engages multiple stakeholders • Establishes policy • Provides resources • Develops instructional leadership skills (ILS) to support schools	*Independent Variable — Professional Development & Technical Assistance*
District Leadership Teams (DLT)	*DLT* Trains *School Leadership Teams (SLT)* • Systems and practices at the school level required for schools to implement RTI • Systems Coach/s attends and provides coaching to school leadership teams and teacher teams	Develop *SLT* Capacity • Assess systems and practices at the school level. • Develop principal and school leaders skills to use data to problem solve, develop solutions, and action plans for on-going evaluation	*SLT* develops a school action plan based on the district's priorities and goals (*DSSP*) • *STL* identifies the resources supports they will require from the DLT	*DLT* has the ILS and commitment to work with each *SLT* • The school plan responds to − Student need − Staff skill − Resources required	*Did We Do What We Said We Would Do...And Is It Working?*
School Leadership Team (SLT)	*SLT* Trains *School Teams (ST)* e.g., reading, PBIS, teacher teams • Systems and practices required in each classroom for RTI • Systems Coach/s attend teacher and individual student team meetings	Develop *ST* Capacity • Assess RTI systems and practices • Develop teacher's skills to implement RTI with fidelity and use data action planning and on-going evaluation	*School Level Data Summary (SLDS) yearly* • Literacy and behavior team • Screening and progress monitoring • Action Plan	*SLT* has the ILS and commitment to work with each *ST* • School plan responds to needs − Staff Skill − Resources To Support Student Need	
Teachers	*ST*s work collaboratively with teachers and paraprofessionals to prevent academic and behavioral problems for *ALL* students	Develop *teacher capacity* • Differentiated core, targeted, individual supports, data based problem solving model	*DLT, SLT* and *ST*s collaboratively measure fidelity and raise everyone's skills to serve students	*DLT & SLT* have the skill and commitment to allocate time to coach school leaders and staff	
Students	*DLT, SLT, ST*s, teachers and school staff develop the instructional and socio-behavioral skills to support the needs *ALL* students and their families	*DLT, SLT* and *ST*s set priorities, define outcome goals, data, systems and practices focused on meeting the needs and engaging *ALL* students	Student Outcome and Practice Fidelity Measures for on-going evaluation	*DLT, SLT, ST*s and Teachers develop ILS, and collective commitment to support ALL student's needs	*Dependent Variable — State Assessment*

Figure 3

EBISS district model. District Leadership includes: board, district and school administrators, specialists, teacher leaders based on district priorities and culture. RTI is defined as a multi-tiered prevention system that implements screening and progress monitoring systems, collaborative data based problem solving and action planning to ensure all students receive the supports they require (NCRTI, 2013).

support from district. Note, at the top right of the figure is the independent variable for this study (EBISS PD and TA) and at the bottom right is the dependent variable (OAKS third grade reading assessment).

EBISS Outcomes. Results from the first two years of EBISS implementation suggest the potential for further development of the EBISS model as a promising practice to strengthen district leadership and improve student outcomes (Chaparro, Smolkowski, et al., 2012). District leadership teams met at least once a year to review their DSSP data, evaluate progress, and revise the district literacy action plan as needed based on student outcome data (e.g., OAKS third grade data, screening and progress monitoring data, office discipline referrals, and attendance). Analysis identified the means of students who *exceeded*, *met* or *did not meet* the OAKS at third grade: The average percent of students who met the OAKS benchmark increased by 3.5% in third grade during year two.

EBISS as a Promising Practice

EBISS implementation and outcome data provided evidence for its promise as an effective district model to implement RTI for early literacy (Chaparro, Smolkowski, et al., 2012). EBISS focused on the development of staff skill (at all levels of the system), through aligned organizational systems that were supported by district policy and implemented through the use of a common district framework (Chaparro, Ryan Jackson, et al., 2012; Sadler & Sugai, 2009). Given the converging research on the importance of district policy to support the effective practices of principals and teachers in schools (Chaparro, Ryan Jackson et al., 2012; Honig et al., 2010; Lane, 2009; Sadler & Sugai, 2009; Slavin et al., 2013; Rorrer et al., 2008) one can appreciate the recent attention of federal policy that calls for district and school accountability measures that improve

principal and teacher effectiveness (ESEA, 2010). This was a fundamental shift from placing the onus on the teacher to implement an EBP to placing the onus on school, district and state leadership to support high quality implementation efforts of teachers in the classroom (Fixsen et al., 2013).

Knapp (2008) talks about the recent attention to the role of district in improving teaching and learning, stating, "a problem of system wide leadership practice is begging to be better understood" (p. 522). Although it is true that we have a great deal of work to do, districts are responding to the challenge of improving teaching and learning through the development of strong leadership models where district policy identifies implementation guidelines and delivers the fiscal and personnel support staff require to meet new accountability measures (Chaparro, Smolkowski, et al., 2012; Lane, 2009; Slavin et al., 2013; Sadler & Sugai, 2009). The organizational systems applied in all of the aforementioned studies focused on the use of EBPs in a common district framework to align and coordinate district policy and resources to provide resources to schools (e.g., curriculum, release time) and direct support to develop staff skill. All critical features of the EBISS model.

Research Summary

High quality implementation of an evidence based district wide reading program is an intricate process that requires careful development of organizational structures at the district and school level to ensure principals and teachers receive the resources and supports they need, from district, to develop the skills required to implement and sustain EBPs in the classroom (Chaparro, Smolkowski et al., 2012; Gersten et al., 2008; Sadler & Sugai, 2009; Stockslager et al., 2013). It is clear that emerging reform models

demonstrate the importance of linking five critical practices of district office leadership to support the effective implementation of successful reading programs in schools: (a) policy, (b) goals and priorities, (c) organizational systems, (d) funding and resources, and (e) staff skill development. Communication of clear expectations, direct support for schools, and collaborative data based decision-making are common threads found among all district reforms models that resulted in effective and efficient problem solving and action planning to improve student outcomes (Chaparro, Smolkowski, et al., 2012; Lane, 2009; Honig et al., 2010; Sadler & Sugai, 2009; Rorrer et al., 2008; Slavin et al., 2013). Yet, the research indicates that changing a districts organizational culture and marshaling a collective commitment to sustain change over time is a difficult task (Rorrer et al., 2008; Stockslager et al., 2013).

Figure 4 provides a strong visual of how the critical features of EBISS TA and PD must be transferred from EBISS coordinators to district leadership. Then, district leadership must transfer their new knowledge to actively support school leadership teams who simultaneously support the individual teachers who ultimately improve student outcomes (see Figure 3 for a detailed description). Instructional leadership skills must flow from the district, through school leadership to teachers in the classroom. In turn, valuable classroom and school-level implementation data (quantitative and qualitative data for improvement cycles) must flow back to the district from teachers (communication loops), so all stakeholders have a voice in the identification of needs and barriers to implementation. This process can ensure that policy supports teacher's implementation efforts and the needs of every student in the district.

Research Questions

My EBISS program evaluation aimed to answer two questions. The first research question asked was there a difference between non-EBISS and EBISS district's students' gains in the reading sub-tests of the OAKS at third grade from time one to time two? The second research question asked were students' gains in the reading sub-tests of the OAKS at third grade from time one to time two different for EBISS districts with high DSSP scores versus low DSSP scores? Answering these research questions will inform future efforts to implement EBISS for early literacy in a common RTI framework to improve 3rd grade reading scores on the OAKS.

Given what is known about the importance of district will and commitment (Rorrer et al. 2008) the question remains did the focus on changing district leadership behaviors improve student reading outcomes on the OAKS at third grade in EBISS districts, over and above a matched comparison sample of non-EBISS districts. The main purpose of this evaluation study is to reject the null hypothesis for research question number one: there is no relationship between EBISS PD and TA and the percent of 3rd grade students passing the OAKS reading assessment from 2006 (EBISS pre-implementation) to 2011 (EBISS post implementation) when EBISS districts are compared to a matched sample of non-EBISS districts. The second purpose of this evaluation study is to reject the null hypothesis for research question number two: there is no relationship between EBISS PD and TA and the percent of 3rd grade students passing the OAKS reading assessment for low-EBISS and high EBISS districts.
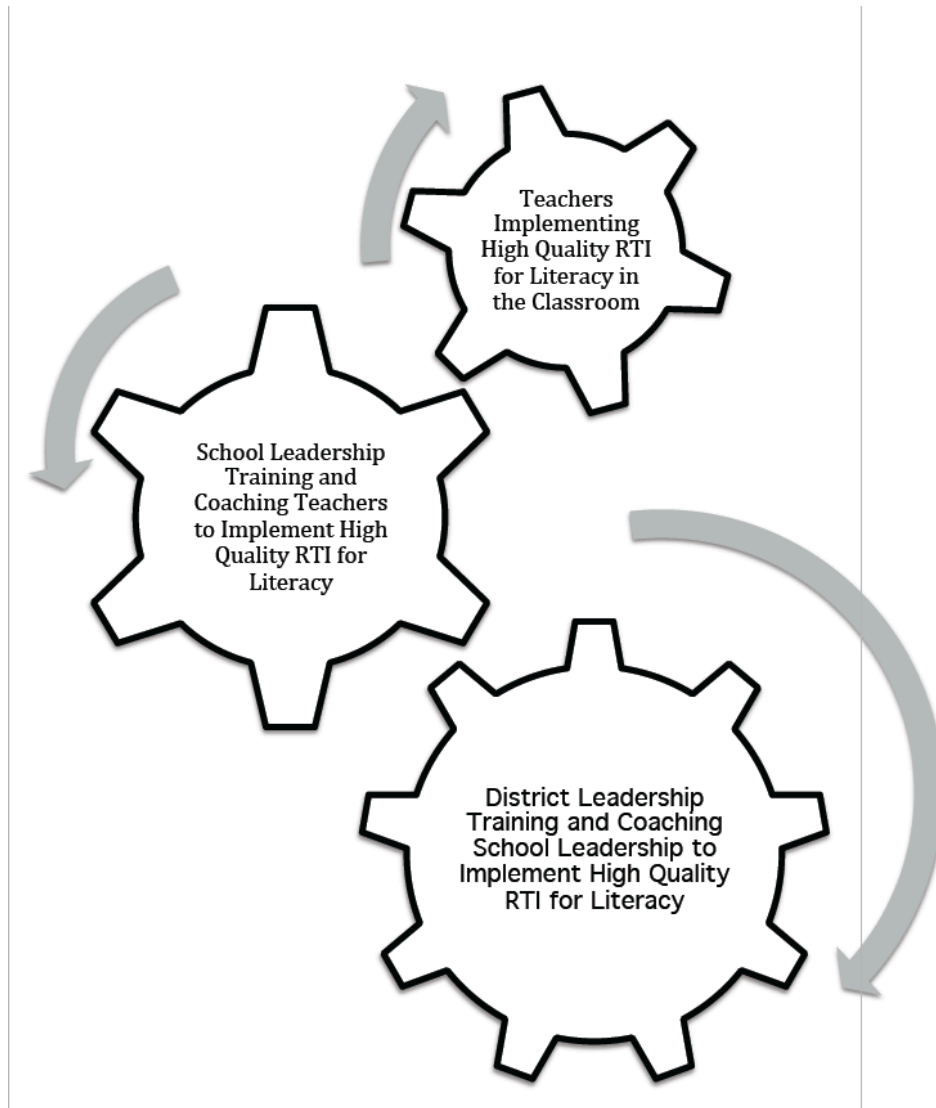
Figure 4

EBISS professional development model. RTI is defined as a multi-tiered prevention system that implements screening and progress monitoring systems to engage in collaborative data based problem solving and action planning to ensure all students receive the supports they require (NCRTI, 2013).

CHAPTER II

METHODS

The purpose of this study was to evaluate the Effective Behavioral and Instructional Support Systems (EBISS) initiative in the State of Oregon. The primary question concerns whether the mean EBISS district's performance was significantly different than non-EBISS districts as measured by the percent of students passing the third grade reading tests on the Oregon Assessment of Knowledge and Skills (OAKS) from time one (T_1) 2006-07 (EBISS pre-treatment year) to time two (T_2) 2010-11 (EBISS post-treatment year). The secondary question focuses on differences within EBISS districts and asks if EBISS districts with high implementation scores made statistically significantly greater gains than EBISS districts with low implementation scores on the percent of students who passed the third grade OAKS reading assessment from to 2007 (first year of administration) to 2011 (post administration). A pre-post quasi-experimental comparison on a matched sample was used to evaluate the four-year EBISS initiative to compare the gains in student performance on the OAKS reading subtests at third grade between non-EBISS and EBISS districts and within EBISS districts.

Design

OAKS reading data was analyzed to conduct an evaluation of the EBISS initiative. A pre-post matched-sample quasi-experimental design was used to evaluate the four-year EBISS initiative. EBISS districts were matched to non-EBISS districts by important district variables. Those two samples were used to determine if student performance significantly differed between non-EBISS and EBISS districts on gains in the percent of student's passing the third-grade OAKS reading assessment from T_1 to T_2.

A second research question was developed to evaluate differences within EBISS districts. EBISS districts with high and low District Systems Support Plan (DSSP) scores were compared on their percentage of student's gains on the OAKS at third grade from T_1 to T_2.

Participants

Twenty-five Non-EBISS districts were matched to 25 EBISS districts by district demographic variables from a population of 197 districts in the State of Oregon (for a complete description of the matched demographic variables see the matching section of this manuscript). In 2006, 25 EBISS districts consisted of 182 elementary schools and 25 non-EBISS districts consisted of 176 elementary schools. Of the original 27 districts that began participation in the EBISS initiative in 2006, two exited. One exited in year two and the other in year four. This attrition resulted in a total of 25 EBISS districts in the State of Oregon who participated in the initiative from $T_{1 \text{ to}} T_2$ for analysis in the current study.

Procedure for Identifying the Matched Sample

Oregon Department of Education (ODE) employees gathered all matching data for this study into an Excel spreadsheet. Data represented the total population of 197 Oregon districts. The 25 EBISS districts were extracted from the total population of Oregon districts and entered into IBM SPSS software. Decision rules were developed that identified the order in which the variables would be entered into SPSS and matched: (a) region, (b) enrollment, (c) Free and Reduced Meals (FARM), (d) white and non-white, (e) male and female, (f) years of teaching experience, and (g) 3rd grade OAKS scores. Then, the remaining 172 non-EBISS districts were entered into SPSS by region and

enrollment (then FARM, white and non-white, male and female, years of teaching experience, OAKS scores). Of the 172 non-EBISS districts in the sample, 40 with an enrollment of under 173 students were deleted (EBISS districts were all well above 173 students) leaving 15 small districts to select four small non-EBISS districts by enrollment (enrollment 205-473). Of the 132 remaining districts, 25 non-EBISS districts were identified and matched to EBISS districts by enrollment (within each region). It was at this point that non-EBISS districts ordered themselves and were matched by FARM, white and non-white, male and female, years of teaching experience, and OAKS scores with EBISS districts. The decision rules for matching were violated on a few counts.

