Use of Data to Inform Instruction in an Urban School System and the Impact on Standardized Assessments

By Jessica Hargest

Submitted in Partial Fulfillment of the Requirements for the

Degree of Master of Education

July 2017

Graduate Programs in Education

Goucher College

Table of Contents

List of Tables

Abstract

Data are increasingly used by teachers and schools to inform instruction. Inconsistency in its use likely affects the clinical utility of data collection and applications for improving teaching and learning outcomes. This study sought to determine whether Partnership for Assessment of Readiness for College and Careers (PARCC) scores for students in grades three to eight differed significantly among schools with varied levels of data use. Data usage was rated based on a qualitative School Effectiveness Review process which reviewed 20 schools in a large urban district. The study found no statistically significant relationship between the level of data use and PARCC scores for the schools. Hence, the null hypotheses, that the mean PARCC scores would not differ significantly among schools with varied levels of data use and that the correlation between PARCC scores and the schools' levels of data use would be insignificant, both were retained. Despite these findings, the review of literature and current demands for accountability in education indicate further studies of how data collection and application to teaching and learning can most effectively benefit students are warranted.

CHAPTER I

INTRODUCTION

Overview

Increasing numbers of schools across the nation are utilizing data to inform instruction. With the use of technology, students' scores on various types of assessments are calculated quickly and areas of need are identified. The proficiency of a school system to be able to use data varies widely, and even when a system decides to use data to inform instruction, the ability of the schools and teachers within that system to use the data varies greatly (Griswold, 2005; Marsh, Pane, & Hamilton, 2006). While many areas of research, particularly those areas that do not need to account for large variation in populations, can give definitive, well-defined answers to questions, questions in education often do not have clear cut answers. Therefore, efforts to determine appropriate interventions to address identified issues, even those using data intended for that purpose, may or may not result in growth in student achievement.

Accountability in school systems also is becoming increasingly reliant on the use of data (Griswold, 2005). Various facets of education and related programs appear to require data to justify funding and even their existence. While accountability certainly is necessary, if school systems and teachers are unable to use data appropriately, the efficiency of data collection and its use is compromised, as is the likelihood of generating desirable outcomes.

Statement of Problem

The researcher was interested in learning about how teachers' use of data to inform instruction in an urban school system relates to student outcomes on a standardized state assessment. It appeared from her professional experience and review of the literature that methods and effectiveness of data use across schools likely were inconsistent.

The researcher rated schools' use of data in grades three through eight by combining Likert scale ratings of two School Effectiveness Review indicators of one key action to yield a score reflecting an estimate of the degree to which "Teachers use data sources to adjust practice" (hereafter referred to as DATA USE). The key action was "Teachers use multiple data sources to adjust practice." Indicator one was "Teachers analyze student progress towards goals" and Indicator two was "Teachers modify instruction in response to data." Schools' aggregated mean Partnership for Assessment of Readiness for College and Careers (PARCC) scores (combined for mathematics and English Language Arts [ELA] for all students in grades three through eight) were then compared across levels of DATA USE. Those variables also were correlated to determine whether schools' use of data and test outcomes were related significantly.

Hypotheses

It was hypothesized that schools with robust, well developed data processes, as determined by the School Effectiveness Review (SER) (see the operational definition), would have better outcomes on the standardized state tests of mathematics and ELA. The null hypotheses, which follow, were tested using a one-way Analysis of Variance to compare the schools' mean aggregated mathematics and ELA PARCC scores (which could range from one to five) across low, medium and high levels of data use.

ho_1: *Mean PARCC scores for Low Data Use = Mean PARCC scores for Medium Data Use = Mean PARCC scores for High Data Use* and ho_2: *(PARCC scores and School Data Use) = 0*

Operational Definitions

School Effectiveness Review (SER) is a qualitative review process that all schools in this urban school setting undergo on a three-year cycle. The process uses a qualitative framework, and requires evidence to be collected through focus groups of teachers, school leadership,

students, parents, community partners, and student support staff; observations of 60 percent of classrooms, using a binary scale to note if particular actions of teachers are observed or not observed and through a review of requested documentation that supports statements made by school leadership. For the purposes of this study, two indicators or statements supporting a broader statement were utilized. These indicators were "Teachers analyze student progress towards goals" and "Teachers modify instruction in response to data." The two indicators were rated by the researcher and those ratings were totaled so that each school received a rating on a 10-point scale of the implementation of teachers using data sources to adjust practice.

