TEACHER DATA USE: IMPACT FROM INTERIM ASSESSMENTS ON STUDENT OUTCOMES

by

JOEL R SEBASTIAN

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

June 2020

DISSERTATION APPROVAL PAGE

Student: Joel R Sebastian

Title: Teacher Data Use: Impact from Interim Assessments on Student 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:

Mark Van Ryzin	Chair
Lina Shanley	Core Member
William Rhoades	Core Member
Kent McIntosh	Institutional Representative

and

Kate Mondloch	Interim Vice Provost and Dean of the Graduate School

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

Degree awarded June 2020

DISSERTATION ABSTRACT

Joel R Sebastian

Doctor of Education

Department of Educational Methodology, Policy, and Leadership

June 2020

Title: Teacher Data Use: Impact from Interim Assessments on Student Outcomes

Born from mandatory state-level high stakes assessment, more sources of student data are available to educators today than any other time in memory. School districts regularly employ some type of internal assessment system in order to understand how student populations are progressing towards expected outcomes. These assessments, often called interim assessments, are administered three to four times throughout the school year. How effectively teachers utilize these assessment data and its impact on student outcomes is the central focus of this study. This study utilized a quantitative design to understand if there is a predictive relationship between how teachers report the use of interim data and the student outcomes on year-end state-level tests.

CURRICULUM VITAE

NAME OF AUTHOR: Joel R Sebastian

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon, Eugene
Portland State University, Portland
Lewis and Clark College, Portland
California State University, Long Beach
California State University, Humboldt

DEGREES AWARDED:

Doctor of Education, 2020, University of Oregon
Master of Arts, Special Education, 1999, CSU, Long Beach
Bachelor of Science, Physical Education, 1996, CSU, Long Beach

AREAS OF SPECIAL INTEREST:

Assessment
Teaching and Learning
Capacity Building

PROFESSIONAL EXPERIENCE:

Principal, West Linn – Wilsonville School District, 2012 - Present
Co-Principal, Canby School District, 2011 - 2012
Principal, Canby School District, 2005 - 2011
Assistant Principal, Canby School District, 2002 - 2005
Teacher, Special Education, Oregon City School District, 1999 - 2002
Teacher, Special Education, Downey Unified School District, 1997 – 1999

GRANTS, AWARDS, AND HONORS:

Peace Educator of the Year, Wholistic Peace Institute, 2019

Administrator of the Year, Oregon Music Educator Association, 2011

ACKNOWLEDGMENTS

I wish to express sincere appreciation to my advisor and dissertation chair, Dr. Mark Van Ryzin. Your support and encouragement inspired me to reach further than I thought possible. It is hard to imagine that I could have completed this project without your ongoing support. I wish to thank my committee for their support and guidance with my project: Dr. Kent McIntosh, Dr. Lina Shanley, and Dr. Bill Rhoades. Also, a big thank you to Dr. Joanna Smith and Dr. Julie Alonzo, the hearts and souls of the EMPL program during my time at the University of Oregon.

I owe a great debt to my colleagues at Athey Creek Middle School and in the West Linn – Wilsonville School District. Particularly Dr. Barb Soisson, Dr. Kathy Ludwig, Khahn Doung, and all staff that participated in my study; this project simply would not have been possible without their support.

It is with a heart full of gratitude that I recognize my good friends, Kevin Egan and Michael Knapp, who never let me forget that I am a regular person attempting difficult things.

Finally and most importantly, thank you to my family. I am truly blessed to be married to Amy Sebastian. Her patience and support during this time was amazing and essential. Thank you to my sons; Ben for assisting me with the mathematics and statistics over my time in the program, and Isaac for providing appropriate distraction when I needed it the most (should we go fishing?). I love you all!

For my parents, to whom I owe so much.

TABLE OF CONTENTS

Chapter Page

Chapter Page

LIST OF FIGURES

LIST OF TABLES

Table Page

Table Page

CHAPTER I

INTRODUCTION

Interim assessments are an important component to an assessment system and provide information regarding student achievement to teachers and administrators. If the interim assessment system is closely aligned with expected outcomes on the Smarter Balanced Assessment (SBA), then this data can be a powerful tool for teachers to understand a student's learning trajectory. According to Alonzo (2016), this assessment information, if it correlates well with year-end assessments, can be used to identify students who may need academic interventions. According to Herte (2007), interim assessments may be much more accurate in predicting student success on standardized assessments than teacher-created tests. Regularly measuring student learning with valid and reliable instruments provides information that teachers can use to change instructional and school practices to (Bulkley, Oláh, and Blanc, 2010). Perie, Marion, and Gong (2009) argue that employing interim assessments is a component of a comprehensive assessment system framework.

Interim assessments can provide a more direct connection between curriculum and student learning goals. When students are informed about what they need to focus on to improve their performance, students become partners in the work. Assessment-informed students and teachers have a common roadmap showing both where to focus and prioritize learning.

The school district that served as the site for this study currently uses the Measure of Academic Progress' (MAP) interim assessments. MAP and the Smarter Balanced Assessment are designed to indicate how well students have progressed towards learning standards. Both assessments are designed to evaluate student learning of the Common Core State Standards (CCSS) and thus are centered on the same content domains.

MAP interim assessments and the Smarter Balanced Assessment both demonstrate substantive validity through asking students to engage in a range of cognitive processes that vary in difficulty and are tightly related to the construct of interest (mathematics and language arts).

US public schools are mandated to assess students annually from third to eighth grade and once during high school. Because of the external pressure on summative state-level assessments results, many districts are using tools to provide interim measures in order to understand predictive trajectories of students. In this project, I am investigating teachers' attitudes and practices toward using interim assessment data and whether these attitudes affect student-learning outcomes on state assessments. Specifically, my research examines (1) the ability of interim assessments (i.e., the MAP) to predict scores on the year-end assessment (i.e., Smarter Balanced), and (2) whether the use of data, reported by teachers, can predict scores on the year-end assessment while controlling for the effects of the interim assessment.

Policy Context

No Child Left Behind (NCLB) was the first national educational law that developed accountability, predominantly based on large-scale testing results, from the public-school systems in the US. The law mandated that states publish results from high stakes (summative) assessments and that the performance on these assessments be disaggregated to highlight achievement gaps for traditionally underserved student populations for the first time (Dee & Jacob, 2010). The approach of publishing disaggregated student assessment data, combined with sanctions applied to school districts for not meeting federally established standards for continuing improvement, was ground breaking (Cusick, 2014). In 2015, The Every Student Succeeds Act (ESSA), replaced NCLB and is current law. ESSA includes many of the aspects of accountability that were framed in the previous version of the nation's educational law. For example, ESSA dictates that all states will have annual math and English/language arts assessments (McGuinn, 2016), and that student achievement results be disaggregated demographically and published publicly.

In accordance with ESSA, state of Oregon policy requires school districts to assess their students annually in Grades 3-8 and Grade 11 in mathematics and English/language arts. These summative assessments, administered towards the end of the school year, attempt to measure student learning in relation to grade-level academic standards and instruction. The grade-level academic standards, called the Common Core State Standards (CCSS), were adopted by Oregon in 2010 along with 41 other states and the District of Columbia. In the 2014-2015 school year, Oregon implemented the Smarter Balanced Assessment (SBA) to assess CCSS-based student learning. As required under ESSA, the state publishes disaggregated results annually for each school and district in the Oregon School Report Card, which outlines school and district progress towards meeting state educational standards. Students are required to pass this state accountability assessment, or prove proficiency through a portfolio assessment, in order to earn a standard diploma upon completing all other credit requirements for graduation. Because state summative assessments measure achievement at the end of the school year, they are often perceived as high-stakes experiences for students (and schools and districts required to demonstrate continuing improvement). Interim standards-based assessments, administered across the school year, measure standards-based learning over time, offering students and teachers insight into their strengths and weaknesses prior to the higher-stakes assessment experience in the spring.

CHAPTER II

LITERATURE REVIEW

Administering educational assessments to students is a familiar practice to most teachers and administrators. The high-stakes nature of assessment-driven accountability systems has accelerated the demand for teachers and administrators to know how each student is progressing toward established expectations of student learning goals (Alonzo, 2016). If school administrators and teachers can accurately determine each student's progression toward mastery of the CCSS over the breadth of the school year, then they can be more informed about instructional and curricular considerations to better ensure adequate student achievement by the end of the school year. As a result, many schools and districts attempt to document and predict students' progress toward meeting standards-based learning goals by administering interim assessments.

Interim Assessments

Teachers administer interim assessments multiple times over the school year to establish baseline achievement and then to check progress toward summative learning goals. In this case, levels of proficient achievement are determined from the results of the Smart Balanced Assessment (SBA). Interim assessments should be based on the same academic content standards on which the summative assessment is based and generate data that provide guidance to teachers' instruction (Braun, 2011). Interim assessments are intended to (a) provide greater perspective on each student's learning trajectory and (b) ascertain if a student might require more or different assistance and supports to reach their individual learning goals (Konstantantopoulos, Li, Miller, and Van der Ploeg, 2017). There are many interim assessment products on the market. One such product is the

Measure of Academic Progress (MAP), published by Northwest Evaluation Association (2017).

MAP assessments are computer-adaptive achievement tests in the areas of reading and math. Computer-adaptive tests alter the difficulty of the test based on the accuracy of the previously submitted answer. MAP assessments are aligned with grade-level CCSS standards and are assigned to students according to their current grade level. Wang, Zhao and Addison (2016) investigated the relationship between the MAP assessment and the Partnership for Assessment of Readiness for College and Careers (PARCC), the statewide summative assessment used in Maryland as well as numerous other states and found a strong positive correlation between MAP and PARCC. More specifically, Wang et al. (2016) reported evidence of predictive validity between fall MAP scores and the year-end PARCC test. All correlation coefficients were greater than .80 across all grade levels and in both reading and math. In a similar study, Ball and O'Connor (2016) examined the predictive validity of MAP in relation to the PARCC assessment used as the statewide achievement test in Wisconsin. In their study, Ball and Connor also reported strong positive correlations; the two reading tests had a correlation of .82, and the MAP assessment explained 68% of the variance in performance on the PARCC test (2016). Shortly after the 2016 Ball and O'Connor study, the state of Wisconsin changed their state assessment from the PARCC to SBA. Klingbeil, Van Norman, Nelson and Birr (2018) investigated the relation between the MAP and the Wisconsin SBA assessment and found statistically significant correlations between MAP math scores and SBA (.84) and MAP reading scores and SBA (.78).

Riggan and Oláh (2011) investigated the implementation of interim assessments within teachers' classroom practices and found that interim assessments, in order to be effective, needed to be utilized as part of a larger assessment system that included teacher-developed assessments, interim assessments, and summative state assessments. One assessment system alone cannot meet the information needs of educators regarding student learning. A brief formative assessment may inform a teacher about how a student understood a recent lesson, but teacher-created formative assessments may not be able to predict accurately how that student will perform on the spring state assessment—both assessments provide important information but are used for different purposes. Interim assessments may be able to provide predictive and diagnostic data to the teacher, but it may not capture the nuance of a particular learning objective inside of a lesson. A summative assessment may provide insight as to how a student is understanding a broad range of standards, yet it is not sensitive or generally timely enough to provide instructional direction or aid in diagnosing learning challenges like a formative or interim assessment might. School district assessment systems that include all three types of assessments have expansive sets of student achievement data that teachers can utilize to make decisions about the instruction, interventions, and curricula used in their classrooms, and how they might be implemented and adapted to best serve their students' needs.

Data-Driven Decision Making (DDDM)

With such data available to evaluate teacher instruction and student learning, it would seem plausible that teacher practices where decisions were supported by data derived from interim assessments would become a consistent practice within schools. Data-Driven Decision Making (DDDM) can be defined "as systematically analyzing existing data sources within the school, applying outcomes of analyses to innovate teaching, curricula, and school performance, and, implementing (e.g. genuine improvement actions) and evaluating these innovations" (Schildkamp and Kuiper, 2010, p. 482). According to Datnow, Park, and Kennedey-Lewis (2012), there has been an upswing in the mandatory use of collaboration structures inside schools and school districts to examine and plan instructional actions based on student data. Improving teachers' capacity for effective use of data has also become an organizational priority in schools and districts (Marsh and Farrell, 2015).