The order of FARM and then white and non-white was reversed in two districts in the central and eastern regions in order to match a district within each region by enrollment. In the southern region, there was only one large district, so a comparable match was not available. The decision was made to stay within the region (south) and select a match from the medium sized districts in the region (however both districts fell into the large ODE classification). The total number of non-EBISS and EBISS districts in each group were equal (i.e., total valley is 30, with 15 EBISS and 15 non-EBISS districts in this region). Matching all Oregon districts by region first and then by size made it very easy to match the student and teacher variables for non-EBISS and EBISS districts. The important variables that define non-EBISS and EBISS districts in this study were identified as important through the literature and include: (a) region of the state, (b) enrollment, (c) percent of student's eligible for FARM, (d) white and non-white, (e) male and female, (f) average teacher's years of experience, and (g) 3rd grade OAKS scores.

(Mathis, 2003; Siran, 2005; Foster & Miller, 2007; Flynn & Rahbar, 2006; Darling-Hammond (1999).

Matching Variables

This matched comparison sample was chosen to minimize selection bias, to the extent possible within this quasi-experimental study, while testing differences between non-EBISS and EBISS districts. The groups were matched by (a) region, (b) enrollment, (c) FARM, (d) white and non-white, (e) male and female, (f) years of teacher's experience, and (g) 3rd grade OAKS scores. Matching Variables were selected with guidance from Shadish et al. (2002), who recommend variables that are similar across conditions (i.e., overlapping distributions), stable, and reliable. In a similar study, Tuttle et al. (2010) evaluated Knowledge is Power Program (KIPP) schools by identifying a matched sample of non-KIPP schools post hoc to compare gains on student's scores on state assessment tests. The same matching variables used in the Tuttle et al. study were applied to the current study and are explained in depth in the next section of this manuscript. First, research on region was considered.

Region. Consideration of the region within a state where a district is located is important due to the vast difference in the population of a locale. The National Center for Educational Statistics (NCES) identifies Locale Codes (2011: http://nces.ed.gov/ccd/rural_locales.asp) that were re-designed by the Census Bureau in 2005-06 to more accurately describe the United States population. For the purpose of this study two terms were used to classify the study population; rural and urban-suburban. Rural is defined by NCES using Locale Codes (or any territory outside of a large or mid-size city). The codes used in this study to define the rural districts are: fringe, distant, and

47

remote. Urban-suburban is also defined by NCES using Local Codes. The coding classification of mid-sized city (population less than 250,000) and small town (greater than 2,500) was used in this study to classify the urban-suburban districts. Research suggests that it is important to consider the region a district is located in when developing matched samples.

The research of Mathis (2003) identified multiple challenges that rural school districts face as opposed to their urban-suburban counterparts. He identified difficulties in recruiting rural teachers, lack of services in isolated areas, inability of rural districts to absorb special education and transportation costs, plus obsolete and rundown rural facilities. Given the differences that can exist between rural and urban-suburban districts, non-EBISS and EBISS districts were matched by geographic regions in the State of Oregon (http://www.oregon.gov/ODOT/pages/highwayregions.aspx). The Oregon Department of Transportation identifies the five regions as: (a) Portland Metro, (b) Willamette Valley, North and mid-coast, (c) Southern Oregon and south coast, (d) Central Oregon, and (e) Eastern Oregon. For the purpose of this study districts were classified as (a) Valley (including Portland metro), (b) South (not south coast), (c) Coast to include north, central, and southern coast, (d) Central Oregon, and (e) Eastern Oregon. Portland metro districts were included in the Willamette Valley and classified as Valley due to the entire areas access to universities and social services. The coastal regions of Oregon were assigned their own geographic classification and separated from the Willamette Valley and Southern Oregon due to the entire areas lack of access to universities and social services. The region a district is located in is important; therefor the percentage of EBISS districts within each region was identified.

48

The majority of the 50 districts in the sample were located in the valley, south and central part of the state (76 percent) in a suburban-urban setting where the population is more concentrated and the citizens have access to services (universities, health services, etc.). Twenty-four percent of the districts were located in the rural areas with less access to large universities and health services. Rural areas were located in coast and eastern regions of the state. Table 2 identifies the percent of districts in each region for non-EBISS and EBISS. There was no statistical mean difference between non-EBISS and EBISS districts for region ($t = 0.000$, $df = 48$, $p = 1.000$).

Table 2

2006-07 Districts by Region: Counts and Frequencies

Region	2006-07 Frequency	2006-07 Percent
Willamette Valley	30	60
South	4	8
Coast	6	12
Central	4	8
East	6	12

Enrollment. The enrollment or number of students enrolled in a district can effect the successful implementation of reform initiatives**.** Hannaway and Kimbal, (1998) conducted a national survey of 2,700 school districts to assess progress in standards-based reform. Enrollment was noted as a factor contributing to successful implementation of reform initiatives. They found that large districts have additional resources produced

through economies of scale (Hannaway & Kimbal, 1998). Putnam (2008) also found that very small districts in the State of Oregon participating in the Oregon Response to Intervention (OrRTI) state initiative often have limited resources that results in failing to meet thresholds that would make them eligible for certain supports or resources.

Given the aforementioned research, enrollment was identified as the second matching variable. ODE classifies district enrollment into 3 groups: (a) small (less than 500), (b) medium (501-1200), and (c) large (over 1201) (ODE, 2007). In order to increase the number of enrollment groups from 3 to 6 this study classified districts as: (a) Small (100-499), (b) Small-medium (500-1299), (c) Medium (1,300-5,999), (d) Medium-large (6,000-11,999), (e) Large 12,000-20,999), and (f) Very large (21,000 +). Expanding the classification to 6 groups made it easier to perform the next step: Matching student and teacher characteristics within each region by enrollment. Table 3 provides a count and percent of the number of districts in each enrollment classification (equal counts for non-EBISS and EBISS are displayed for each classification). There was no statistical mean difference between EBISS and non-EBISS districts for enrollment ($t = -0.121$, $df = 48$, $p = .904$).

Free and reduced meals. FARM was identified as a variable to investigate for matching in this study, because eligibility for FARM is determined by a family's socioeconomic status (SES) and it can affect the resources a family has to support their children in school (Siran, 2005). A student qualifies for FARM based on their family's income or SES. Friedkin and Nocochea (1988) suggest that a relationship exists between district size and student achievement and that it is dependent on the SES variable of students in the district. Conversely, in 2003, Hanushek, Kain, Markman and Rivkin

(2003) analyzed Texas State Assessment scores for grades three through six beginning in

1993 and submit that there is little evidence that SES effects growth in achievement.

However, they qualify their conclusion given the fact that SES is a *noisy measure* (e.g., a

number of variables can effect the measurement of SES). In order to examine the

complex relationship between SES and student achievement the seminal study of White

(1982) was replicated. Using White's procedures, Siran (2005) conducted a meta-analysis

that included 128 school districts and found that the relationship between SES and

academic achievement at the student level had a medium effect (or the two variables

seemed to be related). Interestingly, a large effect was found between SES and academic

achievement at the school level.

Table 3

2006-07 District's Characteristics by Enrollment: Counts and Frequencies

Enrollment	2006-07 Frequency	2006-07 Percent
Small (100-499)	4	8
Small-medium (500-1299)	6	12.5
Medium (1300-5999)	19	38
Medium-large (6000-11999)	13	26
Large 12,000-20999)	7	14
Very large (21000 +)	1	2

My study investigated if there was a difference in achievement for non-EBISS and EBISS districts at the elementary level in reading. Therefore, given the relationship between SES and academic achievement, FARM was identified as an important variable to match. Table 4 provides the means and standard deviations for FARM. There was no statistical mean difference between EBISS and non-EBISS districts for FARM ($t = -0.559$, $df = 48$, $p = .579$).

Table 4

Mean and Standard Deviation for non-EBISS and EBISS on Free and Reduced Meals

Group	n	Mean	SD
Non-EBISS	25	.44	.12
EBISS	25	.46	.13

Non-white and white. Foster (2001) documented that students from different racial and ethnic groups with similar SES entered kindergarten with language knowledge scores that were not significantly different. In 2007, Foster and Miller reported the results of a four-year longitudinal study on 12,621 students (kindergarten through third grade) that were divided into four groups: whites, African Americans, Hispanics and other minorities. They tracked the growth of the student's literacy skills over the course of the study and found the racial and ethnic variable did not predict academic achievement. In conclusion, the authors propose that school literacy readiness is a function of SES status and not racial or ethnic identity.

While Foster and Miller (2007) did not support non-white and white as being a

salient variable for matching, we know that Oregon students from families who are non-white have a lower mean income than white families (http://quickfacts.census.gov/qfd/states/41000.html). Therefore, the non-white and white variable was identified as an important variable to match. Table 5 provides the means and standard deviations for non-white and white participants. There was no statistical mean difference between white and non-white students in EBISS and non-EBISS districts ($t = 0.073$, $df = 48$, $p = .942$). See Table 5.

Table 5

Mean and Standard Deviation for non-EBISS and EBISS on non-white and white

Group	n	Mean	SD
Non-EBISS	25	.76	.13
EBISS	25	.75	.14

Male and female. Flynn and Rahbar (1994) tested the hypothesis that proportionately more boys than girls experience reading failure at the elementary school level. A sample of 708 boys and girls were compared. Those authors examined the difference in male and female scores on a standardized reading test for severely reading disabled students and found the ratio of boys to girls was 1.3:1 at third grade. In a study of 865 children from the age of seven to ten, Eccles, Wigfield, Harold and Blumenfeld (1993) found that boys had more positive beliefs about their competence in mathematics, whereas girls had more positive beliefs about their competence for reading.

Ensuring that districts had a similar percentage of males and females was important when comparing the reading scores of non-EBISS and EBISS districts on third

grade reading scores. Given the research gender was matched in EBISS and non-EBISS districts. Table 6 provides the means and standard deviations for male and female participants. There was no statistical mean difference between the percent of male and female students in either group ($t= 0.106$, $df = 48$, $p = .916$). See Table 6.

Table 6

Mean and Standard Deviation non-EBISS and EBISS on male and female

Group	n	Mean	SD
Non-EBISS	25	.516	.01
EBISS	25	.516	.02

Years of teacher's experience. Darling-Hammond (1999) analyzed data from a 50-state policy survey and found that policy investments in the quality of teachers may be related to improvements in student performance. For example, in a controlled study of middle school mathematics teachers, matched by years of experience, students of fully certified mathematics teachers experienced significantly larger gains in achievement than those taught by teachers not certified in mathematics (Hawk, Coble, & Swanson, 1985). In addition, Rockoff (2004) found that teacher quality is a key factor in raising elementary reading achievement when analyzing the standardized reading scores of 10,000 students in kindergarten through fifth grade and years of teaching experience for 300 teachers in the State of New Jersey.

Given the relationship between years of teaching experience and student achievement this study matched EBISS and non-EBISS on years of teaching experience. Table 7 provides the means and standard deviations for years of teachers' experience.

54

There was no statistical mean difference between years of teaching experience in either group (t= -0.490, df = .48, p = .626). See Table 7.

Table 7

Mean and Standard Deviation non-EBISS and EBISS on years of teacher's experience

Group	n	Mean	SD
Non-EBISS	25	11.98	1.96
EBISS	25	12.27	2.30

OAKS reading 2006 differences. While matching for non-academic variables was important, it was imperative that 2006 student reading achievement between the two groups not be significantly different if a comparison over time was to be made. Thus, a t-test was run to determine if a significant difference existed on the percentage of students passing the OAKS (See Table 8 and see arcsine transformation later in the chapter for complete translation of percent into equal interval data). Table 8 provides the means and standard deviations for the percent of students passing the OAKS reading assessment in 2006. There was no statistical mean difference for 2006 reading between EBISS and non-EBISS districts (t= .617, df = .48, p = .540).

Table 8

Mean and Standard Deviation non-EBISS and EBISS on Percent Passing 2006 OAKS

Group	n	Mean	SD
Non-EBISS	25	66.77	6.51
EBISS	25	65.75	5.13

Professional Development

In addition to matching the two groups on district variables, non-EBISS professional development that was available to all districts in the state was investigated. Understanding the professional development that non-EBISS and EBISS districts had access to is an important component of evaluation research when determining whether the investment in professional development produced the intended result (Babbie, 2007). The professional development provided to non-EBISS and EBISS districts through Oregon's major initiatives (other than EBISS), and local educational service districts (ESDs) are outlined below and give the reader a deeper understanding of the types of professional development provided to both groups in Oregon.

Non-EBISS professional development. An investigation of the professional development and technical assistance that non-EBISS districts received from other Oregon initiatives was conducted. To investigate non-EBISS professional development information was gathered in an Excel spreadsheet and recorded the percent of district's that participated in other major Oregon initiatives from 2007-2011. The initiatives include: (a) Direct Access To Achievement Project (DATA Project: http://www.oregondataproject.org/); (b) Creative Leadership Achieves Student Success

(CLASS: http://chalkboardproject.org/); (c) OrRTI (http://oregonrti.org/) and, (d) Positive

Behavioral Interventions and Supports (PBIS:http://www.pbis.org/). Below is a

description of each initiative.

       DATA Project. The DATA Project is one of Oregon's major statewide initiatives

aimed at improving student achievement through a school's data teaming process (ODE:

2007-2011). As of August 2011, 142 districts and 19 ESD's in the State of Oregon

participated in DATA Project trainings. Project coordinators trained teams to analyze and

use data to inform instructional decisions. Schools volunteered to participate and it was

not unusual for some, but not all schools in a district to participate. The DATA Project

offered data based decision-making training to schools that was similar to the training

EBISS districts received and both aligned with the research findings of Slavin et al.

(2013) who recommended that it was not the analysis of data that mattered but a team's

ability to use the data to make sound instructional decisions to improve student

achievement. Of the 25 non-EBISS districts in this study, 10, or 40% of the districts

participated in the DATA Project for at least one year between 2008 and 2011. Another

major initiative in Oregon is Project CLASS.

       CLASS Project. Twenty-three school districts that served nearly 159,000 students

in the State of Oregon participated in the Project CLASS from 2007 to 2011

(http://chalkboardproject.org). Of the 25 non-EBISS districts in this study, 6 districts, or

24% of the districts participated in Project CLASS for at least one year between 2008 and

2011. While this does not represent a large percentage of the non-EBISS group, the

training that districts received was similar to EBISS in some aspects. Training is focused

on teachers distributing their leadership across the organization (a key component in the

57

EBISS teaming framework) through the development of effective teaching and relevant professional development to improve student outcomes. OrRTI was also focused on relevant professional development and like EBISS it was focused on RTI for literacy.

OrRTI. The ODE provided districts with the opportunity to participate in a statewide RTI initiative called OrRTI (www.oregonrti.org). Districts were trained on the critical components of a tiered model for reading: (a) multi-level prevention system, (b) a screening and progress monitoring system, and (c) collaborative data based decision-making. The professional development and technical assistance provided to OrRTI districts was extremely similar to the professional development provided to EBISS districts. The only difference is that EBISS was focused at the district leadership level and blends RTI for reading and PBIS. Of the 25 non-EBISS districts in this study, 10 districts, or 40% of the districts participated in OrRTI for at least one year between 2008 and 2011.