Partnership for Assessment of Readiness for College and Careers (PARCC) is a standardized state assessment used to gauge school performance across several states, and is used by the state of Maryland. All students in grades three through twelve take this assessment annually. For the purposes of this study, scores for English Language Arts (ELA) and mathematics in grades three through eight were used. The PARCC assessment in grades nine through twelve were not used as students in those grades are tested by specific content areas, such as Algebra and Geometry.

CHAPTER II

REVIEW OF THE LITERATURE

Introduction

This literature review discusses ways to use data to improve instructional outcomes and inform educators regarding using data to inform instructional decisions. The review begins by describing increased use of data in making educational decisions. Challenges for using data effectively are explored in part two of the literature review. Part three discusses ways to assist school systems to use data effectively.

Increase in Educators' Use of Data in Instructional Decision Making

In recent years, data-based decisions have become a mainstay at all levels of education. From the national and state level to the classroom level, data are being used to inform decisions about education. According to Griswold (2005), one reason for this trend in education relates to the No Child Left Behind Act of 2002 (NCLB). This act required that all schools and school systems must make Adequate Yearly Progress (AYP) towards a twelve-year goal of all students being at or above grade level in reading, language arts, math and science. The NCLB required that states utilize standardized state tests to measure student achievement towards these goals. In addition, the NCLB required states to set attendance and graduation goals for students. When schools failed to meet these goals, the NCLB required research-based interventions to be implemented. With high levels of accountability, and evidence of growth or stagnation revealed by data, school systems began to use data in every aspect and level of instruction to make decisions about instruction.

External and Internal Applications of Data

At the school system level, data can be used externally and internally. Externally, data and standardized testing results help school districts inform state education agencies and the public about the relative effectiveness of schools. Secondly, at a system level, data can support state education agencies' decisions to provide additional supports and interventions to schools that are failing and to reward schools that are performing well. Finally, data can assist school systems in determining the types of supports that are provided to their schools (Griswold, 2005).

School systems also use data internally. For example, Marsh et al. (2006) stated that schools use state standardized assessments to make decisions about school effectiveness. Standardized tests, however, have an inherent limitation in that the data produced is summative, and typically done at the end of courses. To address this concern, many school systems have increased testing and therefore data collection, by adding additional assessments to make interim decisions as instruction ensues.

Assessing Teacher Effectiveness with Data

School systems also have begun to consider how to use data to assess and evaluate teacher effectiveness. Through Value Added Models (VAMs), school systems are able to account for variances in student populations and isolate levels of teacher effectiveness. However, a study done by McCaffrey, Koretz, Lockwood and Hamilton (2003) reports that teachers' relationship to student outcomes may be nonlinear, and that teachers may be more or less effective depending on the students' ability to achieve. For example, students who may begin below grade level may not achieve as much growth as students who have consistently been on grade level. Research such as that reported by Rivkin, Hanushek, and Kain (2005) also indicates that teachers' effectiveness demonstrates the most improvement after the first year of teaching and then improves by much smaller increments in future years.

Using Data to Inform School Interventions

Schools have used data to inform a wide range of decisions. For example, Choppin (2002) noted that one school he researched used rates of discipline, office referrals, and resource usage data to encourage teachers to challenge their practices and to try more proactive classroom management measures. Schools also are using academic and nonacademic data, such as attendance and discipline data, to determine whether they are causally related to academic performance and to make action plans that appropriately address causal areas which are and are not limited solely to academics.