Supporting quality DDDM practices requires school organizations to consider a wide range of factors including teacher actions with data, organizational supports (such as principal leadership) and infrastructure to access data (such as dedicated computer programs), improving teacher competence in using data, and supporting and enhancing teacher attitudes towards data and its use (Wayman, Cho, Jimerson and Spikes, 2012).

Wayman, Wilkerson, Cho, Mandinach and Supovitz (2017) have presented a conceptual framework and instrument for understanding and measuring how teachers are using data in their practice: The Teacher Use of Data Survey (TDUS). Figure 1 displays the interaction of the TDUS scales and how those teacher beliefs and actions are related to student outcomes.

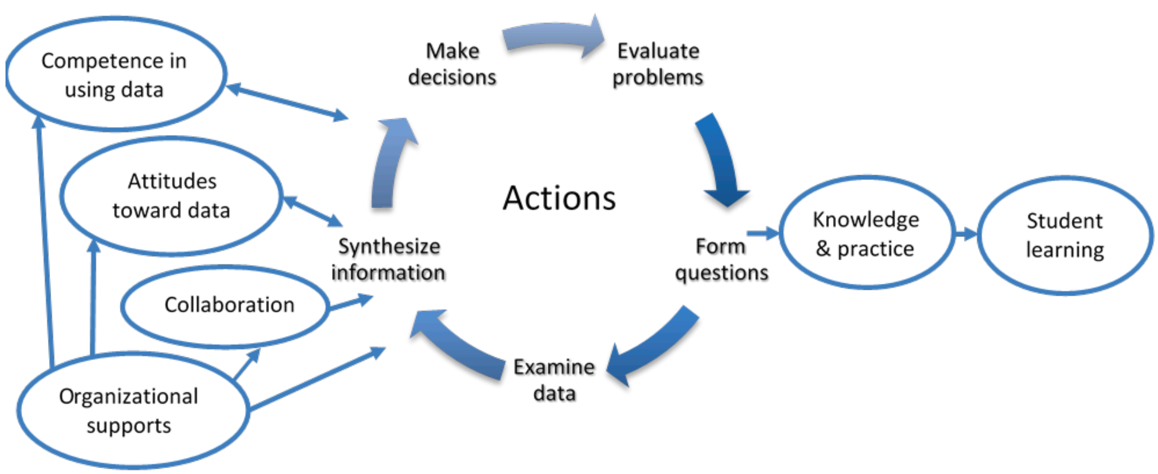

Figure 1. Data use survey conceptual framework (Wayman et al., 2017)

The following sections define the dimensions of the TDUS.

Individual actions with data. Research has indicated that teachers utilize DDDM in a wide range of processes. Beaver and Weinbaum (2015) found teachers identified 15 different activities that were underpinned by the use of data. These activities were organized into three overarching categories that included "planning for future collection and exploration, school-wide improvement efforts, and the individual targeting of students" (Beaver & Weinbaum, 2015, p. 487). Schidkamp and Kuiper (2010) found a wide range of purposes behind data use in six different school district systems, including self-directed studies, accountability systems, collaboration, and targeting and allocating resources.

Collaborative actions with data. These data-driven decision making processes are frequently associated with another common structural and processes found in schools: collaborative teacher teams (Abrams, Varier and Jackson, 2016; Kallemeyn, 2014). Collaborative teacher teams are focused on instructional improvement and use data from various sources to inform the team about student success in their classrooms. Structures that support collaboration are believed to be beneficial, in that when teachers work together, they will be able to assist each other in making sense of the data, engage in joint action planning, and share instructional strategies (Datnow, Park, and Kennedy-Lewis, 2012).

Collaborative Team Trust with Data Use. It is difficult to understate the importance of high-functioning collaborative teacher teams as the driver for improving instruction and learning. As stated earlier, it is in this structure that most school districts rely on for implementing professional development. Being able to work effectively in this structure with other colleagues is key in adopting new practices that lead to improved instruction (DuFour, 2015). Often when establishing these teacher teams, principals prescribe expected behaviors, including practices concerning the use of data (Wayman and Jimerson, 2013). According to DuFour, the crux of how effective a team may become is centered on an interdependent, trusting relationship between the members of the team (2008).

Organizational supports. Part of the draw for school districts to purchase MAP and employ it is the online tools and support provided with purchase. The online portal provides numerous reports and graphs to assist teachers in understanding the performance of their students. Student data can be organized in any number of ways, depending on teacher preference. Providing technology, such as online data portals described above, are examples of how a school or district can promote and enable the use of student data (Wayman, Shaw, and Cho, 2017).

Principal leadership in this area is vitally important. According to Wayman and Jimerson (2013), the research regarding principal influence on teachers using data in their practice is universal. Principal leadership is key in establishing a set of practices that support data use including the provision of time and training regarding these DDDM practices.

Supports for data use. School districts often employ data experts (instructional coaches) to work alongside teachers to train staff to use data more effectively in DDDM processes. Marsh, McCombs, and Martorell (2010) found that instructional coaches spent time working with teachers around ten different activities across a typical two-week work period including reviewing assessment data, administering assessments, training teachers to analyze data and co-planning lessons with teachers.

Competence in using data. Engaging in the DDDM process is complex and requires individuals and collaborative teams of teachers to assume a learning stance when they begin the practice. As Datnow et al. (2012) concluded in their research, even school leaders have certain levels of anxiety related to their skills in interpreting data. It can be the case that teacher teams find interacting with data to be highly complex. For example, with data analytical skills intact, teachers can use interim assessment data to determine certain patterns over time related to student learning. The same data set may also inform the teacher team what next steps ought to be taken immediately in order to improve student-learning trajectories. Bridging the gap between outcomes and teacher practice becomes a paramount skill for both individuals and collaborative teams who are engaged in DDDM for the purpose of improved student outcomes.

Attitudes toward data. Teacher attitudes and beliefs regarding DDDM may be either the greatest asset or hindrance in the process of appropriately using data to improve student learning outcomes. Wayman, Cho, Jimerson and Spikes (2012) found that teachers generally had positive attitude about data and its potential to improve learning outcomes for students. In the same study, Wayman et al. (2012) also uncovered barriers that impacted teachers' beliefs about data including ease of access, day to day difficulties of using data, lack of time to reflect on results as well the possible labor intensiveness of translating data into instructional actions. Abrams, McMillian and Wetzel (2015) discovered that teacher beliefs about data were impacted when the interim assessment did not align well with the state test, or if the time for administration was inordinately long.

Data's Effectiveness for Pedagogy. Remesal (2011) interviewed 50 primary and secondary teachers regarding their attitudes towards using data derived for interim assessments and found contrasting views about the role of assessments (either determining instructional effectiveness or quality of student learning) and valid use of results for lesson planning. Kippers, Wolterinck, Schildkamp, Poortman and Visscher found data-driven decision making was underutilized by a majority of their sample in which the authors hypothesized attitudes, structures, time, and skills related to making data actionable (for improved instruction) were paramount in the observed behaviors (2018). DDDM is a complex process, but a necessary endeavor. Thus, it is essential to establish frameworks that define terms and behaviors in order to develop educators' understanding and perhaps improved educator attitudes about better implementation of DDDM.

Misuse of Data

The process of DDDM is crucial to understand so that teachers and administrators can improve its practice within schools to improve student learning. If assessment data are misused it can be potentially harmful to student outcomes. Frequently, data misuse occurs when school personnel use assessment results to underpin decisions that the assessment and/or resultant data were never intended to inform. For example, state assessment data is a common vehicle for an incongruent decision basis regarding individual intervention decisions because the information is not sensitive to or contiguous with instruction. The argument being that these assessments do not measure learning related to a single or small set of standards (i.e., how instruction is generally planned and implemented), rather it measures proficiency related to a broad set of standards over the school year to drive school and district improvement as a whole. While data is often used ethically in response to intervention (RTI) systems to target students' skills that are lagging behind in grade-level achievement expectations (Alonzo, 2016), poor DDDM may also be used to trap a student in a cycle of remediation, consisting of low rigor and disappointing outcomes, particularly for traditionally underserved student populations (Garner, Thorne, and Horn, 2017; Konstantopoulos et al., 2017). Most concerning yet are the attitudes and perceptions of historically disadvantaged students that result from the misuse, mischaracterization, and publication of achievement data that appear to promote the achievement gap and have little influence on accelerating learning outcomes (Gutierrez, 2008).

Conceptual Framework

Gill, Gorden and Hallgren (2014) developed a theory of action for DDDM in educational settings. This model encompasses three sequential steps that interdependently provide a foundation for teachers to appropriately use data to improve learning outcomes for their students. Figure 2 represents a theory of action for DDDM in educational settings developed by Gill et al., 2014.

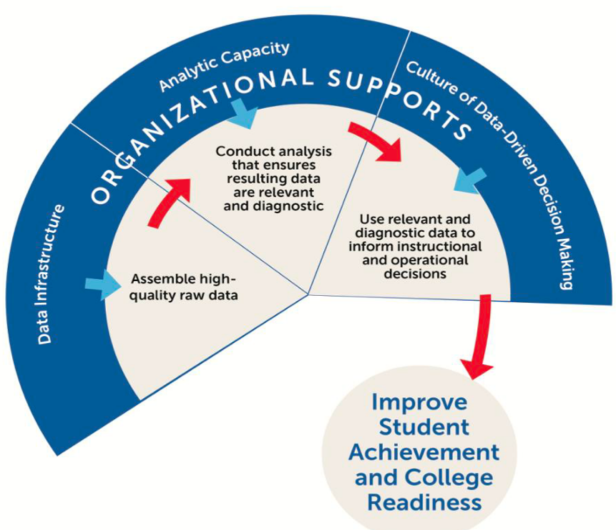

Figure 2: A Theory of Action for DDDM in educational settings developed by Gill, Gorden and Hallgren (2014).

Step 1: Assemble high-quality data. This framework allows for the use of a wide range of data from a variety of sources. Those sources (interim assessments, formative assessments, observations, interviews, etc.) are generated from the data infrastructure of an organization. Collecting high quality data requires a data structure that is robust enough to house and deliver information to all users in a seamless manner. The data infrastructure should also be able to link to other systems in order to combine and disaggregate information from various data bases. Lastly, this system needs be navigable and useable by teaching staff.

Step 2: Conducting analysis. The data from step one is scrutinized to ensure that it is appropriate and relevant to the task at hand. If the data is not relevant to the teacher, such as Oregon state data that may arrive the following school year, then it may not be timely or useful. If analysis reveals that the data is incomplete, then the next steps related to instructional decisions cannot be made. The underpinning structure in this step is the analytic capacity of the teacher(s), which needs to be in place to successfully complete this step. Supporting the analytic capacity of staff becomes paramount as a theme in ongoing professional development.

Step 3: Use the relevant diagnostic data to inform decisions. Once the raw data has become diagnostic in nature due to step two, it is ready to inform and influence instructional and operational decisions. This step is dependent on the systems and culture of the organization to use data upfront in processes.

Summary and Study Context

As the results of published high stakes tests continue to be the vehicle of accountability of public schools in the US, understanding each student's trajectories towards grade level benchmark standards is an essential aspect of educator practice (Alonzo, 2016). Interim assessments that are well correlated to the state assessments are a useful tool for educators to understand this trajectory and make adjustments in instruction, curriculum, or adding additional targeted resources in the effort to help each student learn to their potential. How teachers are making sense of data significantly impacts the data-driven decision making process. Therefore, measuring how teachers make sense as well as ritualized uses of data and the relation of those two factors with actual student outcomes on statewide tests is essential to understand. In this study, I have examined how teachers reported they are accessing and using data derived from their school district's interim assessment program to improve student-learning outcomes with the following research questions;

1. Do the participating school district's third through eighth grade fall MAPS scores in reading and mathematics predict the spring SBA scores in Language Arts and mathematics;
2. Do specific scale scores of the Teacher Use of Data Survey (Actions, Competence in Using Data, Organizational Supports, and Attitudes Toward Data) of the participating school district's third through eighth teachers predict their students' spring SBA scores in language arts and mathematics, controlling for fall MAPS scores, and is this relationship moderated teacher assignment level (i.e., elementary vs. secondary)?