PBIS. The ODE identified PBIS as one of Oregon's major initiatives (Sugai & Horner, 2009). Of the 25 non-EBISS districts in this study, 25 districts, or 100% of the districts participated in some type of PBIS training at least once between 2008 and 2011. Since the 1990s, districts in Oregon have had access to high quality professional development to implement PBIS from experts at the University of Oregon, as well as highly trained PBIS coaches in local ESDs. PBIS identifies the same critical implementation components that are used in implementing RTI for reading (Sugai & Horner, 2009). Furthermore, when districts understand that there is an interaction between student behavior and reading outcomes (undesirable behavior can affect access to classroom instruction), plus seek to put a blended model of RTI in place for behavior

and reading, student outcomes have been shown to increase on both measures (Stewart, Benner, Martella, & Marchand-Martella, 2007). See Table 9 for the percent of districts that participated in one or more of the non-EBISS Oregon initiative.

Table 9

Other Major Oregon Initiative Professional Development 2006-2011

Professional Development	Non-EBISS (n=25)	EBISS (n=25)
DATA Project	40%	76%
Project CLASS	24%	20%
OrRTI	40%	52%
PBIS	100%	100%

EBISS professional development. EBISS districts submitted applications to ODE requesting permission to participate in the EBISS initiative. Applicants were required to identify the type of literacy systems the district's schools had in place. For example, did schools have screening and progress monitoring measures in place to identify the number of students at each benchmark (Tier 1, Tier 2 or Tier 3) in reading? What programs did the district have in place to increase the school's reading skills of students through a coordinated set of data, systems and practices that were guided by clearly defined outcomes (Sugai & Horner, 2002; Fuchs & Fuchs, 2006)? The data gathered from applicants provided valuable information regarding district practices. This information was used to determine the basic and advanced training districts would require to implement the EBISS model and meet the ODE goal of improved student outcomes.

All school districts received yearly funding to participate in EBISS. However not all districts participated in all of the trainings and not all districts submitted the requested data. Lack of participation in the required trainings or submission of required data did not result in a loss of district funding. Districts were also required to use some of their funds to support the role of a systems coach who provided on-going professional development and technical assistance to schools. The systems coaches were selected by the district for their knowledge of and skill to implement RTI for elementary literacy and PBIS. Systems coaches received training from EBISS coordinators yearly. During year three and four training webinars were recorded on RTI implementation features and two full day trainings were offered to provide districts with resources to use beyond the life of the initiative (sustainability and systems coaching). In year four the district systems coaches were required to communicate with their EBISS coordinator by email and/or telephone on a monthly basis. Table 10 identifies the professional development and technical assistance EBISS coordinators provided to district teams and systems coaches from 2006-07 to 2010-11. Note that most trainings were identified as required trainings, however not all districts attended all trainings, and the funding they received to participate in EBISS was not held back if they did not attend. In addition to EBISS professional development, some EBISS districts participated in the professional development provided by other major Oregon initiatives. Seventy-six percent of EBISS districts also participated in the DATA Project and 20% participated in CLASS. Fifty-two percent participated in OrRTI and 100% participated in PBIS. Participation in these initiatives may have resulted in EBISS districts having a deeper understanding of RTI implementation and an additive treatment effect. See Table 9 for a detailed description of the percent of EBISS districts

that participated in each of the major Oregon initiatives.

Table 10

Summary of Yearly EBISS Professional Development 2007-2011

Audience	Professional Development Content	Method of Delivery and Frequency
District Team	Welcome back webinar	Required: One-hour webinar, yearly
District Team	District level data review	Required: One site-visit by EBISS Coordinator, yearly Optional: Additional visits upon request
District Team	Data based decision-making within the context of efficient & effective team meetings	Required: 1 day workshop for entire district team, year 1 and 2 Optional: Attendance at State PBIS conference, 1-2 days, year 1 and 2
Systems Coach	District team facilitation & Data analysis	Required: Communication via phone, email, or webinar to help facilitate district level challenges to EBISS implementation, yearly
Systems Coach	SWRM implementation	Required: 2 day workshop (one in the fall and one in the spring), year 1 and 2 Required: 1 day fall workshop, year 3 and 4
Systems Coach	SWRM explicit reading instruction for grades K-5 Webinars on demand	9 day training of trainers provided, optional year 2 Optional: Webinars to reinforce and sustain EBISS practices beyond the life of the initiative, year 3 and 4
Systems Coach	Sustainability of EBISS: Systems Coaching	Required sustainability training, year 4 Optional: Systems Coach Manual Training, year 4

Measures

The OAKS reading scores at third grade were compared for EBISS and non-EBISS districts using a between subject design to measure if there was a significant difference in student performance between the two groups from T_1 to T_2. In addition, a within subject analysis of EBISS districts was conducted to determine if there was a differences between high-EBISS and low-EBISS districts (based on DSSP scores in 2011) and students' gains on the OAKS reading sub-tests at third grade from T_1 to T_2.

Oregon Assessment of Knowledge and Skills. The OAKS (ODE, 2000) is a statewide, multiple-choice assessment developed to evaluate students' mastery of content knowledge in the area of reading. To prepare students to reach the *meets* level on the OAKS at third grade, ODE suggests that all students should be screened three times per year in grades 1-3 in the areas of reading fluency and comprehension.

Technical adequacy of the OAKS. On the OAKS, students are presented with a variety of reading tasks to assess vocabulary, reading to perform a task, demonstration of general understanding, interpretation, and content and structure for informational and literary text. Students earn a scaled score that represents specific performance criteria (*Nearly Meets*, *Meets* and *Exceeds*). The OAKS reading-literature assessment is given to all students in grades three to eight and grade 11. Reliability of achievement classification ranged form 84-99%, with most falling above 90% (McCall, 2003).

ODE commissioned a series of technical adequacy reports in 2006 to provide the public with information regarding the technical and procedural characteristics of the OAKS (ODE, 2006). Reports are updated as new evidence is collected. For example, ODE reports on the reliability of the OAKS using test information curves, standard error

of measurement, the reliability of achievement classification, strands, and precision at cut scores. They also provide extensive reports on content, concurrent, and criterion validity.

Reliability of OAKS. Reliability refers to the consistency, stability, and accuracy expected from test scores (ODE, 2006). Internal consistency of the OAKS was established (.95) using 4 alternative assessments for OAKS reading using the Kuder-Richardson Formula 20 (ODE, 2006). Analysis of the standard error of measurement was also calculated and demonstrated that students with the same overall proficiency level is measured with the same reliability regardless of the demographic subgroup being tested (except at the 99th percentile) (ODE, 2006).

Validity of OAKS. The OAKS tests are examined using the guidelines from the Standards for Educational and Psychological Testing (1999). ODE ensures the validity of the OAKS by providing evidence and theory that supports the interpretation of the test scores and their intended use (district accountability). The Standards for Educational and Psychological Testing (1999) recommends an examination of three types of validity: content, construct, and criterion. To establish content validity ODE developed OAKS test items to match Oregon's content standards. Test items range in difficulty and sample four major categories: (a) what students should know and do, (b) clear link to test content, content standards and corresponding performance levels, (c) consensus driven test item development, and (d) on-going evaluation to ensure instruction, assessment and standards are aligned.

ODE also established construct and criterion validity (OAKS, 2000). Construct validity was established by relating student's scores on the OAKS with scores on other tests measuring the same construct. The OAKS is correlated in reading with the

California Achievement Test (.75) and with the Iowa Test of Basic Skills of (.78) (ODE, 2000). Criterion validity was established by developing a set of performance standards that identified minimum test scores for meeting and exceeding the content standards and then investigating if there was a relationship between performance on the OAKS (i.e., 10th grade benchmarks and college performance as a freshman). ODE found that students who met or exceeded the 10th grade benchmark had higher Scholastic Assessment Test (SAT) scores and higher freshman retention rates, as opposed to students who did not meet the OAKS benchmarks.

District System Support Plan. This study also investigated if there was a difference in EBISS districts OAKS performance at third grade based on districts high or low DSSP scores. The DSSP is a *district-level* self-assessment used to evaluate a district's implementation efforts over time and test the strength of district-wide organizational systems for early literacy. DSSP scores guide the development of district and school's action plans through on-going assessment and evaluation of scores on the DSSP's seven constructs (and 23 items). The seven constructs include: (a) leadership and commitment, (b) action planning with schools, (c) coordination and coaching, (d) professional development and training, (e) ongoing assessment and evaluation, (f) visibility and stakeholder support, and (g) funding. The DSSP scores were used to measure the growth of districts organizational systems, so districts could enter into data based problem solving and action planning. The district leadership team identified problems, developed solutions and evaluated fidelity of implementation plans overtime to determine if: (a) the critical features of RTI for literacy were being implemented with fidelity, (b) organizational systems and practices were in place at the district and school

level, and (c) continuous improvement of staff skills was supported through adequate

funding, professional development and coaching.

Twenty-five EBISS district leadership teams completed the DSSP and a copy of

the survey can be found in Appendix A. The team scored each item (23) on a 4 point

scale with a 0 = *not in place,* 1 = *partly in place*, 2 = *mostly in place*, and 3 = *completely*

in place. A percentage was assigned to each of the seven constructs and then an overall

average of implementation score was calculated. Teams recorded their initial score and

their progress scores across the four years. See Table 11 for a brief description of each

construct of the DSSP and Chaparro, Ryan Jackson et al. (2012) for a detailed

description.

Each year a Webinar was provided that explained the DSSP team self-assessment:

what it was, why the data was important, and how it should be collected and used for

decision-making. An EBISS state coordinator attended at least one district leadership

meeting a year to provide technical assistance to the team as they scored the DSSP. A

representative district leadership team was instructed to score the DSSP individually and

then come together to discuss and agree on the score for each of the 23 items within the

seven constructs. Then, districts entered the data into an on-line data portal or they sent

their EBISS coordinator an email attachment or faxed copy of the DSSP. The EBISS

coordinator was also available for follow-up support via phone and email as districts used

the DSSP data to set and revise action plans across the years. DSSP data was entered into

an Excel spreadsheet and then into the IBM SPSS system for analysis in this study.

Table 11

District System Support Plan (DSSP)

Constructs	Characteristics
1. Leadership & Commitment	Seven items assess the strength of the team and districts core implementation components.
2. Action Planning with Schools	Three items assess schools performance based on their unique context to identify required district support.
3. Coaching	Three items ensure staff receives the coaching required to implement district reform initiatives through development of a comprehensive coaching plan.
4. Professional Development	Two items to assess professional development needs and develop a comprehensive training plan.
5. On-Going Assessment and Evaluation	Three items identify required data systems and the support schools needed to collect and analyze data to develop action plans.
6. Stakeholder Support	Two items to garner support for initiatives.
7. Funding	Two items to ensure adequate funding for district reform initiatives.

Technical adequacy of the DSSP. From 2008-09 to 2010-11 twenty-five districts completed the DSSP. In 2010-11 the DSSP was scored twice, once in November of 2010 and then again in May of 2011. Districts were required to score the DSSP again in the spring of 2011 because of the significant number of district leaders who were exiting EBISS districts. EBISS coordinators wanted to ensure DSSP scores were updated prior to the arrival of a new administration that would not be familiar with the DSSP or the district's level of implementation fidelity. Collection of DSSP data from 2008-2011

66

resulted in 125 cases for analysis and in 2011 technical adequacy of the DSSP was

established. Chaparro, Park, Baker and Ryan Jackson (2011) established technical

adequacy of the DSSP and found that the DSSP had the potential to be an effective

fidelity measure for implementing RTI for literacy.

Reliability of DSSP. Appendix B presents the descriptive statistics for each item

of the DSSP, as well as Cronbach's alpha for each of seven constructs or factors. All

seven factors in the DSSP survey demonstrated alpha over .70 regardless of the small

numbers of items per factor (2 to 8). For example, funding and professional development

constructs only have two items per construct. Given the alpha of .70 and above for each

item in the DSSP Cohen (1988) would suggest that the DSSP has a medium or typical

effect. Therefore, the DSSP has evidence of internal reliability.

Validity of DSSP. Standards for Educational and Psychological Testing (1999)

recommend an examination of three types of validity: content, construct, and criterion. A

representative team of researchers, practitioners and policy makers developed the DSSP

in 2006. In order to ensure that the DSSP represented the intended content domain the

measure was developed through logical analysis and evaluation of the instrument

(examining its items, formats, wording, testing processes). Then, twenty-five EBISS

districts tested the DSSP in 2007. The DSSP was revised through logical analysis and

released to EBISS districts in 2008 for scoring. Next, construct validity was examined.

Construct validity examines the extent to which the internal components of an

instrument match the defined constructs. This is usually estimated as a model fit indices

by confirmatory factor analysis (Kline, 2011). To investigate the internal consistency of

the DSSP three indices were used: (a) the root mean square error of approximation

(RMSEA), (b) the comparative fit index (CFI), and (c) the standardized root mean square residual (SRMR). The RMSEA index was .09 (90% confidence interval: .80 - .10) indicating reasonable fit. The model demonstrated CFI = .87 which is very close to the standard implying a reasonable fit. The SRMR index for the model was .06 implying a good fit. Taken together, the multiple model fit indices indicated that the DSSP survey has reasonable to good construct validity (Chaparro, Park et al., 2011).

Criterion validity concerns how well one measure predicts an outcome based on information from another measure and it is often divided into concurrent validity and predictive validity (Messick, 1995). The criterion validity of the DSSP was examined by looking at correlations with the seven constructs of the DSSP and reading outcomes at third grade (presented in q`1). Improving student outcomes is the goal of EBISS and the real test of validity (Messick, 1994). The DSSP demonstrated low correlations with district-level performance (district leadership and commitment) measured by the percentage of students meeting OAKS at third grade. The model suggests there is a small or smaller than typical relationship between *district leadership and commitment* and *action planning with schools* and improvement in the number of students meeting the state reading assessment at third grade. It was not statistically significant (*p* < .10). See Table 12 for a complete description of the correlations between all DSSP constructs and OAKS distal outcomes at third grade.

Table 12

Correlations Between DSSP Results and OAKS at Third Grade

	District Leadership	Action Planning	Coaching	Professional	Assessment	Visibility	Funding
Grade 3 Reading (OAKS)	.189[†]	.170[†]	.019	.122	-.035	.050	.044

[†] $p < .10$

Statistical Methods

First, to control for group differences the non-EBISS and EBISS groups, dichotomous and continuous data in the form of counts and percents were examined to ensure the groups were not significantly different for the district characteristic of region and enrollment in 2006. A mean and standard deviation was computed for FARM, white and non-white, male and female, years of teacher's experience, and (g) 3rd grade OAKS scores in 2006 to determine if both groups were relatively similar and not significantly different. Then, a *t*-test was calculated to determine whether the mean difference between the two groups was significantly different. My analysis showed that there was no statistical evidence of a mean difference between the percent of students in each group for region ($p = 0.100$), enrollment ($p = 0.904$), FARM ($p = .579$), white and non-white ($p = .942$), male and female ($p = .916$), the average teacher's years of experience ($p = .626$, and 3rd grade OAKS scores ($p = .540$).

Between subjects repeated measures analysis of variance (rANOVA). To test if EBISS districts outperformed non-EBISS districts an rANOVA was conducted to

determine if there was a mean difference between the gains in the percent of student's passing the third grade OAKS for non-EBISS and EBISS districts from T_1 (2006-07) to T_2 (2010-11). The rANOVA tests an interaction between the two independent variables (non-EBISS and EBISS) and time (2006-07 and 2010-11) on the dependent variable for the percent of students passing the OAKS reading test at third grade. The rANOVA is "mathematically equivalent to [a t-test comparison] of simple gain scores" (Anderson et al., 1980, p. 238). Greater gains in the percent of students who pass the OAKS reading assessment at third grade for EBISS districts suggests responsiveness to the professional development that EBISS districts received to implement RTI for early literacy. It is important to note that the data under analysis is in the form of percentages of students in a district that did not meet, met or exceeded OAKS reading and is not equal interval data. Therefor, arc-sine transformation was applied to stabilize the variance in the data set and normalize the proportional data (Hogg & Craig, 1995).

Arc-sine transformation. The percentages that represented the met or exceeded OAKS standard for districts are not equal interval data. Each data point represents a group with different means and different variances. Arc-sine transformation is used in education research to stabilize variance and normalize proportional data (Jaeger, 2008; Stallings, Robbins, Presbrey, & Scott, 1986). As a preparation for doing statistical analysis (rANOVA), the data in this study was converted through an arc-sine transformation in Excel to radians and then to degrees to allow for appropriate comparison of data points. This allowed for analysis that more accurately represented the attributes of the data.

Within subject rANOVA. To examine if there was a difference within EBISS districts, the district's DSSP scores were converted into two groups based upon the median DSSP score from 2011. The median DSSP score in 2011 was 2.375 for the 25 EBISS districts. Districts with an average DSSP score greater than 2.375 were identified as high-EBISS and districts with a DSSP score lower than 2.375 were designated as low-EBISS). Then, a 2x2 rANOVA was conducted to determine if there was a mean difference between high and low DSSP groups on the gains in the percent of student's passing the third grade OAKS from T_1 to T_2.

CHAPTER III

RESULTS

Following is a summary of the two research questions this evaluation aimed to answer when studying the effects of the Effective Behavioral and Instructional Support Systems (EBISS) initiative. To answer the first research question a quasi-experimental matched comparison analysis was used to determine if there was a difference between non-EBISS and EBISS district's students' gains in the reading sub-tests of the Oregon Assessment of Knowledge and Skills (OAKS) at third grade from time one (T_1) (2006-07) to time two (T_2) (2011-12). To answer the second research question, District Systems Support Plan (DSSP) scores were used to create high-DSSP and low-DSSP groupings. Then the OAKS score for those two groups were analyzed to determine if there was a difference on the reading sub-tests of the OAKS at third grade from T_1 to T_2.

Question One – Comparison of non-EBISS and EBISS Districts

A 2x2 rANOVA was conducted to answer the first research question. That question asked if students in EBISS districts made statistically significantly greater gains than non-EBISS districts on the percent of students who passed the third grade OAKS reading assessment from T_1 to T_2. The rANOVA indicated no significant interaction ($p = .524$) between non-EBISS and EBISS districts from T_1 to T_2 on their OAKSr scores. Also, there was no main effect for group ($p = .312$) or for time ($p = .488$). The means and standard deviations for the EBISS and non-EBISS districts OAKS scores at third grade were: non-EBISS (T_1 $m = 66.77$ $sd = 6.51$ and T_2 $m = 67.67$ $sd = 5.20$) and EBISS (T_1 $m = 65.75$ $sd = 5.13$ and T_2 $m = 65.79$ $sd = 5.18$). The strength of the relationship between EBISS treatment (PD and TA) and the percent of students who passed the third grade

OAKS reading assessment from T_1 to T_2, as assessed by partial eta squared, was not

significant for time (.010), interaction (.008), or group (.021). See Table 13 and Table 14

for complete descriptive and rANOVA statistics.

Table 13

Means and Standard Deviations for EBISS and non-EBISS

	Time 1 OAKS		Time 2 OAKS	
Group	M	SD	M	SD
non-EBISS	66.77	6.51	67.67	5.20
EBISS	65.75	5.13	65.79	5.18

Table 14

rANOVA Effects for EBISS and non-EBISS

Source	SS	Df	MS	F	Sig.	Eta
Group	52.53	1	52.53	1.05	.312	.021
Time	5.40	1	5.41	.489	.488	.010
Interaction	4.55	1	4.55	4.11	.524	.008
Within	4.55	1	4.55			

Question Two – EBISS Within Group Comparison on DSSP Scores

A within subject analysis of EBISS districts was conducted to answer the second

research question. That question asked if students' gains in the reading sub-tests of the

OAKS at third grade from T_1 to T_2 was different for districts with high DSSP scores versus low DSSP scores.

A 2 x 2 rANOVA was conducted to determine if there was a difference between high-EBISS and low-EBISS DSSP scores and district's students' gains in the reading sub-tests of the OAKS at third grade from T_1 to T_2. The rANOVA indicated no significant interaction ($p = .452$) between high EBISS (median > 2.375 in 2011) and low-EBISS (median ≤ 2.375 in 2011) districts from T_1 to T_2 on their OAKSr scores. Also, there was no main effect for group ($p = .613$) or for time ($p = .947$). The means and standard deviations for OAKS reading at third grade between high-EBISS and low-EBISS DSSP groups were: High-EBISS (T_1 $m = 65.82$ $sd = 3.58$ and T_2 $m = 66.69$ $sd = 5.26$) and low-EBISS (T_1 $m = 65.69$ $sd = 6.40$ and T_2 $m = 64.96$ $sd = 5.17$). The strength of the relationship between high-EBISS and low-EBISS and the percent of students who passed the third grade OAKS reading assessment from T_1 to T_2, as assessed by partial eta squared, was not significant for interaction (.025), or group (.613). See Table 15 and 16 for complete descriptive and rANOVA statistics.

Table 15

Means and Standard Deviations for EBISS DSSP on OAKS

Source	DSSP 2011 Median Score	Time 1 OAKS M	SD	Time 2 OAKS M	SD
High-EBISS	>2.375	65.82	3.58	66.69	5.26
Low-EBISS	≤2.375	65.69	6.40	64.96	5.17

Table 16

rANOVA Effects for High-EBISS and Low-EBISS DSSP Scores

Source	SS	Df	MS	F	Sig.	Eta
Group	10.79	1	10.79	.264	.613	.011
Time	.062	1	.062	.055	.947	.000
Interaction	8.033	1	8.033	.585	.452	.025
Within	8.033	1	8.033			

Results Summary

In summary, the results of the first research question indicated no significant interaction between EBISS and non-EBISS districts scores on the OAKS reading test at third grade, $p = .524$ from T_1 to T_2. Also, no main effect for group ($p = .312$) or time ($p = .488$) was found, and no strength of relationship was noted, $n^2 = .008$. Similarly, the results of the second research question indicated no significant interaction between the grouping variable (DSSP scores) and gains in the reading sub-tests of the OAKS at third grade, $p = .452$ from T_1 to T_2. Furthermore, no main effect for group ($p = .613$) or time ($p = .947$) was found, and no strength of relationship was noted, $\eta^2 = .025$.

CHAPTER IV

DISCUSSION

The findings from this EBISS program evaluation have led to the following statistical conclusions based upon our collected data. After four years of providing technical assistance (TA) and professional development (PD) to 25 EBISS districts in the State of Oregon, EBISS districts did not meet the Oregon Department of Education (ODE) goal of improving student outcomes when compared to non-EBISS districts. After matching EBISS districts to a comparison sample of non-EBISS districts, no statistically significant *between group* difference was indicated in the gains in the percent of students passing the Oregon Assessment of Knowledge and Skills (OAKS) Reading at third grade across time, $p =.524$, and no strength of relationship was noted, $\eta^2 = .008$. Moreover, no statistically significant difference was found *within* EBISS districts when comparing districts with high and low implementation scores on the District Systems Support Plan (DSSP) and their gains in the percent of students who passed the OAKS Reading at third grade, $p = .452$, and no strength of relationship was noted, $\eta^2 = .025$.

Research Question One

For this study, a matched comparison sample was carefully designed to ensure that non-EBISS and EBISS districts were similar and not significantly different in the pre-implementation year of the EBISS initiative in 2006. The two groups were matched by demographic factors of (a) region, (b) enrollment, (c) FARM, (d) white and non-white, (e) male and female, (f) years of teacher's experience, and (g) 3rd grade OAKS scores. After demographic matching, a *t*-test was then computed to confirm that no statistical difference between groups existed on the percent of students' passing the

OAKS reading test at third grade in 2006, $p = .540$ (see Methods chapter for complete matching statistics).

When looking at change over time, my current program evaluation found no statistically significant difference in the gains of the percent of students' passing the OAKS at third grade from 2006 (pre-implementation) to 2011 (post implementation), $p = .524$. This finding will be discussed at length, mostly in relation to using student-level data to measure district (adult) interventions. Several limitations exist in the evaluation of proximal data (district leadership implementation effectiveness) based on distal outcomes (student's scores on state assessments) (Grissom et al., 2012) because the variation that is found in implementation quality (adult behavior) is related to variation in distal outcomes (standardized reading assessments) (Cook, 2000).

Using distal student outcomes to measure administrative outcomes. Federal and state policy is drawing public attention to the evaluation of a district's administrators based on state assessment results (ESEA, 2010; Grissom et al., 2012). Grissom at al. (2012) recognized multiple problems that come with measuring administrator effectiveness based on student test scores and suggested an alternative holistic evaluation (non-test measures) of district effectiveness that isolates administrator's contributions to its school's environments. The authors identified three potential problems when using distal data (student's scores on state assessments) to evaluate proximal effects (district implementation effectiveness): (a) clarity regarding what construct is being captured (theoretical ambiguity), (b) error in the identification and measurement of administrator behaviors (potential bias), and (c) confounding variables that interact with the measurement of performance (reliability). Therefore, these noteworthy limitations must

77

be considered when interpreting the results of research question number one.

Theoretical ambiguity. In question number one the existence of theoretical ambiguity must be considered given how districts within a state vary (Mathis, 2003). Districts in both groups (non-EBISS and EBISS) varied. Some districts were rural and others were urban-suburban creating significant differences in funding, resources, and access to on-going professional development (PD) and technical assistance (TA) from EBISS or the other major Oregon initiatives. Districts within easy driving access (urban-suburban) may have received additional TA from EBISS coordinators or other major Oregon initiative providers to develop administrator skills. At the same time nine EBISS and six non-EBISS districts in my study were also part of Reading First (Baker et al., 2011) and entered the study with considerable skill to implement Response to Intervention (RTI) for reading at the district and school level because of their Reading First training. The possibility that some districts may have received additional TA or that some entered the study with strong district skill sets may have resulted in theoretical ambiguity or clarity regarding what construct was being captured (Grissom et al., 2012). Monitoring the amount of TA districts received (Fixsen et al., 2009; Cook, 2000; Grissom et al., 2012) and measuring the differences in districts skills pre and post implementation (Stockslager et al., 2013) could have explained why there was no difference between non-EBISS and EBISS districts gains in student reading outcomes.

Potential bias. The conditions associated with improvement (antecedents to district and school effectiveness: Leithwood & Jantzi, 2008) were not measured in non-EBISS districts). Non-EBISS districts did not collect DSSP data to measure and evaluate the growth of system-wide alignment to district's priorities. Consequently, bias may exist

in the interpretation of non-EBISS results exists due to the error in (or lack of) identification and measurement of administrator behaviors for non-EBISS districts (Grissom at al., 2012). EBISS districts were measured on the seven constructs known to be present in successful system-wide reform models focused at the district level (EBISS DSSP measure). The DSSP constructs and items can be compared to Leithwood and Jantzi's antecedents to district effectiveness and the variables the Evaluation Blueprint (Algozzine et al., 2010) measures. However, Grissom et al. (2012) caution interpreting gain scores due to the turnover of district and school administrators across evaluation years (for EBISS and non-EBISS). Reliability of the distal measurement when comparing EBISS to non-EBISS is also a limitation in this study.

Confounding variables. Confounding variables (specifically other Oregon RTI initiatives) that interact with the measurement of performance (reliability) must be considered when interpreting the comparison results for non-EBISS and EBISS (Grissom et al., 2012). Both groups are nested in the State of Oregon where considerable effort (funding to districts, PD, TA) has been provided to implement all or some of the critical features of RTI (PBIS, OrRTI, DATA Project). Therefore, understanding the counterfactual (Cook, 2000) or districts that did not receive EBISS funding, PD, TA may inform why no statistically significant difference between non-EBISS and EBISS was detected when comparing student outcomes. Investigating the conditions in the state (e.g., parsing the contributions of all the various RTI TA and PD) that affected my dependent variable (distal student outcomes) may inform why no statistically significant difference between groups was detected. For example, the ODE has been providing districts with the opportunity to participate in initiatives that teach schools how to implement the core

components of RTI since 2005. One example is the Oregon Response to Intervention (OrRTI) initiative. Districts can apply to formally participate in OrRTI or attend one-time trainings at local conferences sponsored by the Confederation of School Administrators or Northwest Positive Behavioral Interventions and Supports. District participation in state initiatives other than EBISS may have confounded the measurement of the dependent variable when comparing non-EBISS and EBISS districts.

Oregon's Major Initiatives

Babbie (2007) suggests that measuring the dependent variable in evaluation research is only the beginning for further investigation of all of the variables affecting the dependent variable must be considered. To this end, understanding all of the initiatives that non-EBISS districts had access to from 2006 to 2011 is important (and a difficult task to complete). For example, ODE has been focused on providing districts support to implement RTI for early literacy since 2001 when one Oregon district received federal money to develop an RTI model that could be scaled up across the state (Sadler & Sugai, 2009). Therefore, consideration of Oregon's major initiatives that focus on the implementation of RTI that non-EBISS districts participated in (other than EBISS) is an important aspect of this program evaluation. The other major initiatives included: (a) Direct Access To Achievement Project (DATA Project); (b) Creative Leadership Achieves Student Success (CLASS); (c) Oregon Response to Intervention (OrRTI), and (d) Positive Behavioral Interventions and Supports (PBIS). See the Methods section for a description of each initiative. Initiatives provided districts with TA and PD (and small supplemental grant money in some cases) to implement one or all of the core components of RTI. However, it is important to bring attention to the fact that all major Oregon

initiatives (except EBISS) focused on student interventions at the school level. They only involved district indirectly. For example, district may have identified a school that was interested, but district was not required to participate in any PD or TA that was provided. District was not identified as the impetus for change.