CHAPTER III

METHODS

This chapter begins with an overview of the study and concludes with a detailed description of the measures used, as well as a presentation of the analytic plan.

Research design

This quantitative study aimed to understand if there was a predictive correlational relationship (Creswell, 2018; Babbie, 2013) between teachers' self-reported utilization of interim assessment data (MAP results) and the outcomes of their students on statewide tests (SBA). This study investigated relationships between variables, which is typical of quantitative studies (Creswell, 2018). Each variable will be measured and evaluated against other variables in the study, however the variables are observed and not manipulated or controlled and they occur naturally. The study included two types of quantitative questions outlined by Creswell (2018) that include descriptive as well as predictive questions.

Study Setting

This study took place within a suburban school district in northern Oregon. The district covers 42 square miles, situated south of the metropolitan area of Portland. The district is located in Clackamas county and includes two towns. The district serves approximately 10,000 students with approximately 4,200 students at the nine primary schools, 2400 students at the four middle schools, and 3,500 students enrolled in the three high schools. The student population is comprised of 7% Ever English Learners, 12% students with disabilities, and 17% of the students qualify for the federal free and reduced-price meal program.

This study aimed to understand the use of data derived from the district's interim assessment program. Interim assessments were administered to all students in grades three through eight. Table 1 displays the demographics of the schools that administered the district interim assessments and are included in this research study.

Table 1. Student demographics by school in research study.

Type of School	Number of students	Free/Reduced Rate %	Sped %	ELL%	Minority population %
Primary #1	522	32	15	10	26
Primary #2	378	11	10	*	13
Primary #3	597	38	15	16	35
Primary #4	291	10	15	*	15
Primary #5	573	27	13	19	40
Primary #6	451	6	13	*	17
Primary #7	329	16	12	<5	20
Primary #8	609	6	10	*	25
Primary #9	562	16	12	<5	20
Middle #1	637	11	11	<5	21
Middle #2	558	29	11	15	35
Middle #3	354	32	17	13	31
Middle #4	774	10	10	<5	20
Total/Average	6,635	17	12	7	25

Note: * Not a large enough number of students to report as a sub-group. Minority population includes American Indian/Alaska Native, Asian, Black/African American, Hispanic/Latino, Multiracial, and Native Hawaiian/Pacific Islander.

Unit of Analysis and Time Frame

This study sought to objectively measure and analyze data derived from individual teachers that administered and had access to the data outcomes of the interim assessment system within the school district. The nature of data in this study were considered to be nested: units organized inside of superordinate units. In the current study, both student variables and teacher variables were units included in the analysis. Students were assigned and organized within the school scheduling systems (students in a particular class or assigned to a particular teacher). Both groups (referred to as levels) were associated with variables that were subsequently evaluated (Raudenbush and Bryk, 2002). Therefore, this multi-level analysis investigated student performance outcomes (level-1) as they were influenced by teacher survey responses (level-2). While individual teachers completed responses on quantitative surveys, it is my aim to better understand "how different groups of individuals behave as individuals" with data-driven decision making processes (Babbie, 2013, p. 97).

Confidentiality. This longitudinal study used a combination of teacher survey data and student assessment data collected by the school district. The survey was administered in the winter of 2020 and extant student data was collected from the fall and spring (MAP and SBA scores, respectively) of the 2018-2019 school year. All information that could identify an individual was removed by the district prior to my access and analysis of the data.

Study Sample

I used a purposive sample design in my study. As Babbie (2013) contends, a purposive sample allows for the selection of the sample based on a particular purpose or set of knowledge, in this case, the use of data derived from the school district's interim assessment system (MAP). The sample included as many teaching staff members as possible who teach grades three through eight in the school district. Approval was sought from the University of Oregon's Institutional Review Board (Appendices B) as well as the school district (Appendices C) in order to ensure that the procedures and methods provided the participants with appropriate protections (Creswell, 2018).

The number of total licensed staff was 342 people from across the 13 primary and middle schools. Through email, all teaching staff at the 13 schools were invited to participate in the Teacher Data Use Survey (TDUS). Once the survey closed, the sample responses were culled to only include third through eighth-grade classroom teachers at the primary level (See Appendices D for email invitation script). The number of participates at the primary level were 51 teachers representing a 54% survey completion rate.

At the middle school level, the sampled was culled to include language arts and mathematics teachers. These were the teaching assignments that administered and have access to the results of both the SBA and MAP tests. The number middle school mathematics and language arts teachers participated were 21, representing a survey completion rate of 46%. Table 2 summarizes the demographic characteristics of the sample.

Table 2. Demographics of teachers included in the Sample.

Characteristic:	N	%
Gender		
Female	56	77
Male	16	23
Ethnicity		
Asian	3	4
Caucasian	65	90
Hispanic	4	6
Years of Experience		
1 - 5	14	19
6 - 10	17	24
11 - 15	17	24
16 +	24	33

Measures

The following section describes the measures that I used in my study. Information about the measures were derived from research articles as well as technical manuals. The three instruments used were the Measure of Academic Progress (MAP), the Smarter Balanced Assessment (SBA), and the Teacher Data Use Survey (TDUS).

Measure of Academic Progress. The Measure of Academic Progress (MAP) assessments are computer adaptive achievement tests in the areas of reading and math. These assessments are administered to nearly 8,000,000 students annually (NWEA, 2015). Computer-adaptive tests alter the difficulty of the test based on the accuracy of the previously submitted answer. MAP assessments are aligned with grade-level content standards and are assigned to students according to their current grade level. Items in each administration of MAP are drawn from a pool of 34,000 items, which ensured a zero item repetition on assessments taken within 14 months (NWEA, 2015). In the current study, the school district MAP assessments are administered three times per year: fall, winter, and spring. Outcome scores from MAP are in the form of Rasch Units (RIT), and range from 100 to 300 (NWEA, 2015). MAP RIT scores represent scale difficulties, and as such, MAP RIT scores within the same content area are able to be compared across grade levels.

The MAP assessment has been designed to provide predictive performance estimates on the Smarter Balanced Assessment (SBA), the year-end state assessment used in Oregon. Klingbeil et al. (2018) investigated the relation between the MAP and the SBAC assessment and found statistically significant predictive correlations between MAP math scores and SBAC (.84) and MAP reading scores and SBAC (.78).

Smarter Balanced Assessment. The SBA test is administered to all students in grades three through eight and once during high school. This year-end state assessment is the vehicle to measure student learning across a broad range of standards in English language arts and mathematics. Scores from this assessment are collected by the state and serve as the foundational data source for student, school and district accountability outcomes. The test contains two subtests for each content area. Subtest 1 is a computer-adapted assessment (similar to MAP) and subtest 2 is a written performance task within the content area. SBA scores are scaled vertically on the same scale across grade levels. This allows for comparisons between students across grades. Scores from SBA constituted the student outcome variables for the current study.

The Teacher Data Use Survey. In order to understand how teachers use data to support instruction, their attitudes towards data and the supports that help teachers use data, I utilized a modified version of the Teacher Data Use Survey (TDUS) developed by Wayman et al., (2016).

The survey is intended to measure attitudes and behaviors that teachers have when working with data (see Appendices E for complete survey). These include measuring particular actions such as teacher competence in using data or teacher's beliefs regarding required organizational supports for using data. The TDUS also addresses collaborative behaviors with team members when reviewing data and personal attitudes or orientations towards interim assessment (MAP) data. Survey questions are all set on a four-point Likert scale (1 = strongly disagree, 4 = strongly agree). The instrument was developed through the Regional Education Laboratory (REL) and was funded by the U. S. Department of Education's Institute of Educational Sciences (IES). The survey was subjected to the internal review process of REL and the peer review process of IES. Descriptive statistics were generated during the review process. Standard errors were between 0.10 and 0.20. Reliability was established with alpha reliabilities ranging between 0.84 and 0.97. Inter-item correlations were greater then .70 for most scales.

The TDUS does not provide an overall score for each participant. Instead, it provides outcomes according to its scales. In the current study, I focused on the four scales of the TDUS, the composite score of the survey and selected subscales that I think are most aligned with research question 2 of this project.

Actions with Data Scale. My study investigated teacher behavior with data as the central focus of the study. Particularly, I focused on the use of interim data (MAP Assessment). Sets of questions in this scale are designed to reveal the use of interim data. A second component of the Actions with Data scale is Teacher Collaboration. As presented earlier, expectations regarding teacher collaboration are becoming universal in most school districts (Datnow et al., 2012; DuFour, 2015). Thus, the teacher collaboration component of the Actions with Data Scale were included in my analysis.

Organizational Supports Scale. The Organizational Supports Scale addresses how teachers view and interact with the district's data infrastructure. Accessing high quality data is a key component in the conceptual framework used in this study (Gill et al., 2014). Organizational supports in such systems as the MAP portal (Computer Data System Subscale), school organizational and priority factors (Principal Leadership Subscale), and expert support/professional development (Support for Data Use Subscale) are all essential components that support systemic capacity for data use (Wayman et al., 2016; Marsh & Farrell, 2014).

Attitudes toward Data Scale. As outlined above in the literature review (Wayman et al., 2012; Abrams et al., 2015; Kippers et al. 2018) all contend that teacher attitudes toward the use of data in their planning and instruction are of paramount importance. I investigated the relation of these two subscales (Attitudes Towards Data and Data's Effectives for Pedagogy) with student outcomes (SBA) as part of the Attitudes toward Data scale in my analysis.

Competence in using Data Scale. The competence scale allowed me to investigate teacher perceived efficacy in relation to analyzing and effectively using data in their practice (Gill et al., 2014). This scale interested me as I hoped to provide recommendations to the school district related to professional development as an outcome of this study.

The second subscale of this component provided insight into challenges that teachers face when they share their students' data with others (listed as Collaboration Team Trust). The Collaboration Team Trust subscale gets to the heart of how effective teacher teams are when working with student data (Datnow et al., 2012). This is a crucial aspect because all schools in the sample school district are organized into teacher teams and working with student data (MAP assessment results as well as other data) is a common expectation across the sample school district.

Validation of the TDUS. The TDUS was administered according to implementation recommendations from Wayman et al. (2016). An Exploratory Factor Analysis (EFA) was preformed to verify and substantiate the underlying factor structure of the survey with my sample. Based on the resultant scree plot, a ten-factor solution with Eigenvalues above 1.0 met my criterion for extraction. I executed a series of factor analysis in which I deleted items with factor loadings below .40 or which loaded on more than one factor. Two subscales in the Organizational Supports scale mapped onto the same factor. I decided to combine these two subscales as it matched with the scree plot for a total of ten factors. Overall, six individual items were deleted to create a clean pattern matrix. Each component accounted for greater than two percent variance and the variance of the final structure was estimated at 76.45%. Table 4 displays the final pattern matrix solution from the EFA.

Table 4. Factor Loadings for Exploratory Factor Analysis of the Teacher Use of Data Survey results.

Scale	Component									
	1	2	3	4	5	6	7	8	9	10
6.1 Act/ID									.433	
6.2 Act/ID									.806	
6.4 Act/ID									.764	
6.5 Act/ID							.986			
6.6 Act/ID							.555			
6.7 Act/ID							.860			
6.8 Act/ID							.800			
7.1 Act/PD				.866						
7.2 Act/PD				.917						
7.3 Act/PD				.818						
7.4 Act/PD				.924						
7.6 Act/PD				.824						
7.8 Act/PD				.513						
8.1 Act/CD					.883					
8.2 Act/CD					.870					
8.3 Act/CD					.863					
8.4 Act/CD					.727					
8.5 Act/CD					.637					
8.6 Act/CD					.675					
8.7 Act/CD					.467					
17.1 Act/PLC	.745									
17.2 Act/PLC	.720									
17.3 Act/PLC	.893									
17.4 Act/PLC	.711									
17.5 Act/PLC		.654								
17.6 Act/PLC		.729								
17.7 Act/PLC		.794								
17.8 Act/PLC		.867								
17.9 Act/PLC		.857								
17.10 Act/PLC		.864								
11.1 Att/EP			.615							
11.2 Att/EP			.799							
11.3 Att/EP			.774							
11.4 Att/EP			.836							
11.5 Att/EP			.865							
11.6 Att/EP			.858							
11.7 Att/EP			.706							
11.8 Att/EP			.839							

11.9 Att/EP			.813							
10.1 Org/SDU		.715								
10.2 Org/SDU		.554								
10.3 Org/SDU		.749								
10.4 Org/SDU		.622								

Table 4 (continued).

Scale	Component									
	1	2	3	4	5	6	7	8	9	10
10.6 Org/SDU		.795								
12.2 Org/PL		.835								
12.3 Org/PL		.843								
12.4 Org/PL		.846								
12.5 Org/PL		.799								
12.6 Org/PL		.725								
13.1 Org/CDS										.735
13.2 Org/CDS										.701
13.3 Org/CDS										.531
13.4 Org/CDS										.811
14.1 Comp/DC								.621		
14.2 Comp/DC								.741		
14.3 Comp/DC								.767		
14.4 Comp/DC								.797		
16.1 Comp/PLC						.849				
16.1 Comp/PLC						.874				
16.1 Comp/PLC						.853				
16.1 Comp/PLC						.885				

Note. Factor loadings that were < .40 were removed from the final Pattern Matrix.

In the current study, a small pilot study was run with two primary and two middle school teachers. Based on feedback, adjustments were made to question stems to reflect local terminology. Items that included jargon, such as interim assessments, were written to reflect local common terms, such as the MAP assessment. All questions regarding state testing and data were eliminated to reduce confusion and focus the instrument on interim assessment data. Average time to complete the survey was approximately eleven minutes.

Reliability of the TDUS for the current study sample was established using Cronboach's Alpha coefficients. The reliability estimates were performed after the final pattern matrix solution was determined and the six items were deleted from the scales. The reliability coefficients ranged between .81 and .95 for the four scales and the composite scale. The reliability coefficients for the subscales ranged between .80 and .95. Table 3 displays the scales and coefficients for the study sample.

Table 3. Cronbach's Alpha Coefficients of the TDUS scales.

Scale	Items Comprising the Scale	α
TDUS Composite	All Scales	.95
Actions		.91
Collaborative team actions	17a, b, c, d, e, f	.95
Actions with interim test Data	6a, b, d, e, f, g, h	.81
Actions with publisher Data	7a, b, c, d, f, h	.95
Actions with classroom data	8a, b, c, d, e, f, g,	.88
Organizational Supports		.93
Computer Data Systems and Principal Leadership	13a, b, c, d, 12a, b, c, d, e, f	.80
Support for Data Use	10a, b, c, d, e, f	.87
Attitudes toward Data		.94
Attitudes toward Data	11f, g, h, i	.93
Data's effectiveness for Pedagogy	11a, b, c, d, e	.88
Competence in using Data		.81
Data competence	14a, b, c, d	.90
Collaborative Team Trust	16a, b, c, d	.89

Data Analytic Plan

Variables. My study included both independent and dependent variables. The dependent variable was the student test outcomes on the SBA test. The independent variables were the fall student MAP data as well as averaged scores on the scales and minor scales of the Teacher Use of Data Survey.

It was suggested by a committee member to investigate and test for multicollinearity of the TDUS scale score averages. These scores, as noted above, were independent variables and should, therefore, be independent from each other. Multicollinearity can significantly impact various statistical coefficient estimations, therefore reducing precision and power of the study (Spiegelhalter, 2019). In order to understand the degree of multicollinearity of the scales of the TDUS, I performed a bivariate Pearson correlation test. Table 5 displays the correlations.

Table 5. Pearson correlations for the TDUS scales.

Scale	Action	Organization	Attitudes	Competence
Action	1	.241	.455**	.360*
Organization	.241	1	.624**	.458**
Attitudes	.455**	.624**	1	.446**
Competence	.360*	.458**	.446**	1

*$p < 0.05$. **$p < 0.01$.

From table five, the largest correlation within the scales was .624. While this is correlation was statistically significant, correlations between .5 and .7 are generally referred to as moderately correlated (Spiegelhalter, 2019). In order to address any effects from the influence of multicollinearity, all independent variables in this study were centered around their means to reduce structural multicollinearity. Centered variables were also used for interaction terms in certain applications of my analytic model, explained below.

Analytic Models. Research question one (RQ1) aims to investigate if the MAP reading and math assessments predict the spring SBA English language arts and mathematics test scores. The analytical model for RQ1 is as follows:

(level 1 - student) $SBA_{ij} = \pi_{0j} + \pi_{1j}$ (student MAP Score)$_{ij} + e_{ij}$
(level 2 - teacher) $\pi_{0j} = \beta_{00j} + r_{0j}$
(level 2 - teacher) $\pi_{1j} = \beta_{10j} + r_{1j}$

Level-1 represents the degree that MAP scores predict SBA scores. Level-2 in this model is left "open" with no teacher predictors from the TDUS scales in order to isolate the effect of the fall MAP score on the spring SBA score.

The aim of research question two (RQ2) was to understand the relationship between teacher use of data and student achievement. Schools are organized into classrooms. Each class is taught by a teacher that had strengths and weaknesses in their instructional and interpersonal skill sets. Since a set of students experience a teacher together in a classroom, it is impossible to ignore the assumption of independence due to the nested nature of the study setting and the nature of the resultant outcomes. The multilevel nature of the study data (i.e., student within teacher classroom) requires that we used nested random coefficients analysis with HLM 8.0 software, which allocated variance either "within" or "between" groups, accounting for dependencies introduced by nesting (Raudenbush & Bryk, 2002).

The basic analytical model for RQ2 was:

(level 1 - student) $\text{Outcome}_{ij} = \pi_{0j} + \pi_{1j}\,(\text{student MAP Score})_{ij} + e_{ij}$
(level 2 - teacher) $\pi_{0j} = \beta_{00} + \beta_{01}(\text{School Level})_j + \beta_{02}(\text{TDUS Scale})_j + \beta_{03}(\text{Interaction Term})_j + r_{0j}$
(level 2 - teacher) $\pi_{1j} = \beta_{10j} + \beta_{11}(\text{School Level})_j + \beta_{12}(\text{TDUS Scale})_j + \beta_{13}(\text{Interaction Term})_j + r_{1j}$

Level-1 represents the individual student, and their outcome on the SBA and controls for the predictive MAP score. With Level-2, I began with investigating the interaction between school level (dichotomously coded for either primary or middle level job assignment) and the scales of the TDUS. I theorized that school level would moderate the influence of the TDUS scale being evaluated. I developed interaction terms for school level and each scale average for the TDUS and tested them within the model to estimate the moderator effect. If the interaction was not significant, meaning the effect of the TDUS scale was not different according to school level, then the interaction terms were removed from the model and the main effects of school level and the TDUS were investigated by themselves. As discussed above, these predictors were centered before the interaction term was created to reduce multicollinearity. The effect of interest was $\beta_{02,}$ which indicated the degree to which the TDUS scale predicting student SBA test scores in the spring, controlling for MAP scores in the fall (assuming the interaction effect was not significant). In an exploratory manner, I also examined $\beta_{12,}$ which indicated the degree to which the TDUS scale at Level-2 moderated the relationship between MAP and SBA scores at Level-1 (i.e., a cross-level interaction).

Hierarchical linear model justification. In order to justify that the Hierarchical Linear Model was required for evaluating these data sets, I began by calculating the Intraclass Correlation (ICC) of the unconditional model, that is, a model with the dependent variable (SBA) only and not including any predictors at level one or two. The HLM equation for the unconditional model is:

(level 1 - student) $SBA_{ij} = \pi_{0i} + e_{ti}$
(level 2 - teacher) $\pi_{0i} = \beta_{00} + r_{0i}$

The equation to calculate the ICC is:

$$P = \frac{\sigma_0^2}{\sigma_0^2 + \sigma^2} = \frac{9383.44}{9383.44 + 18144.73} = .34$$

The variance of the unconditional model at level-2 (teacher) was $p = .34$. This estimate explains that about a third of the variance of math SBA scores can be accounted for by which class the student is enrolled in. The ICC estimation of variance at level 2 for the language arts SBA scores was .21. With a substantial amount of variation unaccounted for in the unconditional model for both language arts and math SBA scores, an HLM modeling approach could help further explain the remaining unknown variance by introducing level-1 predictors (MAPS Scores) and level-2 predictors (TDUS scale scores).

Relative Fit of the model. Model fit was investigated and subsequently found to be improved with the inclusion of random effects at level-2 for the slope of the MAP score predicting SBA (r_1 in the model above). The deviance statistic along with the number of estimated parameters that are included in the model (which serve as degrees of freedom) were compared. The reduction in deviance between the two models was statistically significant, given the degrees of freedom used via a chi-square table of critic values. The deviance value was significantly improved by reducing the statistic by a value of 147.31 in the language arts results model and by 1278.25 in the mathematics model. Thus, the slope of MAP scores predicting SBA scores were allowed to vary at Level-2 in all subsequent models.

CHAPTER IV

RESULTS

Research Question 1

Research question 1 investigates if fall MAP scores in reading and mathematics predict the spring SBA scores in language arts in the sample school district. I found significant predictive relationships between MAP and SBA in the content area of language arts. Table 6 summarizes the estimation of fixed effects. Table 7 presents the estimation of the variance components.

Table 6. Estimation of fixed effects for fall reading MAP test scores predicting spring SBA language arts test scores.

Fixed Effect	Coefficient	Standard Error	*t*-ratio	Approx. *d.f.*	*p*-value
Intercept 1, π_0					
Intercept 2, β_{00}	2555.25	4.31	592.6	43	<0.001
MAP Read slope, π_1					
Intercept 2, β_{10}	5.28	0.33	16.11	43	<0.001

The estimation of the variance components of the fixed effects for fall MAP reading test scores predicting the spring SBA language arts tests scores are displayed in table 7.

Table 7. Estimation of variance components for fall reading MAP test scores predicting SBA language arts test scores.

Random Effect	Standard Deviation	Variance Component	*d.f.*	X^2	*p*-value
Intercept 1, r_0	20.60	424.1	43	100.45	<0.001
MAP Read slope, r_1	1.86	3.47	43	100.45	<0.001
Level -1, *e*	82.98	6884.02			

The amount of variance remaining at level-1 once MAP was introduced as a predictor was 6884.02. The remaining variance in the unconditional model was 14454.72. It appears that the proportion of variance explained by MAP was 47%.

I found significant predictive relationships between fall math MAP scores and spring math SBA scores. Table 8 summarizes the estimation of fixed effects.

Table 8. Estimation of fixed effects for fall math MAP test scores predicting spring SBA math test scores.

Fixed Effect	Coefficient	Standard Error	*t*-ratio	Approx. *d.f.*	*p*-value
Intercept 1, π_0					
Intercept 2, β_{00}	2580.50	21.73	118.77	42	<0.001
MAP Math slope, π_1					
Intercept 2, β_{10}	7.01	0.88	7.94	43	<0.001

The estimation of the variance components of the fixed effects for fall MAP math tests scores predicting the spring SBA math tests scores are displayed in table 9.

Table 9. Estimation of variance components for fall math MAP test scores predicting SBA math test scores.