Continued investigation. Using the guidance of Babbie (2007), this study evaluated whether TA and PD to implement the core components of RTI for early literacy had an impact the OAKS reading performance at third grade for EBISS districts over and above that of non-EBISS districts. Given that there was no difference found between non-EBISS and EBISS distal outcomes, one could infer that the TA and PD from the other Oregon initiatives was equally effective when measuring the dependent variable, gains in the OAKS reading scores at third grade. So, further investigation was warranted to shed light on why the EBISS initiative did not show statistically significant gains in OAKS reading scores at third grade, over and above non-EBISS districts. Therefore, a second research question was investigated to determine if there was a difference in the gains of third grade OAKS reading scores within EBISS districts when considering the range of scores districts assigned themselves on the DSSP (see methods for a detailed description of the DSSP).

Research Question Two

A repeated measure ANOVA was conducted to determine if there was a difference within EBISS districts on students' gains in the reading sub-tests of the OAKS at third grade from 2006 to 2011. High-DSSP and low-DSSP districts were identified based on each district's overall average DSSP score in 2011 (the DSSP was scored on a scale of 0-3, a district score of ≤ 2.37 = low and > 2.37 = high). However, no statistically

81

significant difference was noted, $p = .452$ between high-DSSP and low-DSSP gains in the OAKS reading scores at third grade. Consequently, further investigation of the literature on measuring distal outcomes based on proximal data was warranted, because evaluation of the EBISS initiative aimed to demonstrate a relationship between gains in OAKS reading scores at third grade (the distal measure) and changing adult behaviors (DSSP proximal data as a measure of changing adult behavior).

Klinger, Boardman and McMaster (2013) emphasized the need for program evaluation to provide data that indicates a program had a positive effect on student achievement. However, Cook (2000) and Grissom et al. (2012) submit that measuring gains on student achievement is not a direct measure of changing adult behavior (or linking adult behavior to student outcomes). To address this measurement issue, Leithwood and Jantzi, (2008) suggested that evaluation first identify strong and weak leadership characteristics (e.g., surveys or seven constructs of the DSSP) and describe the size and nature of leadership's impact (proximal measurement of district antecedents on organizational change, e.g., change in DSSP scores), to measure its relationship to student outcomes (e.g., change in OAKS reading scores at third grade). Still, Grissom et al., (2012) warned against program evaluators attempts to link adult behavior to gains in student outcomes because, "there simply may not be enough signal in average student achievement gains at the school level to get acceptably reliable measures of improvement" (p. 13). Given the warning of Grissom et al. and the advise of Leithwood and Jantzi one could question whether the DSSP assessment identified and accurately measured antecedents to district behavior, for evaluation must identify and carefully measure antecedents (what they term strong and weak district behavior) or what Grissom

et al. identify as non-test variables (e.g., goal setting, organizational skills) when attempting to link adult behavior to student outcomes.

Leithwood and Jantzi, (2008) suggested that non-test measures of implementation gathered though self-assessment surveys can inform the difference between strong and weak leadership behaviors. Yet, Gersten, Dimino, Jayanthi, Kim and Santoro (2010) warned that the majority of studies on the effects of PD on student outcomes rely on self-report data and fail to directly measure implementation fidelity by trained facilitators who have established inter-observer reliability. Inter-observer reliability for scoring the DSSP was not established among EBISS coordinators. Nor were district sessions taped, so accurate score agreement could be established by at least two EBISS coordinators (e.g., ensure a score of three, fully in place, was the same in each district). Consequently, the reliability of DSSP scores is suspect. For example, Gersten et al. (2010) established inter-rater reliability for facilitators (a role similar to that of EBISS coordinators who deliver TA and PD) prior to observation and subsequent scoring of the implementation checklist they used in a randomized control study on the effects of PD on teacher behavior and student outcomes (the DSSP can be considered a checklist as well).

This brings attention to the fact that the DSSP was designed as a self-report measure to assess fidelity to the seven constructs known to be present in districts who successfully implement the critical components of RTI (Fuchs & Fuchs, 2006; Sadler & Sugai, 2009). EBISS teams self-scored the DSSP, however, a scoring protocol to establish inter-rater reliability within a district team was not established to ensure that all districts had a common way for a team to score the DSSP. Ensuring the inter-observer

and inter-rater reliability of DSSP scores could have more accurately identified the antecedents of strong and weak district behaviors.

Strong and Weak District Behaviors

Developing a deeper and more accurate understanding of the difference between strong and weak leadership behaviors within EBISS districts may have identified antecedents of organizational characteristics (DSSP constructs) and the amount of TA and PD individual districts would require to achieve the states goal of improved student outcomes. For while no statistically significant difference was noted within the EBISS group, the EBISS initiative did have one distinct difference that stood out from all other major Oregon initiatives that must be considered. EBISS placed the onus on district to change adult behaviors system-wide (Sugai & Horner, 2006; Slavin et al., 2013). Identifying and carefully measuring the antecedents to district behavior prior to implementation would have allowed EBISS coordinators to differentiate the TA and PD districts required.

Antecedents to changing district behavior have been identified (e.g., clear district goals, development of staff skill) (Leithwood & Jantzi, 2008). Accurately identifying antecedents and describing the size and nature of leadership's impact on organizational change could have informed the additional TA and PD districts required. While the DSSP was designed as a fidelity measure to test if district leadership (changes in adult behavior) was improving implementation of the critical features of EBISS in schools it had its limitations (see Table 11 in Methods for a complete description of the DSSP). The technical adequacy of the DSSP demonstrated a small or smaller than typical effect between *district leadership and commitment* and the number of students meeting the state

84

reading assessment at third grade, ($p < .10$). Therefore, one could surmise that the DSSP

measure did not capture the necessary antecedents of strong and weak leadership

behavior to inform the PD and TA that was required to improve distal outcomes.

       District leadership behaviors and distal measurement. Given the fact that no

statistically significant difference was found within EBISS districts one could surmise

that EBISS TA and PD was not sufficient to generate a statistically significant increase in

student outcomes on the OAKS at third grade. However, to gain a deeper understanding

of the independent variable (EBISS TA and PD), Cook (2000) identified the need for

program evaluation to understand the variation in implementation quality and how it is

related to change in distal outcomes. Implementation quality has been identified as the

critical component to improved student outcomes (Gersten et al., 2010). In EBISS this

meant using the DSSP scores as a measure of system-wide implementation fidelity. Cook

specifies the need for prerequisite knowledge of program theory (RTI) and an

understanding of the mediating processes (DSSP constructs as antecedents to district

behavior) that must be in place to realize district-wide change.

       EBISS was Oregon's first endeavor to implement RTI focused solely on

providing TA and PD to district in an attempt to change adult behaviors system-wide.

Many lessons were learned following the four-year initiative. For example, the EBISS

initiative did not collect qualitative data other than the DSSP (e.g., surveys, interviews

with district and school staff) to identify strong and weak leadership characteristics, nor

was inter-observer or inter-rater reliability established to ensure DSSP scores reflected

the size and nature of change in adult behavior (Leithwood & Jantz1, 2008). Grissom et

al. (2012) advocated for the collection of multiple non-test measures (e.g., surveys,

interviews) to quantify the variance in districts' effects that separate district behaviors from principal's behavior, and school staff behaviors (Robinson et al., 2008).

Linking changes in adult behaviors at the district and school level (seven constructs of the DSSP) and subsequent changes in student outcomes (distal measure) would have informed how EBISS implementation components could be strengthened when delivering TA and PD on an individualized basis (Grissom et al., 2012). For example, also measuring strong and weak principal behavior could have allowed for EBISS PD that focused on developing the skills of district administrators to support principal's instructional and organizational skills (Honig et al., 2010). Then, EBISS coordinators could have followed-up with the appropriate amount of TA to ensure the PD was transferred to practice by school administrators (Gersten et al., 2008).

Because multiple measures of leadership (surveys, interviews) were not collected by EBISS, continued investigation must try to examine the presence or absence of the conditions present in EBISS districts (commitment, organizational components, capacity). Further investigation of the conditions in EBISS districts may indicate if (a) the TA and PD provided was sufficient to improve student reading outcomes, over and above non-EBISS districts, (b) non-compliance to some of the critical implementation components of EBISS existed, and (c) if district leaders had the will and commitment to change district leadership behaviors and support schools' implementation of EBISS with fidelity. First, let us consider if the TA and PD provided to EBISS districts was sufficient.

EBISS technical assistance and professional development. Fixsen et al. (2009) describe TA as specialized support that is delivered to build capacity and achieve desired outcomes. Given the fact that EBISS state coordinators only delivered one site visit per

86

year to each district (as opposed to 23 per year by the Florida RTI initiative: Stockslager et al., 2013) one could argue that the TA was not sufficient and possibly not delivered in a timely manner as implementation problems arose. Fixsen et al. (2009) identifies the need to differentiate the delivery of TA based on an organizations capacity: (a) basic TA (have capacity to institute system-wide change), and (b) intensive TA (do not have capacity and are challenged by long held beliefs). While EBISS coordinators were available for phone consultation as needed, it may not have been sufficient for districts who required intensive on-site TA to change current beliefs of district staff and to realign goals, priorities, policy, and resources to change adult behaviors. As states and districts consider implementing a common RTI framework in all of its schools they will want to consider the type and level of TA district and school leadership will require when transferring PD to practice.

The transfer of PD to practice is not a one-time event (Dow et al., 2012). Important components like an attendance policy that is monitored and a system for follow-up PD (trainer of trainer model: Stockslager et al., 2013) must be present (Dow et al., 2012). Collection and monitoring of attendance would have been an important variable to measure when considering the antecedents to changing district behavior in EBISS. Grissom, et al. (2012) suggested that this non-test measure could reduce theoretical ambiguity when measuring changes in adult behaviors and its relationship to improved student outcomes (you can not measure PD that was never received). Without doubt some EBISS districts did not attend all trainings and some did not attend with a representative district leadership team. While EBISS required attendance at all PD and clearly communicated that district leadership should attend (newsletters, emails)

attendance was not collected nor was funding held back when districts did not attend. Of equal importance is the knowledge of mediating processes in a district (Cook, 2000) or the beliefs and skills of adults (Stockslager et al., 2013).

Identifying and differentiating the TA and PD provided to each school, based on their unique needs, (Sadler & Sugai, 2009; Chaparro, Ryan Jackson et al., 2012) would require the administration of a survey (non-test measure: Grissom, et al., 2012) or needs assessment to determine the current level of staffs' skills and their beliefs that a common RTI framework will improve student outcomes (Stockslager et al., 2013). Aligning TA and PD with the needs and beliefs of all staff, at all levels of the system, could have served to transfer TA and PD components to practice in EBISS districts. However, TA and PD alone will not change district behaviors, because a lack of sustained district leadership and commitment can present a barrier to district and schools implementation efforts.

Establishing district leadership and commitment. Like Oregon's EBISS initiative, the Florida RTI initiative reported that sustained district leadership and commitment: (a) inconsistent district policies, (b) procedures for implementation of RTI, and (c) inconsistent attendance at the PD posed a barrier to districts successful implementation (Stockslager et al., 2013). Henceforward, initial establishment of district willingness and commitment may be the critical component required to establish system-wide change that improves student outcomes. In short, many EBISS district leaders may not have possessed the capacity or the will and commitment to provide the support required to transfer EBISS PD to practice in schools following all EBISS PD events. Understanding each district's will and commitment to implement EBISS could have

88

given EBISS state coordinators a better understanding of the challenge at hand. For, districts had to have the will and commitment to increase their instructional leadership skills and support all staff in the system-wide change (Honig et al., 2010; Lane, 2009; Rorrer et al., 2008).

System-wide implementation of RTI. EBISS focused on changing district leadership behaviors, system-wide, to meet state and federal accountability measures through alignment of organizational systems and policy in a common RTI framework (Lane, 2009). Technical assistance and PD sought to strengthen district's instructional leadership skills as district administrators and coaches worked with school's leadership (principal, coaches, teacher leaders) to support the unique needs of each school (Honig et al., 2010). In turn, school leadership worked with teacher teams and individual teachers to ensure school staff understood the district's priorities and goals (EBISS Teaming Framework: Chaparro, Ryan Jackson et al., 2012; Sadler & Sugai, 2009) and had the knowledge (Robinson et al., 2008), skills (Slavin et al., 2013), resources (Rorrer et al., 2008), and policy (Fixsen et al., 2013) to implement RTI using a common district framework in each school (Fuchs & Fuchs, 2006; Lane, 2009). Clearly, implementing RTI system-wide for early literacy will require more than TA and PD to improve student outcomes. The critical ingredient that must be present is districts sustained will and commitment (Putnam, 2008; Rorrer, et al., 2008) to align organizational systems through district's instructional leadership skills, goals and priorities (Lane, 2009) that are clearly articulated in policy through equitable resource allocation (Carnine, 1999; Fixsen et al., 2013).

89

It is clear that leading successful district-wide reform initiatives is not for the faint of heart. District must clearly define and articulate its goals and priorities (Sadler & Sugai, 2009; Rorrer et al., 2008) and align the organizational systems (Fixsen et al., 2013; Lane, 2009) with sound policy to support schools' implementation efforts (Carnine, 1999; Fixsen et al., 2013) and the development of staff skill (Slavin et al., 2013; Fixsen et al., 2013) through adequate resource allocation (Rorrer et al., 2008, Sadler & Sugai, 2009). Districts must also have the will and commitment to develop capacity by aligning resources to meet district goals and the identified need of all staff and all students. Rorrer et al. (2008) stress that it is the alignment that reveals a districts organizational purpose, values and desired outcomes. This alignment can be thought of as a *commitment to a decision* (Firestone, 1989) even when others *oppose the decision* by proactively educating all stakeholders, raising expectations, and ensuring expectations are met (Jacobson, 1986).

Limitations

Clear limitations exist in the EBISS program evaluation that is worth noting. First, it is important to understand that the EBISS initiative was a professional development initiative and it was not designed as a research study. Therefore, threats to interval and external validity were not controlled for or equalized prior to districts implementation efforts in the 2007-08 school year. In addition, threats were not controlled for during the four-year EBISS initiative.

Internal validity. Internal validity is concerned with the extent to which extraneous variables are controlled for in order to eliminate sources of error variance (Parker, 1993). Failure to control for extraneous variables can limit the interpretation of

the conclusion that seeks to suggest that the independent variable was related to changes in the dependent variable. In the EBISS evaluation concerns with internal validity existed. Threats to internal validity include selection, instrumentation, treatment diffusion, and sample size.

Selection. The EBISS group can be identified as a convenience sample. EBISS districts volunteered to participate in the EBISS initiative so attention was not given to equalizing the differences within the EBISS group (e.g., geographic location, capacity to implement EBISS) to ensure districts that volunteered to participate were different from those that did not. Parker (1993) brings attention to the fact that volunteers can be different from a non-volunteers and effect the outcome variable.

Instrumentation. Instrumentation within the EBISS group posed a threat to internal consistency. The DSSP was a fidelity measure that districts used to collect data on changes in adult behavior system-wide. However, a methodology to ensure inter-rater reliability (e.g., a protocol to ensure teams agreed on self-assessment scores) and inter-observer reliability (e.g., EBISS coordinator agreement on what a score of three fully in place represented) was not established prior to or during the four years when districts collected and submitted this data. Parker (1993) suggests that a threat to instrumentation exists when accuracy of the responses is violated. Without a protocol to ensure reliability of the assigned scores a threat to instrumentation existed. In addition, the DSSP technical adequacy report (Chaparro, Park, et al., 2011) did not provide evidence to support that the seven constructs of the DSSP are related to implementation fidelity and improved student reading outcomes at third grade (i.e., small or smaller than typical relationship between *district leadership and commitment* and *action planning with schools* and improvement in

91

the number of students meeting the state reading assessment at third grade ($p < .10$).

Treatment diffusion. Treatment diffusion also posed a threat to internal validity because some EBISS participants only received some of the EBISS treatment (not all districts attended all trainings). Attendance was not collected at EBISS trainings so there was no way to identify the districts that attended the PD trainings or if a district attended with a representative district leadership team. Understanding the PD dosage that each district did or did not receive, because of attendance, could have informed why no statistically significant difference between EBISS and non-EBISS and low-DSSP and high-DSSP was indicated. In addition, some EBISS districts received some or all of the treatment from another Oregon initiatives that focused on aspects of implementing RTI (e.g., full participation in DATA Project or selected participation in PBIS training). Shadish et al., (2002) suggest that lack of clarity about which treatment variable occurred first can yield confusion about which variable is the cause and which is the effect.

Sample size. The small sample size posed a significant limitation in this study for both research questions. A commonly accepted sample size to yield reasonably accurate *p* values for the rANOVA analysis is 30 subjects (Green & Salkind, 2011). Abelson (1995) agrees and stresses the fact that significance tests are highly dependent on the sample size of what is under study and he explains that when the sample size is small, strong and important effects can be non-significant. Therefore, *p*-values from significance tests will reflect both the sample size and the magnitude of the effects under study. Given the small sample size (25 districts per group) caution must be taken when interpreting the results that no significant difference between the two groups was detected. Furthermore, the second research question examined differences within the EBISS group. Splitting the

EBISS sample in half based on a district's average DSSP score in 2011 resulted in two very small groups (12 districts in one group and 13 in the other). Given the fact that the second research question was less than half the size of what is reasonable, caution must be used when interpreting the results that indicate there was no significant difference between high and low DSSP districts within the EBISS group.

External validity. External validity is concerned with the degree that research findings can be generalized across persons, time, and settings (Parker, 1993). When a study fails to obtain a sample that represents the population in a range of settings (e.g., states other than Oregon), and time (e.g., wait-listing treatment and measurement) and populations (e.g., equal sample size from urban-suburban and rural) external validity cannot be generalized to other persons, times, and settings.

Sampling. External validity is compromised when the participants in a sample are not randomly selected. This creates an interaction between selection and treatment. EBISS was a convenience sample of volunteers; therefore the results of this study cannot be generalized to non-volunteers. In other words, participants may have demonstrated characteristics that enhanced (strong district conditions) or diminished (weak district conditions) the effects of EBISS.

Interaction of setting and treatment. The interaction of setting and treatment also poses a threat to external validity; treatments in EBISS districts in Oregon may not be transferable to districts in other states. Implementing the EBISS model will be limited to settings that have similar priorities, knowledge, skills, and infrastructure to implement RTI for literacy. Districts in other states will have to consider the capacity Oregon has to

support district's implementation of RTI (PBIS, OrRTI, EBISS, DATA Project). Generalizability to districts outside of the State of Oregon is limited.

Interaction of testing within treatment and treatment diffusion. Finally, there may be an interaction of testing within treatment and treatment diffusion. While EBISS districts did not know they would be compared to non-EBISS districts they were aware of the fact that their DSSP scores were being evaluated for change over time. Saretsky (1972) termed this the John Henry effect; because once performance is made public the result may be competitive motivation as participants put forth extra effort to increase implementation. When interpreting the within group rANOVA caution is advised. Districts DSSP scores may have demonstrated an increase or decrease in the participants responsiveness or sensitivity to the treatment resulting in inflated scores on some DSSP measures. For example, some districts may have inflated some scores on the DSSP to demonstrate that the district was improving. At the same time, it is likely that treatment diffusion between EBISS and other major Oregon initiatives existed. EBISS districts also participated in OrRTI and received extensive TA and PD to implement RTI for literacy. In addition all EBISS districts participated in PBIS and some in the DATA Project and CLASS that stressed the ongoing use of data to problem solve, develop solutions and evaluate progress overtime.

Future Studies

EBISS sought to change district leadership behaviors to improve student outcomes, however in the final analysis of the EBISS initiative there were no statistically significant *between (p =.524) or within (p = .312)* group differences indicated in the gains in the percent of students passing the OAKS Reading at third grade across time. While

EBISS was not designed as a research study, future efforts to teach district how to implement RTI in a common district framework would benefit from the development of a research design (prior to implementation) that would more closely detect the differences between the treatment (EBISS) and the control (non-EBISS) conditions. To this end, future EBISS initiatives could rely on recent studies that were aimed at changing district leadership behaviors and develop a research design that captures strong leadership behaviors that can be correlated with improved student outcomes and scaled up in other Oregon districts (Fuchs & Fuchs, 2006; Leithwood & Jantzi, 2008; Grissom et al., 2012; Slavin et al., 2013). Therefore, I propose four different future studies that will increase the sensitivity of the independent variable (IV) to more accurately measure the dependent variable (proximal and distal).

Future Study One – Proximal Measure of Student Level Outcomes

A non-equivalent pretest-posttest matched comparison design would examine student growth over time by the proportion of students who were at benchmark on the Oral Reading Fluency (ORF) measure from time one to time two. This design would seek to answer the question did implementation of EBISS TA and PD (IV) result in an increase in the proportion of students in EBISS districts meeting benchmark on the ORF measure (DV) at third grade, over and above that of students in non-EBISS districts. Measuring proximal student growth would increase the sensitivity of the student outcome measure if sufficient time was allowed between pre and post-test measures (Cook, 2000). A proximal measure would also be more sensitive detecting if there were a significant relationship between district (board, district and school administrators behaviors) and changes in student outcomes (Grissom et al., 2012; Leithwood & Jantzi, 2008). However,

95

detecting if there were a significant relationship between district behaviors and student outcomes would be dependent on the degree of sensitivity in the IV measures (Grissom et al., 2012; Leithwood & Jantzi, 2008).

Future Study Two – Increase Sensitivity of Independent Variable

A non-equivalent pretest-posttest matched comparison design would examine student growth over time and answer the question: Did implementation of EBISS TA and PD (IV) result in an increase in district RtI skills and RtI beliefs (blocking variables) and demonstrate an increase in the proportion of students meeting benchmark on the ORF (DV) at third grade for EBISS districts, over and above that of non-EBISS districts from time one to time two. Revising the EBISS intervention components could increase the sensitivity of the IV. For example: (a) providing an introduction to EBISS program philosophy and required components prior to commitment of participation, (b) administering the RTI skills and beliefs survey to all EBISS district staff (January and May of each year), and (c) strengthening the EBISS PD and TA data based decision-making component (see Methods Summary of Yearly EBISS Professional Development 2007-11) by providing training and follow-up TA on how to use a problem solving model when analyzing data.

Introduce EBISS prior to district commitment. When districts agreed to participate in the EBISS initiative they did not know they were making a sustained commitment to the establishment of a district team that would reorganize and align policy to fund and implement the critical components of RTI district-wide. I propose providing districts with a one-day introduction to EBISS training. An introduction to EBISS would allow districts to make an informed decision regarding their readiness, willingness and

96

capacity to commit to EBISS goals, priorities and requirements (e.g., attend PD, transfer PD to practice in the district). This would allow districts the opportunity to evaluate their capacity (skills) and commitment (belief) that participation in EBISS may result in improved student outcomes.

Collect RtI skills and beliefs survey data. I recommend that Florida's RtI Skills and Beliefs Survey (Stockslager et al., 2013) be administered to all staff (board, administration, teaching and support staff) in January and May of each year to measure if staff skill and belief to implement RTI changed over time as a result of EBISS PD and TA. Measuring skills and beliefs would serve to identify what Cook, (2000) identifies as mediating processes (implementation quality, change in behavior). This would serve as an important non-test measure (Grissom et al., 2012) to guide the on-going development of differentiated PD and follow-up TA, as well as the degree to which adult skill and belief was improving in order to implement RTI in practice with fidelity (Gersten et al., 2008).

We know that district must communicate their belief that all students can achieve (Rorer et al., 2008) and commit to developing the skills of principals who in turn support teachers effective implementation of practices to improve student's outcomes (Lane, 2009; Leithwood & Jantzi, 2008; Sadler & Sugai, 2009). Identifying the skills and beliefs of adults across the system would provide districts with a clear understanding of their capacity, as well as their will and commitment to implement the critical components of RTI system-wide.

Strengthen data based decision-making. Finally, in reflection of the Rorer et al. (2008) study a major component that was missing in the design of EBISS PD and TA

was a steadfast and explicit focus *on maintaining an equity focus and owning past inequities*. A steadfast and explicit focus can be identified as a process where leaders use data to identify need and implement programs and practices that address the identified need (Slavin et al., 2013; Todd et al., 2011). Providing a two day training to teach district how to examine the data specific to their sub-group populations (e.g., economically disadvantaged, students of color, students in special education) using the TIPS model (Todd et al., 2011) could increase districts awareness of the needs of students and the practices that must be put in place to meet the need (Slavin et al., 2013). At the same time accurate measurement of a district's systems and adult behaviors would inform the differentiated support individual districts would require to improve student outcomes.

Future Study Three – Conduct Validity Study of District System Support Plan

I propose the development of a validity study of the DSSP survey; an important non-test measure used in the current EBISS initiative (Grissom et al., 2012). The validity study would revise and strengthen weak constructs within the current DSSP to increase the validity (and reliability) of the survey. I recommend using the data from the DSSP technical adequacy report (Chaparro et al., 2011) to revise weak constructs within the current DSSP. For example, a review of other district leadership measures could be conducted to strengthen the content validity of the current DSSP to ensure the data identifies the difference between weak and strong leadership behaviors (Leithwood & Jantzi, 2008). Furthermore, given the small number of internal components on some of the current DSSP constructs (e.g., funding only has two items related to funding of DSSP goals and curricula) a review of other leadership measures could inform the addition of items that would increase the construct validity of a revised and new DSSP survey. For

example, revisions could include strengthening the funding construct of the current

DSSP. Items could be added to the funding construct of the new DSSP to make explicit

reference to equitable resource allocation based on need (Rorrer et al., 2008; Slavin et al.,

2013). In addition, reliability of the DSSP scores could be strengthened.

Reliability of the DSSP scores is suspect. Coordinators did not receive training to

calibrate DSSP scoring across districts and teams did not have a common system for

scoring the DSSP. I recommend the development of a system of training EBISS

coordinators to calibrate the scoring of the new DSSP to increase the likelihood that there

is common agreement on what a score of, for example, one or three means in practice

(inter-observer reliability). Of equal importance is inter-rater reliability. I propose the

development of a common protocol for how teams would score the DSSP (e.g., all team

members score the measure individually and then discuss scores and agree on the

assignment of a common score for each item). Taken as a whole, using proximal data

(versus distal data) to measure student outcomes, increasing the sensitivity of the IV and

conducting a validity study of the DSSP would strengthen the design of the study in

question. However, consideration must be given to the fact that federal and state

accountability measures require district to improve student outcomes through the

measurement of distal outcomes (ESEA, 2010).

Future Study Four – Increase Sensitivity of Distal Measurement

Three potential problems can exist when using distal data (OAKS scores) to

evaluate proximal effects (district implementation effectiveness): theoretical ambiguity,

potential bias, and reliability (Grissom et al., 2012). In addition, district percentages are

not equal interval data, so each data point will represent a group with different means and

99

different variances (Jaeger, 2008; Stallings, Robbins, Presbrey, & Scott, 1986).

Therefore, to alleviate the problem associated with using district percentages to measure

student outcomes, I recommend that future studies use actual student scores to detect

differences in the groups (EBISS and non-EBISS) OAKS scores at third grade (pre and

post), because individual student scores are more sensitive to change. A non-equivalent

pretest-posttest matched comparison design would answer the question: Is there a mean

difference between the number of student's passing the third grade OAKS for non-EBISS

and EBISS districts from time one to time two?

Strengthening the Matched Comparison Sample Design

In order to test for interaction effects and differences between EBISS and non-

EBISS a matched comparison of non-EBISS districts should be identified prior to the

intervention to ensure the groups are equivalent by demographics (e.g., location, size,

student, teacher demographics), outcome measures (proximal or distal) and levels of

implementation (i.e., DSSP, RtI Skills and RtI-Beliefs surveys). Ensuring groups were

matched by demographics, the outcome measure, and levels of implementation (at pre-

test) could inform whether EBISS TA and PD resulted in increased implementation

scores and improved student outcomes for the EBISS group, over and above the non-

EBISS group. Therefor, a non-EBISS posttest should also be administered at the end of

the initiative. At the same time specific behaviors (items within each construct of skills,

beliefs, DSSP measure) could be isolated and linked with strong district implementation

behaviors that effected the DV (proximal or distal) using the evaluation theory of

Leithwood and Jantzi (2008). Measuring the skills, beliefs, and DSSP scores pre and post

implementation for both groups could detect the degree to which other non-EBISS

100

interventions had an effect on districts RTI skills, beliefs and DSSP scores. This design would serve to answer the question: Were EBISS and non-EBISS groups matched on pre-test measures and did EBISS district's RTI *skills,* RTI *beliefs* and DSSP scores increase across time, over and above non-EBISS?

As districts consider implementing RTI for early literacy (and the blending of PBIS) they will want to be cognizant of the tremendous will and commitment this level of district-wide reform will entail (Rorrer et al., 2008). Districts must commit to redesigning and aligning systems that focus on the resources staff and students will require to close the research-to-practice-to-policy gap (Carnine, 1999; Fixsen et al., 2013).

Summary

District leaders are held accountable for improving student outcomes and recent scholarship on the role of district in this important work is gaining the attention of researchers and practitioners alike (Cook, 2000; Fixsen et al., 2009; Fixsen et al., 2013; Grissom et al., 2012; Honig et al., 2010; Lane, 2009; Leithwood & Jantzi, 2008; Rorrer et al., 2008; Slavin et al., 2013). The findings from this program evaluation on the EBISS initiative implies that changing district leadership behaviors (organizational systems, resource allocation, instructional leadership skills of all staff, and policy) takes tremendous will and commitment from district leadership to improve the outcomes for all students. Nonetheless, given the longitudinal data from the National Center for Educational Statistics (NCES) that reported the gap in reading scores widened for students from 1971 to 2009 with no significant changes for Black and Hispanic students from 2009 to 2011 one can appreciate the attention that is needed to maintain an equity

focus when implementing RTI initiatives. Rorrer et al. (2008) suggest that districts must take a moral response to *owning past inequities* and have the courage to acknowledge low achievement and seek viable solutions to change that moves beyond the *all students can learn rhetoric*. However, this moral response, even when others oppose (Firestone, 1989) will take courageous leaders at the district, school, and community level because advocating for the needs of every student a district serves is often riddled with conflict (Rorrer et al., 2008). Yet, the students educators were called to serve rely on courageous leaders to ensure that they receive equitable access to a quality education that prepares them for successful adult outcomes.