Random Effect	Standard Deviation	Variance Component	*d.f.*	X^2	*p*-value
Intercept 1, r_0	138.01	19044.58	42	942.03	<0.001
MAP Math slope, r_1	5.59	31.19	43	972.38	<0.001
Level -1, *e*	82.39	6789.02			

The amount of variance remaining at level-1 once MAP was introduced as a predictor was 6789.02. The remaining variance in the unconditional model was 18144.73. It appears that the proportion of variance explained by MAP is 37%.

Research Question 2

Research question two introduces the quantitative survey results from the TDUS. The results of the survey, organized by average scores on the 4 scales and overall survey average, served as predictors in the level-2 equations of the statistical models for research question 2. A total of 72 primary and middle school teachers completed the TDUS survey. The overall mean of the instrument for all the scales was 2.68, and the standard deviation was .4. Table 10 displays the descriptive statistics of the TDUS survey.

Table 10. Descriptive statistics for the Teacher's Use of Data Survey based on validated EFA analysis.

Scale	N	M	SD
TDUS	72	2.68	.40
Action Scale	72	1.95	.37
Collaborative Team Actions	72	2.26	.64
Action with Interim Test Data	72	1.21	.37
Action with Publisher Test Data	72	2.15	.63
Action with Classroom Test Data	72	2.24	.63
Organizational Supports	72	2.60	.46
Supports for Data Use	72	2.51	.51
Computer data Systems	72	2.65	.54
Attitudes Toward Data	72	3.03	.50
Attitudes Toward Data	72	3.02	.54
Data's Effectiveness for Pedagogy	72	3.05	.50
Competence in Using Data	72	3.16	.42
Data Competence	72	3.56	.57
Collaborative Team Trust	72	3.45	.50

Note. A response of 1 = strongly disagree, and a response of 4 = strongly agree.

I began with investigating the interaction between school level (dichotomously coded for either primary or middle level job assignment) and the scales of the TDUS. I theorized that the school level would moderate the influence of the TDUS scale being evaluated. I developed interaction terms for school level and each scale average for the TDUS and tested them within the model to estimate the moderator effect. If the interaction was not significant, meaning the effect of the TDUS scale was not different according to school level, then the interaction terms were removed from the model and the main effects of school level and the TDUS were investigated by themselves.

School level and TDUS Composite scale for effects on Language Arts SBA. Model 1 tested the interaction of school level and TDUS composite scale score on the SBA reading test outcomes. The interaction term of -7.40 ($d.f. = 40$, $p = 0.200$) was found to be not significant. The interaction term was removed for Model 2 and the main effects were examined. Both the TDUS Composite term of 1.60 ($d.f. = 41$, $p = 0.85$)

and School Level term of -3.21 (*d.f.* = 41, p = 0.56) were found not significant (see table 11).

Table 11. TDUS Composite Scale Results Predicting State Reading Test Scores (N = 1522 students and N = 44 teachers).

	Model 1 (Interaction)		Model 2 (Main Effects)	
Predictors	*Coefficient*	*SE*	*Coefficient*	*SE*
Student level				
MAP	5.31***	.34	5.31***	.34
Teacher level				
Intercept	2554.04***	4.51	2554.01***	4.55
TDUS - Composite	1.60	.5.84	1.15	5.87
School Level	-8.54	6.88	-3.22	5.53
Interaction	-7.40	5.68		
Error Variance				
Level 1 (e)	6881.58		6879.93	
Level 2 (r_0)	458.75		472.42	
Level 2 (r_1)	3.65		3.61	

*p < .05. **p < .01. ***p < .001.

School level and TDUS Actions with Data scale for effects on Language Arts SBA. Model 1 tested the interaction of school level and TDUS Actions with Data scale scores on the SBA reading test outcomes. The interaction term of -4.92 (*d.f.* = 40, p = .303) was found not significant. The interaction term was removed for Model 2 and the main effects were examined. Both the TDUS Actions with Data scale score of -0.38 (*d.f.* = 41, p = 0.94) and School Level term of -4.93 (*d.f.* = 41, p = 0.41) were found not significant (see table 12).

Table 12. TDUS Actions with Data Scale Predicting State Reading Test Scores (N = 1522 students and N = 44 teachers).

	Model 1 (Interaction)		Model 2 (Main Effects)	
Predictors	*Coefficient*	*SE*	*Coefficient*	*SE*
Student level				
MAP	5.30***	.33	5.31***	.33
Teacher level				
Intercept	2553.90***	4.54	2553.96***	4.55
TDUS - Action	1.16	5.39	-.38	5.21
School Level	-6.30	5.45	-4.11	4.97
Interaction	-4.93	4.72		
Error Variance				
Level 1 (e)	6882.47		6881.10	
Level 2 (r_0)	470.27		472.35	
Level 2 (r_1)	3.58		3.61	

*p < .05. **p < .01. ***p < .001.

School level and TDUS Attitudes Towards Data scale for effects on Language Arts SBA. Model 1 tested the interaction of school level and TDUS Attitudes towards Data scale scores on the SBA reading test outcomes. The interaction term of -3.64 ($d.f.$ = 40, p = .438) was found not significant. The interaction term was removed for Model 2 and the main effects were examined. Both the TDUS Attitudes towards Data scale score of 3.23 ($d.f.$ = 41, p = 0.54) and School Level term of -2.04 ($d.f.$ = 41, p = 0.69) were found not significant (see table 13).

Table 13. TDUS Attitude towards Data Scale Predicting State Reading Test Scores (N = 1522 students and N = 44 teachers).

	Model 1 (Interaction)		Model 2 (Main Effects)	
Predictors	*Coefficient*	*SE*	*Coefficient*	*SE*
Student level				
MAP	5.31***	.39	5.31***	.37
Teacher level				
Intercept	2554.02***	4.59	2554.07***	4.54
TDUS-Attitude	5.46	5.98	3.233	5.23
School Level	-3.19	5.44	-2.04	5.14
Interaction	-3.64	4.64		
Error Variance				
Level 1 (e)	6878.02		6878.91	
Level 2 (r_0)	489.7		471.38	
Level 2 (r_1)	3.70		3.66	

*p < .05. **p < .01. ***p < .001.

School level and TDUS Competence with Data scale for effects on Language Arts SBA. Model 1 tested the interaction of school level and TDUS Attitudes towards Data scale scores on the SBA reading test outcomes. The interaction term of -2.89 (*d.f.* = 40, p = .59) was found not significant. The interaction term was removed for Model 2 and the main effects were examined. Both the TDUS Competence with Data scale score of -2.14 (*d.f.* = 41, p = 0.67) and School Level term of -4.79 (*d.f.* = 41, p = 0.29) were found not significant (see table 14).

Table 14. TDUS Competence with Data Scale Predicting State Reading Test Scores (N = 1522 students and N = 44 teachers)

	Model 1 (Interaction)		Model 2 (Main Effects)	
Predictors	*Coefficient*	*SE*	*Coefficient*	*SE*
Student level				
MAP	5.30***	.34	5.30***	0.39
Teacher level				
Intercept	2554.01***	4.64	2553.97***	4.54
TDUS- Competence	-2.30	5.04	-2.14	.67
School Level	-6.30	5.41	-4.79	.29
Interaction	-2.89	5.36		
Error Variance				
Level 1 (e)	6877.05		6879.75	
Level 2 (r_0)	506.47		468.27	
Level 2 (r_1)	3.84		3.71	

*p < .05. **p < .01. ***p < .001.

School level and TDUS Organizational Supports scale for effects on Language Arts SBA. Model 1 tested the interaction of school level and TDUS Organizational Supports scale scores on the SBA reading test outcomes. The interaction term of -7.74 (*d.f.* = 40, p = .11) was found not significant. The interaction term was removed for Model 2 and the main effects were examined. Both the TDUS Organizational Supports scale score of 1.66 (*d.f.* = 41, p = 0.76) and School Level term of -3.07 (*d.f.* = 41, p = 0.54) were found not significant (see table 15).

Table 15. TDUS Organizational Supports Scale Predicting State Reading Test Scores (N = 1552 students and N = 44 teachers).

	Model 1 (Interaction)		Model 2 (Main Effects)	
Predictors	*Coefficient*	*SE*	*Coefficient*	*SE*
Student level				
MAP	5.31***	.4.42	5.30***	.33
Teacher level				
Intercept	2554.35***	4.42	2554.07***	4.55
TDUS – Org. Sup.	4.53	5.62	1.66	5.48
School Level	-6.47	5.21	-3.07	-0.62
Interaction	-7.74	4.72		
Error Variance				
Level 1 (e)	6886.01		6879.84	
Level 2 (r_0)	421.60		471.88	
Level 2 (r_1)	3.55		3.63	

*$p < .05$. **$p < .01$. ***$p < .001$.

School level and TDUS Composite scale for effects on Math SBA. Model 1 tested the interaction of school level and TDUS Composite scale scores on the SBA math test outcomes. The interaction term of 9.73 (*d.f.* = 39, $p = .69$) was found not significant. The interaction term was removed for Model 2 and the main effects were examined. The TDUS Composite scale score of -4.35 (*d.f.* = 40, $p = 0.84$) was found not significant. The School Level term of 61.46 (*d.f.* = 40, $p = 0.005$) was found to be significant (see table 16).

Table 16. TDUS Composite Scale Results Predicting State Math Test Scores (N = 1534 students and N = 43 teachers)

	Model 1 (Interaction)		Model 2 (Main Effects)	
Predictors	*Coefficient*	*SE*	*Coefficient*	*SE*
Student level				
MAP	6.40***	.68	7.03***	.83
Teacher level				
Intercept	2565.08***	0.68	2580.46***	19.99
TDUS - Composite	10.93	17.37	-4.35	-0.21
School Level	40.85	16.93	61.46**	2.97
Interaction	9.73	23.91		
Error Variance				
Level 1 (e)	5944.63		6788.86	
Level 2 (r_0)	7814.23		15973.46	
Level 2 (r_1)	17.67		27.96	

*$p < .05$. **$p < .01$. ***$p < .001$.

School level and TDUS Actions with Data scale for effects on Math SBA.

Model 1 tested the interaction of school level and TDUS Actions with Data scale scores on the SBA math test outcomes. The interaction term of 25.96 (*d.f.* = 39, p =0.09) was found not significant. The interaction term was removed for Model 2 and the main effects were examined. The TDUS Actions with Data scale score of -9.72 (*d.f.* = 40, p = 0.65) was found not significant. The School Level term of 59.86 (*d.f.* = 40, p = 0.007) was found to be significant (see table 17).

Table 17. TDUS Actions with Data Scale Predicting State Math Test Scores (N = 1534 students and N = 43 teachers).

	Model 1 (Interaction)		Model 2 (Main Effects)	
Predictors	*Coefficient*	*SE*	*Coefficient*	*SE*
Student level				
MAP	6.39***	.66	7.04***	.84
Teacher level				
Intercept	2565.19***	13.75	2580.62***	19.96
TDUS- Action	14.98	14.56	-9.72	21.02
School Level	47.64	15.14	59.86**	20.87
Interaction	25.96	15.08		
Error Variance				
Level 1 (e)	5948.72		6789.41	
Level 2 (r_0)	6992.06		15858.59	
Level 2 (r_1)	16.73		28.07	

$*p < .05$. $**p < .01$. $***p < .001$.

School level and TDUS Attitudes towards Data scale for effects on Math SBA. Model 1 tested the interaction of school level and TDUS Attitudes towards Data scale scores on the SBA math test outcomes. The interaction term of -23.82 (*d.f.* = 39, p = 0.19) was found not significant. The interaction term was removed for Model 2 and the main effects were examined. The TDUS Attitudes towards Data scale score of -7.46 (*d.f.* = 40, p = 0.71) was found not significant. The School Level term of 61.61 (*d.f.* = 40, p = 0.004) was found to be significant (see table 18).

Table 18. TDUS Attitudes towards Data Scale Predicting State Math Test Scores (N = 1534 students and N = 43 teachers).