EBISS DISTRICT SYSTEMS SUPPORT PLAN

EBISS DISTRICT SYSTEMS SUPPORT PLAN (Please submit to EBISS Coordinator by _____)
DIRECTIONS: This plan is designed to assist district teams to identify, implement, and sustain the essential features of effective, systems-level academic and behavior support. The features include: 1) District Leadership and Commitment, 2) Action Planning with EBISS schools , 3) Coordination and Coaching, 4) Professional Development and Training Capacity, 5) On-going Assessment and Evaluation Systems, 6) Visibility and Political Support, and 7) Funding. This plan should be completed by a team of district and school administrators representing general and special programs. Teams should briefly review and brainstorm each feature and item. A four-point rating scale is provided for each item: Completely in Place, Mostly In Place, Partly in Place, and Not in Place. Space is included within each item for current score and three additional progress scores, and a section is provided for tracking overall progress. Teams are encouraged to use this District Systems Support Plan as a working document. Districts with many features requiring actions should review and update the plan quarterly. Districts with many features "in place" should focus on sustainability and update the plan annually, at a minimum. Completion of the support plan should take approximately one hour. For any items that are not completely in place, districts should identify and describe needed actions briefly and specifically on the EBISS Action Planning Guide (p. 7). Space is provided on the guide for including staff member(s) who will be responsible for each action, completion date, and any documentation used to determine the item response (see guide for further instructions).

PROGRESS ASSESSMENT	Initial Score	Progress Score	Progress Score	Progress Score
EBISS FEATURES (Enter the total scores/percentages for each feature in the space below, and percentages in the columns to the right)	**Date**	**Date**	**Date**	**Date**
1. Leadership & Commitment (__/24 = __%):				
2. Action Planning w/Schools (__/9 = __%):				
3. Coordination & Coaching (__/9 = __%):				
4. Professional Development & Training Capacity (__/6 = __%):				
5. Ongoing Assessment & Evaluation Systems (__/9 = __%):				
6. Visibility & Political/Stakeholder Support (__/6 = __%):				
7. Funding (__/6 = __%):				
Overall Average (% Scores __/ 7):				

EBISS FEATURES and RATING SCALES

1. DISTRICT LEADERSHIP & COMMITMENT

a. District has a leadership team for EBISS that includes district-level administrators (representing C & I, Sped, Title, ELL, Counseling, etc.) and school leaders, as needed. *(Referred to as "Team" in the rest of this document).*

3 = Completely in Place   2 = Mostly in Place (e.g. some recruiting to do)   1 = Partly in Place (e.g. substantial recruiting to do)   0 = Not in Place

      Initial Score:____   Progress Score:____   Progress Score:____   Progress Score:____

b. Team meetings are scheduled/held at least every other month throughout the year.

3 = Completely in Place   2 = Mostly in Place (e.g. set schedule but fewer meetings planned/held)   1 = Partly in Place (e.g. schedule and meetings planned/held sporadically)   0 = Not in Place

      Initial Score:____   Progress Score:____   Progress Score:____   Progress Score:____

d. District has a Continuous Improvement Plan that emphasizes evidence-based academic and behavior practices across multi-tiers (i.e., addresses needs of all students).

3 = Completely in Place (practices identified for reading and behavior at all tiers)   2 = Mostly in Place (e.g. effective practices identified for reading and behavior for SOME tiers)   1 = Partly in Place (e.g. effective practices for reading OR behavior for some tiers)   0 = Not in Place

Initial Score:____ Progress Score:____ Progress Score:____ Progress Score:____

e. District has policies and a procedural manual for LD identification based on a multi-tiered, response to intervention (RTI) framework.

3 = Completely in place, 2 = Mostly in Place (e.g. policies and procedures are in last draft) 1 = Partly in Place (e.g. policies and procedures are in initial draft) 0 = Not in Place

Initial Score:____ Progress Score:____ Progress Score:____ Progress Score:____

f. District provides time and resources to school teams to engage in collaborative teamwork a minimum of: (a) twice/year to analyze and update systems and action plans, and (b) once/month, to meet with grade level teams to analyze student progress and adjust interventions.

3 = Completely in Place 2 = Mostly in Place (e.g. school teams meet several times a year but could be more consistent),   1 = Partly in Place (e.g. school teams meet for annual planning but monthly meetings not consistent, or vice versa)
0 = Not in Place
Initial Score:____ Progress Score:____ Progress Score:____ Progress Score:___

1. DISTRICT LEADERSHIP & COMMITMENT, cont'd.

g. District practices and procedures encourage hiring personnel with skills that will contribute to EBISS systems sustainability.

3 = Completely in Place 2 = Mostly in Place (e.g. no formal policy but practice is consistent) 1 = Partly in Place (e.g. no formal policy, practices are inconsistent)   0 = Not in Place

Initial Score:____   Progress Score:____   Progress Score:____   Progress Score:____

h. Team reviews new programs and initiatives to rate their effectiveness and compatibility with EBISS priorities before adopting (i.e., they are evidence-based <u>and</u> data shows they are needed).

3 = Completely in Place   2 = Mostly in Place (e.g. most initiatives and programs are reviewed for their effectiveness and EBISS compatibility)   1 = Partly in Place (e.g. some initiatives and programs are reviewed for their effectiveness and EBISS compatibility)   0 = Not in Place

Initial Score:____   Progress Score:____   Progress Score:____   Progress Score:____

2. ACTION PLANNING with EBISS SCHOOLS

a. District implementation plan defines goals for the number of schools to implement EBISS priorities (i.e., elementary school-wide reading, secondary school-wide reading, K-12 behavior, and response to intervention for LD identification) over the next 4 years (see EBISS priorities form).

3 = Completely in Place   2 = Mostly in Place (e.g. working out some details)   1 = Partly in Place (e.g. working out many details)   0 = Not in Place

Initial Score:____   Progress Score:____   Progress Score:____   Progress Score:____

b. Team uses tools to help designated schools analyze strengths and needs related to EBISS priorities.

> EXAMPLE: Priority: SW Reading   Tools: Planning & Evaluation Tool for School-wide Reading-Revised-PET-R, DIBELS reports, Literacy Plan template, etc.)

3 = Completely in Place   2 = Mostly in Place (e.g. we have evaluation tools for all areas but need some help using them),   1 = Partly in Place (e.g. we do not have tools for all areas and/or need considerable help using the tools)   0 = Not in Place

Initial Score:____   Progress Score:____   Progress Score:____   Progress Score:____

c. Team has identified time and a strategy for working with designated schools on their action plans for EBISS priorities. School Action plans are in response to data from priority specific tools (see number 2)

3 – Completely in Place   2 = Mostly in Place (e.g. working out some details)   1 = Partly in Place (e.g. working out several details)   0 = Not in Place

Initial Score:____   Progress Score:____   Progress Score:____   Progress Score:____

3. COORDINATION & COACHING

a. Team has identified specific individual(s) to coordinate EBISS activities and plans (e.g., facilitate Team decisions, communication with schools, and professional development).

3 = Completely in Place   2 = Mostly in Place (e.g. FTE needs to be increased some)   1 = Partly in Place (e.g. FTE needs to be increased substantially)   0 = Not in Place

Initial Score:____   Progress Score:____   Progress Score:____   Progress Score:____

3. COORDINATION & COACHING, cont'd.

b. Team has identified specific individual(s) to provide coaching in literacy and behavior to schools implementing EBISS priorities. This person(s) has been or is receiving training in each of these areas.

3 = Completely in Place   2 = Mostly in Place (e.g. FTE needs to be increased some)   1 = Partly in Place (e.g. FTE needs to be increased substantially)   0 = Not in Place

Initial Score:____   Progress Score:____   Progress Score:____   Progress Score:____

c. Team and schools have identified specific individual(s) to serve as site-based experts to support EBISS activities (literacy and behavior specialists, team facilitators, etc.).

3 = Completely in Place   2 = Mostly in Place (e.g. most schools have expertise; we fill in the gaps when needed)   1 = Partly in Place (e.g. schools have literacy but not behavior expertise, or vice versa)   0 = Not in Place

Initial Score:____   Progress Score:____   Progress Score:____   Progress Score:____

4. PROFESSIONAL DEVELOPMENT & TRAINING CAPACITY

a. Release time is allocated for Coordinators, Coaches, and Site-Based Experts to participate in State-coordinated training (e.g., attending PBS conference, EBISS trainings, Literacy Outreach).

3 = Completely in Place   2 = Mostly in Place (e.g. budget accommodates most needs)   1 = Partly in Place (e.g. budget accommodates some needs)   0 = Not in Place

Initial Score:____   Progress Score:____   Progress Score:____   Progress Score:____

5. ONGOING ASSESSMENT & EVALUATION SYSTEMS (Reading & Behavior)

a. District has identified specific measures for screening and progress monitoring of students' reading and behavior skills (e.g., state assessments, DIBELS, other curriculum-based measures, office discipline referrals, counseling referrals, and attendance reports).

3= Completely in Place   2 = Mostly in Place (e.g. all measures are identified but more training on use is needed) ,   1 = Partly in Place (e.g. not all measures are in place and/or substantial training on use is needed)   0 = Not in Place

Initial Score:____   Progress Score:____   Progress Score:____   Progress Score:____

5. ONGOING ASSESSMENT & EVALUATION SYSTEMS (Reading & Behavior), cont'd.

b. District provides schools with usable and easy-to-read reports on academic and behavior data for decision-making (e.g., state assessment reports, DIBELS, SWIS, easyCBM, Aimsweb, etc.).

3 = Completely in Place   2 = Mostly in Place (e.g. schools have all the data but cannot always access it quickly)   1 = Partly in Place (e.g. schools have some of the data access they need)   0 = Not in Place

| Initial Score:____ | Progress Score:____ | Progress Score:____ | Progress Score:____ |

c. Team routinely evaluates: (a) implementation of this Systems Support Plan; (b) implementation of school action plans; and (c) impact of both plans on student achievement.

3 = Completely in Place   2 = Mostly in Place (e.g. evaluation occurs but could be more routine and efficient)   1 = Partly in Place (e.g. evaluation occurs but needs to be substantially more routine and efficient)   0 = Not in Place

| Initial Score:____ | Progress Score:____ | Progress Score:____ | Progress Score:____ |

6. VISIBILITY & POLITICAL/STAKEHOLDER SUPPORT

a. Superintendent and School Board receive at least semi-annual progress reports from the Team.

3 = Completely in Place   2 = Mostly in Place (e.g. one formal report is planned)   1 = Partly in Place (e.g. informal reports are planned)   0 = Not in Place

| Initial Score:____ | Progress Score:____ | Progress Score:____ | Progress Score:____ |

b. Team shares EBISS accomplishments at a regional level at least semi-annually in a public forum (e.g., rotary, ESD, parent groups).

3 = Completely in Place   2 = Mostly in Place (e.g. one formal report is planned)   1 = Partly in Place (e.g. informal reports are planned)   0 = Not in Place

| Initial Score:____ | Progress Score:____ | Progress Score:____ | Progress Score:____ |

7. FUNDING

a. District budget includes funding dedicated to the implementation of this systems support plan, the district action plan, and school action plans.

3 = Completely in Place   2 = Mostly in Place (e.g. funding is dedicated and nearly adequate)   1 = Partly in Place   (e.g. funding is dedicated but much less than adequate)   0 = Not in Place

| Initial Score:____ | Progress Score:____ | Progress Score:____ | Progress Score:____ |

b. District budget includes funds for academic and behavior core curricula, supplemental, and targeted intervention materials.

3 = Completely in Place   2 = Mostly in Place (e.g. funding is dedicated and nearly adequate)   1 = Partly in Place (e.g. funding is dedicated but   not nearly adequate)   0 = Not in Place

| Initial Score:____ | Progress Score:____ | Progress Score:____ | Progress Score:____ |

APPENDIX B

DESCRIPTIVE STATISTICS OF THE DSSP

Construct	Item	Description	*Mean*	*SD*	*Alpha*
District Leadership & Commitment	lead1	Leadership team	2.62	0.63	.84
	lead2	Scheduled team meeting	2.47	0.83	
	lead3	Team meeting attendance	2.36	0.84	
	lead4	Continuous improvement plan	2.26	0.78	
	lead5	LD identification	1.59	1.16	
	lead6	Collaborative teamwork	2.33	0.77	
	lead7	Hiring personnel	1.89	0.93	
	lead8	Review of new programs	1.93	0.95	
Action Planning	act1	Goals for the number of schools	1.84	0.91	.79
	act2	Tools to analyze strength and needs	2.19	0.75	
	act3	Time and strategy to work with schools	1.74	0.88	
Coordination & Coaching	coach1	Individual(s) to coordinate activities	2.54	0.77	.83
	coach2	Individual(s) to provide coaching	2.29	0.84	
	coach3	Individual(s) to serve as site-based experts	2.18	0.80	
Professional Development & Training Capacity	pd1	Release time for state-coordinated training	2.53	0.74	.76
	pd2	Training opportunities	2.05	0.79	
Ongoing Assessment & Evaluation Systems	eval1	Measures for screening and progress monitoring	2.42	0.74	.79
	eval2	Reports for decision-making	2.35	0.80	
	eval3	Routine evaluation	1.62	0.90	
Visibility & Political Stakeholder Support	visible1	Progress report for superintendent and board	1.97	1.01	.74
	visible2	Share of team's accomplishments	1.73	1.09	
Funding	fund1	Funding for the systems support plan	1.82	0.94	.82
	fund2	Funding for tiered intervention materials	1.92	0.89	

REFERENCES CITED

Abelson, R.P. (1995). *Statistics as principled argument*. Hillsdale, NJ: Erlbaum.

Aladjem, D. K., & Borman, K. M. (2006). *Examining comprehensive school reform*. Washington, DC: Urban Institute Press.

Algozzine, B., Horner, R. H., Sugai, G., Barrett, S., Dickey, S. R., Eber, L., Kincaid, D., et al. (2010). *Evaluation blueprint for school-wide positive behavior support*. Eugene, OR: National Technical Assistance Center on Positive Behavior Interventions and Support. Retrieved from www.pbis.org

American Educational Research Association, American Psychological Association, National Council on Measurement in Education, Joint Committee on Standards for Educational, & Psychological Testing (US). (1999). *Standards for educational and psychological testing*. Amer Educational Research Assn.

Babbie, E. (2007). The Practice of Social Research. Belmont: Thomson Wadsworth.

Baker, S. K., Fien, H., & Luft Baker, D. (2010). Robust reading instruction in the early grades: Conceptual and practical issues in the integration and evaluation of tier 1 and tier 2 instructional supports. *Focus on Exceptional Children*, *42*(9).

Baker, S. K., Smolkowski, K., Smith, J. M., Fien, H., Kame'enui, E. J., & Beck, C. T. (2011). The impact of Oregon Reading First on student reading outcomes. *The Elementary School Journal*, *112*(2), 307-331.

Carnine, D. (1997). Bridging the research-to-practice gap. *Issues in educating students with disabilities*, 363-373.

Carnine, D. (1999). Perspective: Campaigns for Moving Research into Practice. *Remedial and Special Education*, *20*(1), 2-6.

Chaparro, E. A., Park, Y., Baker, S. K., & Ryan-Jackson, K. M. (2011). *District system support plan: A district level self-assessment tool*. (Report No. 1104). Eugene, OR: Center on Teaching and Learning, University of Oregon.

Chaparro, E. A., Ryan-Jackson, K., Baker, S. K., & Smolkowski, K. (2012). Effective behavioural and instructional support systems: An integrated approach to behavior and academic support at the district level. *Advances in School Mental Health Promotion, 5*, 161-176.

Chaparro, E. A., Smolkowski, K., Baker, S. K., Hanson, N. & Ryan-Jackson, K. M. (2012). *A* Model for System-Wide Collaboration to Support Integrated Social Behavior and Literacy Evidence-Based Practices. *Psychology in the Schools (5)* 161-176.

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

Cook, T. D. (2000). The false choice between theory-based evaluation and experimentation. *New directions for evaluation*, *2000*(87), 27-34.

Darling-Hammond, L. (1999). *Teacher quality and student achievement: A review of state policy evidence*. Seattle, WA: Center for the Study of Teaching and Policy, University of Washington.

Dow, A. W., Salas, E., & Mazmanian, P. E. (2012). Improving quality in systems of care: Solving complicated challenges with simulation-based continuing professional development. *Journal of Continuing Education in the Health Professions*, *32*(4), 230-235.

Eccles, J., Wigfield, A., Harold, R. D., & Blumenfeld, P. (1993). Age and gender differences in children's self-and task perceptions during elementary school. *Child development*, *64*(3), 830-847.

Ervin, R. A., Schaughency, E., Goodman, S. D., McGlinchey, M. T., & Matthews, A. (2006). Merging research and practice agendas to address reading and behavior school-wide. *School Psychology Review, 35*, 198–223.

Fiester, L. (2010). Early Warning! Why Reading by the End of Third Grade Matters. KIDS COUNT Special Report. *Annie E. Casey Foundation*.

Firestone, W. (1989). Educational policy as an ecology of games. *Educational Researcher*, *18*(7), 18-24.

Fixsen, D. L., & Blase, K. (2009). Implementation: The missing link between research and practice *NIRN Implementation Brief #1*. Chapel Hill, NC: The University of North Carolina, FPG, NIRN.

Fixsen, D. L., Blase, K. A., Horner, R., & Sugai, G. (2009). Intensive Technical Assistance. Scaling-Up Brief. Number 2. *FPG Child Development Institute*.

Fixsen, D., Blase, K., Metz, A., & Van Dyke, M. (2013). Statewide implementation of evidence-based programs. *Exceptional Children*, *79*(2), 213-230.

Flynn, J. M., & Rahbar, M. H. (1994). Prevalence of reading failure in boys compared with girls. *Psychology in the Schools*, *31*, 66-71.
Fowler, F.C. (2013). *Policy studies for educational leaders*. San Francisco: Pearson

Foster, W. A., & Miller, M. (2007). Development of the literacy achievement gap: A longitudinal study of kindergarten through third grade. *Language, Speech, and Hearing Services in Schools*, *38*, 173.

Friedkin, N.E., and Necochea, J. 1988. "School System Size and Performance: A Contingency Perspective." *Educational Evaluation and Policy Analysis* 10: 237–249.

Fuchs, D., & Fuchs, L. S. (2006). Introduction to Response to Intervention: What, why, and how valid is it?. *Reading Research Quarterly*, *41*(1), 93-99.

Gersten, R., Complon, D., Santoro, L. E., Dimino, J., Linan-Thompson,S., & Tilly, D. (2008). *Response to intervention (RTI) & multitier intervention for reading in the primary grades*. Washington, DC: U.S. Department of Education, Institute for Education Sciences.

Gersten, R., Dimino, J., Jayanthi, M., Kim, J. S., & Santoro, L. E. (2010). Teacher Study Group Impact of the Professional Development Model on Reading Instruction and Student Outcomes in First Grade Classrooms. *American Educational Research Journal*, *47*(3), 694-739.

Green, S., & Salkind, N. J. (2011). *SPSS QuickStarts*. Pearson Higher Ed.

Grissom, J. A., Kalogrides, D., & Loeb, S. (2012). *Using student test scores to measure principal performance* (No. w18568). National Bureau of Economic Research.

Greenwood, C. R., & Abbott, M. (2001). The research to practice gap in special education. *Teacher Education and Special Education: The Journal of the Teacher Education Division of the Council for Exceptional Children*, *24*(4), 276-289.

Hannaway, J., & Kimball, K. (1998). *Big isn't always bad: School district size, poverty, and standards-based reform*. US Department of Education, Office of Educational Research and Improvement, Educational Resources Information Center.

Hanushek, E. A., Kain, J. F., Markman, J. M., & Rivkin, S. G. (2003). Does peer ability affect student achievement?. *Journal of applied econometrics*, *18*(5), 527-544.

Hawk, P., Coble, C.R., & Swanson, M. (1085). Certification: It does matter. *Journal of Teacher Education, 36* (3),13-15.

Hogg, R. V., & Craig, A. T. (1995). *Introduction to Mathematical Statistics*, 5th ed. New York: Macmillan.`

Honig, M. I., Copland, M. A., Rainey, L., Lorton, J. A., & Newton, M. (2010). Central office transformation for district-wide teaching and learning improvement. Retrieved from The Wallace Foundation website: http://www.wallacefoundation.org/knowledge-center/school-leadership/district-policy-and-practice/Documents/Central-Office-Transformation-District-Wide-Teaching-and-Learning.pdf

Sugai, G., & Horner, R. (2002). The evolution of discipline practices: School-wide positive behavior supports. *Child & Family Behavior Therapy*, *24*(1-2), 23-50.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logic mixed models. *Journal of memory and language*, *59*(4), 434-446.

Joyce, B. R., & Showers, B. (2002). Student achievement through staff development. *ASCD*.

Kame'enui, E. J., & Simmons, D. C. (1998). Beyond effective practice to schools as host environments: Building and sustaining a school-wide intervention model in reading. *OSSC Bulletin, 41*, 3–24.

Kline, R. (2011). Principles and practices of structural equation modeling (3rd ed.). New York: Guilford.

Knight, J. (2009). Coaching. *Journal of Staff Development*, *30*(1), 18-22.

Klingner, J. K., Boardman, A. G., & Mcmaster, K. L. (2013). What Does It Take to Scale Up and Sustain Evidence-Based Practices? *Exceptional Children*, *79*(2), 195-211.

Knapp, M. S. (2008). How can organizational and sociocultural learning theories shed light on district instructional reform? *American Journal of Education*, *114*(4), 521-539.

Lane, B. (2009). Rapid District Improvement. *Center on Innovation and Improvement*. Retrieved from: http://www.sde.state.nm.us/div/psb/dl10/Exploring_the_Pathway_to_Rapid_District_Improvement.pdf

Leithwood, K., & Jantzi, D. (2008). Linking leadership to student learning: The contributions of leader efficacy. *Educational administration quarterly*, *44*(4), 496-528.

Mathis, W. J. (2003, March). *Equity and adequacy challenges in rural schools and communities.* Paper presented at the annual meeting of the American Education Finance Association, Orlando, FL. (ERIC Document Reproduction Service No. ED 478 057)

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist, 50*, 741-749.

McCall, M.S. (n.d.). Technical review of the Oregon Department of Education assessment system. Northwest Evaluation Association.

McIntosh, K, Chard, D.J., Boland, J.B. & Horner, R.H. (2006). Demonstration of combined efforts in school-wide academic and behavioral systems and incidence of reading and behavior challenges in early elementary grades. *Journal of Positive behavioral Interventions, 8,* 146-154

National Center for Educational Statictics. (2010). *National Center of Educational Progress: 1971-2008 Long-term trend data.* Retrieved from http://nationsreportcard.gov/ltt_2008/ltt0003.asp

National Center for Educational Statictics. (2011). *The Nations Report Card: Reading 2011.* Retrieved from http://nces.ed.gov/nationsreportcard/pdf/main2011/2012457.pdf

National Center for Response to Intervention. (2010). *Essential components of RTI: A closer look at Response to Intervention.* Retrieved from: http://www.rti4success.org/sites/default/files/rtiessentialcomponents_042710.pdf

Newton, J. S., Horner, R. H., Algozzine, B., Todd, A. W., & Algozzine, K. (2012). A randomized wait-list controlled analysis of the implementation integrity of team-initiated problem solving processes. *Journal of School Psychology*, *50*(4), 421-441.

No Child Left Behind Act (NCLB) Act of 2001, Pub. L. No. 107-110, §115, Stat. 1425 (2002).

Odom, S. L., Brantlinger, E., Gersten, R., Horner, R. H., Thompson, B., Harris, K. R. (2005). Research in Special Education: Scientific Methods and Evidence-Based Practices. *Exceptional Children* 71, 137-148.

Oregon Department of Education. (2000). Oregon statewide reading and literature, mathematics, science, and social science multiple choice assessments. Accessed September 29, 2011, from http://ww.ode.state.or.us/asmt/administration/manuals/2000

Oregon Department of Education. (2006). Doran, H., and Cohen, J. Oregon Technical Report.. Accessed October 29, 2013, from http://ww.ode.state.or.us/asmt/administration/manuals/2000

Oregon Department of Education. (2000). Closing the achievement gap: Oregon's plan for success for all students. Salem, Or.: Oregon Department of Education. Retrieved September 29, 2011, from http://www.ode.state.or.us/asmt/administration/manuals/2000

Oregon Department of Education. (2011). Meeting on OAKS reading at grade 3. Retrieved March 03, 2013, from http://www.ode.state.or.us/search/page/?id=3296

Oregon Department of Education. (2008). Oregon Eduvctional Leadership and Administrator Standards. Retrieved January 09, 2014, from http://www.ode.state.or.us/wma/teachlearn/educatoreffectiveness/or-admin-standards.pdf

Parker, R. M. (1993, March). Threats to the validity of research. *Rehabilitation Counseling Bulletin*. p. 130.

Putnam, D. (2008). Guiding RTI System Implementation: The Oregon Experience. *The School Administrator, 8* (65)

Reauthorization, E. S. E. A. (2010). A Blueprint for Reform. *2010-11-29. http://www2. ed. gov/policy/elsec/leg/blueprint/index, html*.

Robinson, V. M., Lloyd, C. A., & Rowe, K. J. (2008). The impact of leadership on student outcomes: An analysis of the differential effects of leadership types. *Educational administration quarterly*, *44*(5), 635-674.

Rockoff, J. E. (2004). The impact of individual teachers on student achievement: Evidence from panel data. *The American Economic Review*, *94*(2), 247-252.

Rorrer, A. K., Skrla, L., & Scheurich, J. J. (2008). Districts as institutional actors in educational reform. *Educational Administration Quarterly*, *44*(3), 307-357.

Sadler, C. (2000). Effective behavior support implementation at the district level. *Journal of Positive Behavior Interventions*, *2*(4), 241-246.

Sadler, C., & Sugai, G. (2009). Effective Behavior and Instructional Support A District Model for Early Identification and Prevention of Reading and Behavior Problems. *Journal of Positive Behavior Interventions*, *11*(1), 35-46.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). Experimental and quasi-experimental designs for generalized causal inference.

Sharratt, L., & Fullan, M. (2006). Accomplishing districtwide reform. *Journal of School Leadership, 16*, 583-595.

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

Slavin, R. E., Cheung, A., Holmes, G., Madden, N. A., & Chamberlain, A. (2013). Effects of a Data-Driven District Reform Model on State Assessment Outcomes. *American Educational Research Journal*, *50*(2), 371-396.

Slavin, R., & Madden, N. A. (2011). Measures inherent to treatments in program effectiveness reviews. *Journal of Research on Educational Effectiveness*, *4*(4), 370-380.

Spillane, J.P., Halverson, R., & Diamond, J.B. (2001). Investigating school leadership practice: A distributed perspective. *Educational Researcher 30*, 23-28. doi: 10.3102/0013189X030003023

Stallings, J., Robbins, P., Presbrey, L., & Scott, J. (1986). Effects of instruction based on the Madeline Hunter Model on students' achievement: Findings from a follow-through project. *The Elementary School Journal*, *86*(5), 571-587.

Stewart, R. M., Benner, G. J., Martella, R. C., & Marchand-Martella, N. E. (2007). Three-tier models of reading and behavior. *Journal of Positive Behavior Interventions,* 9, 239–253. doi: 10.1177/10983007070090040601

Stockslager, K. M., Castillo, J. M., Hines, C. V., Batsche, G. M., & Curtis, M. J. (2013). The Florida Problem Solving/Response to Intervention Project.Sugai, G., & Horner, R. R.(2002). The evolution of discipline practices: School-wide positive behavior supports. *Child and Family Behavior Therapy, 24*, 23–51.

Sugai, G., & Horner, R. R. (2006). A promising approach for expanding and sustaining school-wide positive behavior support. *School Psychology Review*, *35*(2), 245.

Sugai, G., & Horner, R. H. (2009). Responsiveness-to-intervention and school-wide positive behavior supports: Integration of multi-tiered system approaches. *Exceptionality*, *17*(4), 223-237.

Swinnerton, J. (2007). Brokers and Boundary Crossers in an Urban School District: Understanding Central Office Coaches as Instructional Leaders. *Journal of School Leadership*, *17*(2), 195-221.

Todd, A. W., Horner, R. H., Newton, J. S., Algozzine, R. F., Algozzine, K. M., & Frank, J. L. (2011). Effects of team-initiated problem solving on decision making by schoolwide behavior support teams. *Journal of Applied School Psychology*, *27*(1), 42-59.

Togneri, W., & Anderson, S. E. (2003). Beyond Islands of Excellence: What Districts Can Do To Improve Instruction and Achievement in All Schools. A Project of the Learning First Alliance [and] A Leadership Brief.

Tucker, A., Edmondson, A., & Nembhard, I. (2005). *Contributions of know-what and know-how to performance improvement in complex service organizations.* Paper presented at the Academy of Health Meeting 22: abstract no. 3700, Boston, MA.

Tuttle, C. C., Teh, B. R., Nichols-Barrer, I., Gill, B. P., & Gleason, P. (2010). Student characteristics and achievement in 22 KIPP middle schools. *Washington, DC: Mathematica Policy Research*.

United States Department of Education. (2010). ESEA Blueprint for Reform. Washington D.C. Retrieved on November 17, 2012 from http://www2.ed.gov/policy/elsec/leg/blueprint/index.html

Vernez, G., Karam, R., Mariano, L. T., & DeMartini, C. (2006). Evaluating Comprehensive School Reform Models at Scale: Focus on Implementation. *RAND Corporation*.

Walker, H., Colvin, G., & Ramsey, E. (1995) *Antisocial behavior in public school: Strategies and best Practices.* CA: Brookes/Cole.

White, K. R. (1982). The relation between socioeconomic status and academic achievement. *Psychological Bulletin*, *91*(3), 461.