	Model 1 (Interaction)		Model 2 (Main Effects)	
Predictors	*Coefficient*	*SE*	*Coefficient*	*SE*
Student level				
MAP	6.41***	.67	7.04***	.84
Teacher level				
Intercept	2565.41***	14.15	2580.33***	19.97
TDUS - Attitudes	-12.33	14.93	-7.46	20.24
School Level	31.49	14.73	61.61**	20.32
Interaction	-23.82	18.14		
Error Variance				
Level 1 (e)	5945.59		6789.60	
Level 2 (r_0)	76.68.74		15891.19	
Level 2 (r_1)	17.41		27.88	

$^*p < .05$. $^{**}p < .01$. $^{***}p < .001$.

School level and TDUS Competence with Data scale for effects on Math SBA. Model 1 tested the interaction of school level and TDUS Competence with Data scale scores on the SBA math test outcomes. The interaction term of 22.09 ($d.f.$ = 39, p = 0.31) was found not significant. The interaction term was removed for Model 2 and the main effects were examined. The TDUS Competence with Data scale score of 40.74 ($d.f.$ = 40, p = 0.30) was found not significant. The School Level term of 86.21 ($d.f.$ = 40, p = 0.014) was found to be significant (see table 19).

Table 19. TDUS Competence with Data Scales Predicting State Math Test Scores (N = 1534 students and N = 43 teachers).

	Model 1 (Interaction)		Model 2 (Main Effects)	
Predictors	*Coefficient*	*SE*	*Coefficient*	*SE*
Student level				
MAP	6.39***	.67	3.78	2.11
Teacher level				
Intercept	2564.63***	14.13	2558.84***	43.66
TDUS-Competence	23.49	16.19	40.74	38.97
School Level	37.91	14.36	86.21*	33.51
Interaction	22.09	21.27		
Error Variance				
Level 1 (e)	5943.97		5943.54	
Level 2 (r_0)	7447.46		7461.61	
Level 2 (r_1)	17.22		17.92	

*p < .05. **p < .01. ***p < .001.

School level and TDUS Organizational Supports scale for effects on Math SBA. Model 1 tested the interaction of school level and TDUS Organizational Supports scale scores on the SBA math test outcomes. The interaction term of -12.17 ($d.f.$ = 39, p = 0.58) was found not significant. The interaction term was removed for Model 2 and the main effects were examined. The TDUS Organizational Supports scale score of -1.68 ($d.f.$ = 40, p = 0.94) was found not significant. The School Level term of 62.11 ($d.f.$ = 40, p = 0.005) was found to be significant (see table 20).

Table 20. TDUS Organizational Supports Scale Predicting State Math Test Scores (N = 1534 students and N = 43 teachers).

	Model 1 (Interaction)		Model 2 (Main Effects)	
Predictors	*Coefficient*	*SE*	*Coefficient*	*SE*
Student level				
MAP	6.41***	.69	7.04***	.85
Teacher level				
Intercept	2565.46***	14.42	2580.27***	19.99
TDUS – Org. Sup.	-2.11	16.55	-1.68	20.92
School Level	31.86	17.26	62.11**	20.87
Interaction	-12.17	21.74		
Error Variance				
Level 1 (e)	5945.23		6789.59	
Level 2 (r_0)	7784.23		15926.21	
Level 2 (r_1)	18.40		28.76	

*$p < .05$. **$p < .01$. ***$p < .001$.

Summary of Results by Research Question

The main findings of the two research questions are summarized below.

Research question 1 summary. I found significant predictive relationships between the fall reading MAP assessment and the spring language arts SBA test. I also found significant predictive relationships between the fall math MAP assessment and the spring math SBA test.

Research Question 2 summary. Language arts models that included interaction terms for the five different TDUS scales and school-levels did not produce significant results. Subsequently, the interaction term was removed and the models were rerun.

Main effects from these evaluations also produced non-significant results indicating the TDUS scales are not predictive of SBA outcomes.

Mathematics models that included interaction terms from the five different TDUS scales and school-level terms did not produce significant results. Subsequently, the interaction term was removed, and the models were rerun. Main effects from the TDUS scales were not found to be significant predictors of SBA scores. Main effects for school-level were found to be significant. The presence of significant main effects for the school-level terms suggests that secondary school SBA scores were significantly higher than in primary school.

CHAPTER V

DISCUSSION

How did teachers use interim data in order to improve student outcomes on the Smarter Balanced Assessment test? This question was the central tenet of this study. Several researchers have noted the expansive use of interim assessments as school districts attempt to improve student outcomes (Abrams, 2015; Alonzo, 2016; Konstantopoulos, et al., 2017). That was the case in the setting for this study. The sample school district had been contracting with NWEA to provide MAP assessments to all students, third through eighth grade, three times per year.

In order to gain a better understanding how staff used MAP interim data, I surveyed a purposive sample of teachers with the Teacher Data Use Survey. Then, with extant MAP and SBA scores, as well as the TDUS results, I used HLM models to evaluate the three sources of data. HLM proved to be an appropriate approach due to the nested structure of the data and that the math and language arts unconditional models at level-1 (student level) indicated that a large amount of variance remained to be explained at level-2 (teacher level).

Research question 1

With the first research question, significant results were found for both the math and language arts MAP assessment results predicting SBA test results at level-1. This relationship is well established in the literature (Ball, 2016; Northwest Evaluation Association, 2017). These results underpin the use of MAP assessments as a reliable data source for teachers and school administrators when considering if students are on track to meet expected outcomes on the SBA tests. The fall MAP testing window in the sample school district closes in mid-October. Therefore, each classroom teacher, third through eighth grade, has access to valid and reliable information for each student's performance and trajectory outcomes (for SBA) in the content areas of math and language arts. How teachers made use of this data was focus of research question 2.

Research Question 2

Teacher attitudes and behaviors were measured according to the Teacher Data Use survey. The survey included four scales, comprised of 10 subscales, that asked questions pertaining to different teacher actions with data, teacher's attitudes towards data, teacher competence with data, and organizational supports for data use. The results for the TDUS, which were in the form of scale averages, as well as a composite score for the entire survey, became predictors in level-2 (teacher) of the analytical models.

The level-2 model also included interaction terms to test a working theory that the school-level variable would moderate the effects of the TDUS. Much is often made of the differences between the conditions of a primary teaching assignment (teaching approximately 25 students all subjects throughout the day) and secondary teaching assignment (teaching a single subject to approximately 125 students throughout the day). At the conclusion of the level-2 model evaluations, school-level did not moderate the TDUS with either the language arts or the math SBA results.

Once the interaction terms were removed and the model rerun, the main effects were examined. Again, no scales from the TDUS were found to be significant predictors of student outcomes. School level was found to be significant as a main effect across all math HLM evaluations at level-2. This indicated that there was a significant difference between the two school levels, which may be as simple as the SBA scores were higher for secondary teachers (as we would expect). Additionally, all middle schools in this study organized math courses by ability. Due to this organization of instruction, students within the middle schools are receiving high school mathematics content and instruction. This is not the case for language arts instruction. This difference between math and language arts instruction could also be a factor in the difference of significance in the results.

I was surprised that the scales from the Teacher Data Use Survey did not generated any predictability correlations with significance from the level-2 models. It would seem that understanding how teachers reported their actions, attitudes, competence and rated organizational supports for using data would be more indicative on student outcomes than I was able to prove. Exploratory factor analysis revealed a 10 factor structure to the survey. Two scales were found to map on the same factor, but otherwise, the subscales mapped fairly cleanly onto different factors. Additionally, reliability within and correlations between the survey's scales were tenable. A possible explanation for the lack of significance at level-2 within the models may not lie with the TDUS instrument but rather with what was reveled about the use of interim data by teachers.

Teacher Use of Data Survey

Upon closer examination of the TDUS results, as displayed in table 10, the lowest average of all the subscales was the items that related to interim assessments ($M = 1.21$, $SD = .37$). The Actions with Interim Data subscale average nearly a point lower than the next lowest scale across the instrument. The other subscales of the Actions with Data scale reveal a similar pattern of low support for the regular use of different types of data including the use of publisher data ($M = 2.25$, $SD = .63$) and teacher developed data ($M = 2.24$, $SD = .63$). The results of the survey regarding the Actions with Data scale are conclusive ($M = 1.95$, $SD = .37$), indicating that teachers in the sample did not use various forms of data regularly in their practice. Following this logic, if teachers are not using interim data regularly (such as the MAP assessment results), we would not expect to see the Actions with Data scale generating predictive correlations with student outcomes.

It is more difficult to reconcile the remaining TDUS results and the relatively more positive average scores on the other scales. Teachers rated the Organizational Supports scale at an average of 2.6 (SD = .46), Attitudes towards Data was an average of 3.03 (SD = .50), and Competence with Data at 3.16 (SD = .42). It is difficult to understand how these scales did not generate more significant results in the models. They do, however, present a favorable context for the sample school district's future ability to leverage the results of the MAP assessment for improving student outcomes. Teachers report high competence levels in using data (M = 3.56, .57) and support a relatively positive view of data's effectiveness for pedagogy (M = 3.05, .50). This combination of survey results shines a bright light on future professional development opportunities to incorporate powerful data, such as the MAP results into regular practice. In fact, 90% of all teachers in the sample either agreed or strongly agreed with the statement, "Using data helps me be a better teacher."

Educational Implications

Implications from this study fall into three categories. The first is the use of MAP assessment results to guide instruction. The MAP assessment is a strong predictor of expected student outcomes. It would seem valuable for schools and districts to prioritize future professional development on interpretation and best practices of incorporating teacher interpretations of the MAP results into actions related to planning and classroom level assessment.

Secondly, teachers in the sample indicated that many of the variables needed to change practice or orientation towards the use interim data are intact. Teachers reported positive attitudes towards data, high levels of competence in working with data and also strongly rated data's effectiveness for pedagogy. These teacher survey responses provide a strong positive outlook for improving the use of data within the school district organization.

Lastly, organizational supports was rated nearly in the center of the scale (M = 2.6), and several comments from teachers on the survey indicated that staff could benefit from principals prioritizing time and resources for using data more effectively in their practice. These responses included dedicated training and removing unessential tasks from teacher workloads.

Validity of study

When interpreting outcomes from a study that utilizes a predictive correlational relationship, one must exercise caution since the demonstration of a relationship between two groups does not imply that the relationship is causal (Creswell & Creswell, 2018). This non-experimental design allowed me to observe the variables outlined in the study, and as such, there is little control exercised over internal threats to validity. Because the timeframe is different between the extant student data and the teacher survey data, the threat of history becomes a factor as teachers may have changed jobs or left the district since the data was collected. Teachers' attitudes and beliefs about using data could shift over time. However, validity and reliability was established in the study for both of the measures utilized as well as the data analytic plan.

Babbie (2012) cautions against over generalizing the results inferred from a purposive sample. The outcomes and findings of my study are to be used by the sample school district to better understand how the teaching staff are using data from their interim assessment system, the impact that data practices are making in student outcomes on the SBAC, and what structures are supporting this practice. The results of this study could also be generalizable to other similar school districts. Overall, caution may need to be exercised with generalizing the results of this study due to the purposive sample and the non-experimental design.

Limitations

As discussed early, the correlation assumptions in the design of the study cannot be utilized for causal conclusions. With approximately half of the purposive sample completing the survey, caution must be considered when utilizing the results of the models or the TDUS scale outcomes. This study only represents a small step forward in research designs and sources of data to model level-2 teacher influences on level-1 student outcomes. More research and possibly other survey tools are required to identify teacher practices and beliefs that impact student outcomes. These tools, as well as experimental designed studies could possibly explain causality between teacher data beliefs and behaviors that influence student outcomes.

Conclusion

At first glance, the results of this study maybe disappointing. The lack of level-2 predictors from the TDUS scales did feel like a setback. Yet, much was learned. The fact that the MAP assessments predicted both the language arts and math SBA scores, while not new information, was still powerful. Teaching and learning is terrifically complex and difficult. Teachers are constantly striving to improve their practice so that their students can achieve more each year. Teachers also want tried and true approaches.

The evaluation of MAP has proved that the assessment is both tried and true. Administrators in the sample district are armed with two powerful sets of information; the MAP evaluations from data derived from their own students and the positive outcomes regarding competence and attitudes towards data from the TDUS. This combination of information could create the conditions for targeted professional development that may potentially close the gap between research and practice resulting in improved student outcomes.

APPENDIX A

Literature Search Process and Criteria

My literature search process began by accessing the UO + Summit library search function of the University of Oregon Library. Keys words searched included interim assessments, data driven decisions, middle school, public school, and primary school. Inclusion criteria were peer-reviewed journals, education, secondary education and accountability. I excluded anonymous authors, medical index, psychology, high school and college/university.

I wrote to and corresponded with two researches. The first was Dr. Wayman, the lead researcher and first author of the Teacher Use of Data Survey. Dr. Wayman provided invaluable background information about the survey tool as well as provided me with additional research related to the topic. The second was Dr. Lockton, whom I located through a description of her presentation of a paper at the 2018 American Educational Research Association. Dr. Lockton provided me with further research citations.

I initially selected the date range of 2009 through 2019 to reflect the current trends, assessments, and priorities of data use in public schools. However, I did additional reference material that was published prior to 2009 as this research seemed seminal to the background of my study.

APPENDIX B

University of Oregon IRB Approval

UNIVERSITY OF OREGON

DATE: January 23, 2020 **IRB Protocol Number: 01102020.011**

TO: Joel Sebastian, Principal Investigator
Department of Educational Methodology, Policy and Leadership

RE: Protocol entitled, "Teacher Data Use: Impact from Interim Assessments and Student Outcomes"

Notice of Review and Exempt Determination

The above protocol has been reviewed and determined to qualify for exemption. The research is approved to be conducted as described in the attached materials. Any change to this research will need to be assessed to ensure the study continues to qualify for exemption, therefore an amendment will need to be submitted for verification prior to initiating proposed changes.

For this research, the following determinations have been made:

- **This study has been reviewed under the 2018 Common Rule (45 CFR 46) and determined to qualify for exemption under Title 45 CFR 46.104(d)(1).**

Approval period: January 23, 2020 - January 22, 2021

If you anticipate the research will continue beyond the approval period, you must submit a Progress Report at least 45-days in advance of the study expiration. **Without continued approval, the protocol will expire on January 22, 2021 and human subject research activities must cease.** A closure report must be submitted once human subject research activities are complete. Failure to maintain current approval or properly close the protocol constitutes non-compliance.

You are responsible for the conduct of this research and adhering to the Investigator Agreement as reiterated below. You must maintain oversight of all research personnel to ensure compliance with the approved protocol.

The University of Oregon and Research Compliance Services appreciate your commitment to the ethical and responsible conduct of research with human subjects.

Sincerely,

Daniel Berman, MS
Research Compliance Administrator

CC: Mark Van Ryzin

APPENDIX C

School District Authorization of Study

WEST LINN - WILSONVILLE SCHOOL DISTRICT

January 10, 2020

This letter verifies that Joel Sebastian has the permission and support of the West Linn-Wilsonville School District to gain access to middle school student assessment data that includes OSAS/SBAC scores and NWEA Measures of Academic Progress (MAP) scores for the 2018-2019 school year. The district also extends permission for a middle school teacher survey designed as part of Mr. Sebastian's data collection tools and methodology. Mr. Sebastian is conducting this research for his Doctor of Education in Educational Leadership dissertation.

Mr. Sebastian has consulted with me and with the district Director of Information and Technology, provided a detailed description of his methodology and established how data will be encrypted and ensure anonymity for all students, teachers and schools. The West Linn-Wilsonville School District approves this research and supports Mr. Sebastian with the data collection and analysis.

Respectfully,

Barb Soisson

Barbara Soisson, D.Ed.
Assistant Superintendent
Teaching and Learning

APPENDIX D

Survey Script

ATHEY CREEK MIDDLE SCHOOL

Character • Community • Excellence

Tualatin, Oregon 97062
(503) 673-7400

Dear WLWV Educator,

In case we have not met, I am your colleague from Athey Creek Middle School where I serve as the principal. I am currently finishing my dissertation and my goal is to understand the roll of data in teachers' professional practice. More specifically, I am interested in the use of MAP, as well as classroom level data, and if it has an influence on student achievement (the SBAC test).

Your responses will be completely confidential. However, please know that in order for me to run the statistical analysis required to determine which data practices are most promising for improving student achievement, I need to link your survey responses to your students' actual achievement scores. That's why I request your school district email in the first question. Once that link is made, all student, school, and teacher identifiable information will be removed completely from the data file. Additionally, the analysis plan and all results will only be reported in the aggregate.

I hope that this research adds to the body of knowledge related to the most promising teacher data use practices. Additionally, I'll be able to evaluate how well the MAP assessment predicts the spring SBAC scores with our own student population. My intention is to report back to you the results of the study, as well as possible professional development considerations for our school district.

Thank you so much for taking the approximately 10 minutes to complete the survey. I am very appreciative of your professionalism and your support. Please feel free to contact me if you have any questions or concerns.

Sincerely yours,

Joel Sebastian
Principal
Email- sebastij@wlwv.k12.or.us
Cell- 971-832-2159

APPENDIX E

List of The Teacher Data Use Survey Items

1. Where the following forms of data available to you over the last year?

Form of data	Yes	No
Smarter Balanced	■ _	■ _
MAP	■ _	■ _
Publisher Created	■ _	■ _
Teacher Created	■ _	■ _
Other	■ _	■ _

2. Teachers use all kinds of information (i.e., data) to help plan for instruction that meets student learning needs. How frequently did you use the following forms of data over the last year?

Form of data	Do not use	Less than once a month	Once or twice a month	Weekly or almost weekly	A few times a week
Smarter Balanced	■ _	■ _	■ _	■ _	■ _
MAP	■ _	■ _	■ _	■ _	■ _
Publisher Created	■ _	■ _	■ _	■ _	■ _
Teacher Created	■ _	■ _	■ _	■ _	■ _
Other	■ _	■ _	■ _	■ _	■ _

3. If you marked the “other” option above, please specify the form of data here:

4. Now, how useful were the following forms of data to your practice over the last year?

Form of data	Not useful	Somewhat useful	Useful	Very useful
Smarter Balanced	■ _	■ _	■ _	■ _
MAP	■ _	■ _	■ _	■ _
Publisher Created	■ _	■ _	■ _	■ _
Teacher Created	■ _	■ _	■ _	■ _
Other	■ _	■ _	■ _	■ _

5. If you marked the “other” option above, please specify the form of data here:

6. These questions ask about the Smarter Balanced Assessment (SBA). Over the last year school year, how often did you do the following?

Action	**One or two times a year**	**A few times a year**	**Monthly**	**Weekly**
a. Use SBA to identify instructional content to use in class.	■ _	■ _	■ _	■ _
b. Use SBA to tailor instruction to individual students' needs.	■ _	■ _	■ _	■ _
c. Use SBA to develop recommendations for additional instructional support.	■ _	■ _	■ _	■ _
d. Use SBA to form small groups of students for targeted instruction.	■ _	■ _	■ _	■ _
e. Discuss SBA with a parent or guardian.	■ _	■ _	■ _	■ _
f. Discuss SBA with a student.	■ _	■ _	■ _	■ _
g. Meet with a specialist (e.g., instructional coach or data coach) about SBA.	■ _	■ _	■ _	■ _
h. Meet with another teacher about SBA.	■ _	■ _	■ _	■ _

Items adapted from Wayman, J. C., Cho, V., & Shaw, S. (2009). *Survey of Educator Data Use.* Unpublished instrument.

7. These questions ask about the Measure of Academic Progress (MAP) used in your school or district. In a typical month, how often did you do the following over the last year?

Action	**Less than once a month**	**Once or twice a month**	**Weekly or almost weekly**	**A few times a week**
a. Use MAP to identify instructional content to use in class.	■ _	■ _	■ _	■ _
b. Use MAP to tailor instruction to individual students' needs.	■ _	■ _	■ _	■ _
c. Use MAP to develop recommendations for additional instructional support.	■ _	■ _	■ _	■ _
d. Use MAP to form small groups of students for targeted instruction.	■ _	■ _	■ _	■ _
e. Discuss MAP with a parent or guardian.	■ _	■ _	■ _	■ _
f. Discuss MAP with a student.	■ _	■ _	■ _	■ _
g. Meet with a specialist (e.g., instructional coach or data coach) about MAP.	■ _	■ _	■ _	■ _
h. Meet with another teacher about MAP.	■ _	■ _	■ _	■ _

Items adapted from Wayman, J. C., Cho, V., & Shaw, S. (2009). *Survey of Educator Data Use.* Unpublished instrument.

8. These questions ask about data derived from Publisher Created assessments used in your school or district. In a typical month, how often did you do the following over the last year?

Action	Less than once a month	Once or twice a month	Weekly or almost weekly	A few times a week
a. Use data from publisher created tests to identify instructional content to use in class.	■ _	■ _	■ _	■ _
b. Use data from publisher created tests to tailor instruction to individual students' needs.	■ _	■ _	■ _	■ _
c. Use data from publisher created tests to develop recommendations for additional instructional support.	■ _	■ _	■ _	■ _
d. Use data from publisher created tests to form small groups of students for targeted instruction.	■ _	■ _	■ _	■ _
e. Discuss data from publisher created tests with a parent or guardian.	■ _	■ _	■ _	■ _
f. Discuss data from publisher created tests with a student.	■ _	■ _	■ _	■ _
g. Meet with a specialist (e.g., instructional coach or data coach) about data from publisher created tests.	■ _	■ _	■ _	■ _
h. Meet with another teacher about data from publisher created tests.	■ _	■ _	■ _	■ _

Items adapted from Wayman, J. C., Cho, V., & Shaw, S. (2009). *Survey of Educator Data Use.* Unpublished instrument.

9. These questions ask about data derived from teacher created assessments. In a typical month, how often did you do the following over the last year?

Action	Less than once a month	Once or twice a month	Weekly or almost weekly	A few times a week
a. Use data from teacher created tests to identify instructional content to use in class.	■ _	■ _	■ _	■ _
b. Use data from teacher created tests to tailor instruction to individual students' needs.	■ _	■ _	■ _	■ _

c. Use data from teacher created tests to develop recommendations for additional instructional support.	■ _	■ _	■ _	■ _
d. Use data from teacher created tests to form small groups of students for targeted instruction.	■ _	■ _	■ _	■ _
e. Discuss data from teacher created tests with a parent or guardian.	■ _	■ _	■ _	■ _
f. Discuss data from teacher created tests with a student.	■ _	■ _	■ _	■ _
g. Meet with a specialist (e.g., instructional coach or data coach) about data from teacher created tests	■ _	■ _	■ _	■ _
h. Meet with another teacher about data from teacher created tests.	■ _	■ _	■ _	■ _

The remainder of this survey asks general questions about the use of data to inform your education practice over the last year. For the rest of this survey, please consider only the following when you are asked about "data":

- *Smarter Balanced Assessments.*
- *MAP Assessment.*
- *Publisher Created Tests.*

10. These questions ask about supports for using data. Please indicate how much you agree or disagree with the following statements:

Statement	**Strongly disagree**	**Disagree**	**Agree**	**Strongly agree**
a. I am adequately supported in the effective use of data.	■ _	■ _	■ _	■ _
b. I am adequately prepared to use data.	■ _	■ _	■ _	■ _
c. There is someone who answers my questions about using data.	■ _	■ _	■ _	■ _
d. There is someone who helps me change my practice (e.g., my teaching) based on data.	■ _	■ _	■ _	■ _
e. My district provides enough professional development about data use.	■ _	■ _	■ _	■ _
f. My district's professional development is useful for learning about data use.	■ _	■ _	■ _	■ _

11. These questions ask about your attitudes and opinions regarding data. Please indicate how much you agree or disagree with the following statements:

Statement	**Strongly disagree**	**Disagree**	**Agree**	**Strongly agree**
a. Data help teachers plan instruction.	■ _	■ _	■ _	■ _
b. Data offer information about students that was not already known.	■ _	■ _	■ _	■ _

c. Data help teachers know what concepts students are learning.	■ _	■ _	■ _	■ _
d. Data help teachers identify learning goals for students.	■ _	■ _	■ _	■ _
e. Students benefit when teacher instruction is informed by data.	■ _	■ _	■ _	■ _
f. I think it is important to use data to inform education practice.	■ _	■ _	■ _	■ _
g. I like to use data.	■ _	■ _	■ _	■ _
h. I find data useful.	■ _	■ _	■ _	■ _
i. Using data helps me be a better teacher.	■ _	■ _	■ _	■ _

12. These questions ask how your principal and assistant principal(s) support you in using data. Principals and assistant principals will not be able to see your answers. Please indicate how much you agree or disagree with the following statements:

Statement	**Strongly disagree**	**Disagree**	**Agree**	**Strongly agree**
a. My principal or assistant principal(s) encourages data use as a tool to support effective teaching.	■ _	■ _	■ _	■ _
b. My principal or assistant principal(s) creates many opportunities for teachers to use data.	■ _	■ _	■ _	■ _
c. My principal or assistant principal(s) has made sure teachers have plenty of training for data use.	■ _	■ _	■ _	■ _
d. My principal or assistant principal(s) is a good example of an effective data user.	■ _	■ _	■ _	■ _
e. My principal or assistant principal(s) discusses data with me.	■ _	■ _	■ _	■ _
f. My principal or assistant principal(s) creates protected time for using data.	■ _	■ _	■ _	■ _

13. Your school or district gives you programs, systems, and other technology to help you access and use student data. The following questions ask about these computer systems. Please indicate how much you agree or disagree with the following statements:

Statement	**Strongly disagree**	**Disagree**	**Agree**	**Strongly agree**
a. I have the proper technology to efficiently examine data.	■ _	■ _	■ _	■ _
b. The computer systems in my district provide me access to lots	■ _	■ _	■ _	■ _

of data.				
c. The computer systems (for data use) in my district are easy to use.	■ _	■ _	■ _	■ _
d. The computer systems in my district allow me to examine various types of data at once (e.g., attendance, achievement, demographics).	■ _	■ _	■ _	■ _
e. The computer systems in my district generate displays (e.g., reports, graphs, tables) that are useful to me.	■ _	■ _	■ _	■ _

14. These questions ask about your attitudes toward your own use of data. Please indicate how much you agree or disagree with the following statements:

Statement	**Strongly disagree**	**Disagree**	**Agree**	**Strongly agree**
a. I am good at using data to diagnose student learning needs.	■ _	■ _	■ _	■ _
b. I am good at adjusting instruction based on data.	■ _	■ _	■ _	■ _
c. I am good at using data to plan lessons.	■ _	■ _	■ _	■ _
d. I am good at using data to set student learning goals.	■ _	■ _	■ _	■ _

The following questions ask about your work in collaborative teams.

15. How often do you have scheduled meetings to work in collaborative team(s) over the last year? (Check only one.)

■ Less than once a month.
■ Once or twice a month.
■ Weekly or almost weekly.
■ A few times a week.
■ I do not have scheduled meetings to work in collaborative teams.

16. As you think about your collaborative team(s), please indicate how much you agree or disagree with the following statements:

Statement	**Strongly disagree**	**Disagree**	**Agree**	**Strongly agree**
a. Members of my team trust each other.	■ _	■ _	■ _	■ _
b. It's ok to discuss feelings and worries with other members of my team.	■ _	■ _	■ _	■ _
c. Members of my team respect colleagues who lead school improvement efforts.	■ _	■ _	■ _	■ _
d. Members of my team respect those colleagues who are experts in	■ _	■ _	■ _	■ _

their craft.				
e. My principal or assistant principal(s) fosters a trusting environment for discussing data in teams.	■ _	■ _	■ _	■ _

17. How often do you and your collaborative team(s) do the following over the last year?

Action	**Never**	**Sometimes**	**Often**	**A lot**
a. We approach an issue by looking at data.	■ _	■ _	■ _	■ _
b. We discuss our preconceived beliefs about an issue.	■ _	■ _	■ _	■ _
c. We identify questions that we will seek to answer using data.	■ _	■ _	■ _	■ _
d. We explore data by looking for patterns and trends.	■ _	■ _	■ _	■ _
e. We draw conclusions based on data.	■ _	■ _	■ _	■ _
f. We identify additional data to offer a clearer picture of the issue.	■ _	■ _	■ _	■ _
g. We use data to make links between instruction and student outcomes.	■ _	■ _	■ _	■ _
h. When we consider changes in practice, we predict possible student outcomes.	■ _	■ _	■ _	■ _
i. We revisit predictions made in previous meetings.	■ _	■ _	■ _	■ _
j. We identify actionable solutions based on our conclusions.	■ _	■ _	■ _	■ _

18. What else would you like to share with us about data use?

REFERENCES CITED

Abrams, L., McMillan, J., & Wetzel, A. (2015). Implementing benchmark testing for formative purposes: Teacher voices about what works. *Educational Assessment, Evaluation and Accountability, 27*, 347-375.

Abrams, L., Varier, D., & Jackson, L. (2016). Unpacking instructional alignment: The influence of teachers' use of assessment data on instruction. *Perspectives in Education, 34*, 15-28.

Alonzo, J. (2016). The relation between easyCBM and Smarter Balanced reading and mathematics assessments. *Journal of School Administration Research and Development, 1*, 17-35.

Babbie, E. (2012). *The Practice of Social Research* (13th ed.). Belmont, CA: Wadsworth.

Ball, C. R., & O'Connor, E. (2016). Predictive utility and classification accuracy of oral reading fluency and the Measures of Academic Progress for the Wisconsin Knowledge and Concepts Exam. *Assessment for Effective Intervention, 41*, 195–208. doi:10.1177/1534508415620107

Beaver, J., & Weinbaum, E. (2015). State test data and school improvement efforts. *Educational Policy, 29*, 478-503. doi:10.1177/0895904813510774

Braun, H. (2011). Can road maps tell us whether we are off course? *Measurement: Interdisciplinary Research and Perspectives, 9*, 130-133.

Creswell, J., & Creswell, J., D. (2018). *Research design : Qualitative, quantitative, and mixed methods approaches* (Fifth ed.). Thousand Oaks, California: SAGE Publications.

Cusick, P. A. (2014). The logic of the U.S. educational system and teaching. *Theory Into Practice, 53*, 176-182.

Datnow, A., Park, V., & Kennedy-Lewis, B. (2013). Affordances and constraints in the context of teacher collaboration for the purpose of data use. *Journal of Educational Administration, 51*, 341-362.

DuFour, R. (2015). *In praise of American educators and how they can become even better*. Bloomington: Solution Tree.

DuFour, R., DuFour, R.B., & Eaker, R. E. (2008). *Revisiting professional learning communities at work : New insights for improving schools*. Bloomington: Solution Tree.

Dee, T., & Jacob, B. (2010). The impact of No Child Left Behind on students, teachers, and schools. *Brookings Papers On Economic Activity,* 149-207.

Garner, B., Thorne, J. K., & Horn, I. S. (2017). Teachers interpreting data for instructional decisions: Where does equity come in? *Journal of Educational Administration, 55*, 407-426.

Gill, B., Borden, B.C., & Hallgren, K. (2014). A conceptual framework for data-driven decision making (Report No. 40019.190). Princeton, NJ: Mathematica Policy Research. Retrieved from https://www.mathematica-mpr.com

Gutierrez, R. (2008). A "Gap-Gazing" fetish in mathematics education? Problematizing research on the achievement gap. *Journal For Research In Mathematics Education*, *39*, 357-364.

Kallemeyn, L. M. (2014). School-level organizational routines for learning: Supporting data use. *Journal of Educational Administration*, 52, 529-548.

Klingbeil, D. A., Van Norman, E. R., Nelson, P. M., & Birr, C. (2018). Evaluating screening procedures across changes to the statewide achievement test. *Assessment for Effective Intervention, 44*, 17–31. doi:10.1177/1534508417747390

Kippers, W. B., Wolterinck, C. H. D., Schildkamp, K., Poortman, C. L., & Visscher, A. J. (2018). Teachers' views on the use of assessment for learning and data-based decision making in classroom practice. *Teaching and Teacher Education*, *75*, 199-213. doi:10.1016/j.tate.2018.06.015

Konstantopoulos, S., Li, W., Miller, S. R., & Van der Ploeg, A. (2017). Do interim assessments reduce the race and SES achievement gaps? *Journal of Educational Research, 110*, 319-330. doi:10.1080/00220671.2015.1103685

Marsh, J. A., McCombs, J. S., & Martorell, F. (2010). How instructional coaches support data-driven decision making: Policy implementation and effects in Florida middle schools. *Educational Policy, 24*, 872-907.

McGuinn, P. (2016). From No Child Left Behind to The Every Student Succeeds Act: Federalism and the education legacy of the Obama administration. *Publius: The Journal of Federalism, 46*, 392-415.

Northwest Evaluation Association. (2017). *Linking the Smarter Balanced Assessments to NWEA MAP Growth Tests.* Retrieved from Northwest Evaluation Association website: https://www.nwea.org/content/uploads/2015/06/SBAC-MAP-Growth-Linking-Study_OCT2017.pdf

Raudenbush, S. W., & Bryk, A., (2002). *Hierarchical linear models: Applications and data analysis methods*. Newbury Park: Sage Publications.

Remesal, A. (2011). Primary and secondary teachers' conceptions of assessment: A qualitative study. *Teaching and Teacher Education, 27*, 472-482. doi:10.1016/j.tate.2010.09.017

Riggan , M. & Oláh L. N. (2011) Locating interim assessments within teachers' assessment practice, *Educational Assessment, 16*, 1-14. doi:10.1080/10627197.2011.551085

Schildkamp, K., & Kuiper, W. (2010). Data-informed curriculum reform: Which data, what purposes, and promoting and hindering factors. *Teaching and Teacher Education: An International Journal of Research and Studies, 26*, 482-496. doi:10.1016/j.tate.2009.06.007

Spiegelhalter, D. (2019). *The art of statistics: How to learn from data*. New York: Basic Books.

Spillane, J. P. (2012). Data in practice: Conceptualizing the data-based decision-making phenomena. *American Journal of Education, 118*, 113-141. doi:10.1086/663283

Spillane, J. P., & Jennings, N. E. (1997). Aligned instructional policy and ambitious pedagogy: Exploring instructional reform from the classroom perspective. *Teachers College Record, 98*, 449-481.

Wang, H.Y., Zhao, H., & Addison, K. (2016). *Linking PARCC and MAP Assessments for Students in Montgomery County Public Schools.* Rockville, MD: Montgomery County Public Schools.

Wayman, J. C., Cho, V., Jimerson, J., & Spikes, D. D. (2012). District-wide effects on data use in the classroom. *Education Policy Analysis Archives*, 20 (25), 1-31.

Wayman, J., & Jimerson, J. (2014). Teacher needs for data-related professional learning. *Studies in Educational Evaluation, 42*, 25-34.

Wayman, J. C., Shaw, S., & Cho, V. (2017). Longitudinal effects of teacher use of a computer data system on student achievement. *AERA Open*, *1*, 1-18. doi:10.1177/2332858416685534

Wayman, J. C., Wilkerson, S. B., Cho, V., Mandinach, E. B., & Supovitz, J. A. (2016). *Guide to using the Teacher Data Use Survey* (REL 2017–166). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Appalachia. Retrieved from http://ies.ed.gov/ncee/edlabs.

Wayman, J. C., Wilkerson, S. B., Cho, V., Mandinach, E. B., & Supovitz, J. A. (April 2017). *Helping administrators get data on how teachers use data: The teacher data use survey*. Paper presented at the meeting of the American Educational Research Association, San Antonio.