Benefits and Challenges of Using Data

The use of data provides methods for school systems, individual schools, and teachers to be held accountable, to reflect systematically on practices, and to determine the effectiveness of interventions and action plans. However, data are not always as clear, or as simple to use and interpret as policy makers and data users would like them to be. Sometimes using data to inform programmatic or individual educational decisions ~~can~~ inadvertently can have a negative effect on teaching and learning.

Challenges for Using Data Effectively

One widely expressed criticism of the NCLB Act of 2002 was that AYP used only one indicator, assessment data, to determine schools' failure and success. NCLB did not consider that schools have different resources and funding allocations, historical achievement gaps, and settings which could affect assessment outcomes despite implementing appropriate practices (Celio & Harvey, 2005).

Each school system must consider all implications of utilizing data as indicators of success or failure with regard to improving student achievement, and each system should select

indicators that provide meaningful and accurate depictions of its students' needs and progress. Celio and Harvey (2005) provide examples of indicators of success or failure which are aggregates of data that society understands and suggests that schools use such indicators to indicate what is happening both in society and in schools. One such example was the cost of living indicator. The cost of living indicator has 124 variables, and a national average is developed which allows national level decisions and comparisons to be made by those using it. Schools systems need to use data in a similar way so that causal information is not eliminated. However, systems must proceed with caution to ensure that assumptions being made about the information are fair and accurate given each district's unique characteristics.

Celio and Harvey (2005) suggest that school systems fail at developing strong indicators for several reasons. First, while schools may elicit support from consultants who provide insight into additional data points that could be collected, the overwhelming amount of information obtained can make it difficult for school systems to determine which are the most appropriate indicators for their particular system. Second, snapshots of schools' needs and characteristics tend to be inexpensive and the data are acquired easily, but these data only provide a picture of what currently is happening. Snapshots limit the ability to detect or monitor trends or determine the effective interventions as opposed to cross-sectional and longitudinal studies, which are costlier, but can show changes over time. Third, Celio and Harvey state that systems appropriately or inappropriately may decide to use comparison data versus absolute data. In some instances, comparing two schools in a district may be more important than examining each individual school's overall success in any one measure. For example, two low performing schools both may be categorized as not achieving and may be grouped together for strategic supports. However, upon utilizing a comparison and including more detail, it may be revealed

that one school is stronger than another in certain aspects, thereby informing the level and type of support needed by each school. Finally, Celio and Harvey cite parsimony versus complexity as being an important deciding factor which systems should consider in determining how to use data. Systems can decide to over-simplify and create one data point that indicates passing or failing. Conversely, they can use so many indicators that it becomes difficult to interpret what is needed or what progress has been made, thus hindering decision making and efficient use of resources.

Challenges for Teachers

Teachers have increasing access to large amounts of data, and just as systems, they must understand how data can be used effectively and potentially can be misused. Marsh et al. (2006) cite four ways in which teachers might inadvertently misuse data. First, teachers may review test results from an assessment that is poorly aligned to state standards and make incorrect assumptions about the level of student preparedness for completing the state assessment. Second, teachers may lack the technical or content ability to detect small variances in data which have implications for student outcomes. Third, teachers may utilize insufficient data to make assumptions regarding student progress instead of gathering more data to ensure the data yield an accurate representation of student needs and growth. Finally, systems and politics may influence the way in which individual teachers and schools interpret and analyze certain data points. For example, a school system may have made a large investment in technology, and when a state computer-based standardized test is administered and that school system performs better than a neighboring school system which has not made the same investment, the school system may inaccurately attribute portions of its success to its investment in to technology when other factors influenced the scores.

Consequences of Data Use for Students

Ultimately the outcomes of data use by systems and teachers can affect students in both positive and negative ways. Bertrand and Marsh (2015) suggest that teachers may use data inappropriately due to individual sense-making. In other words, the teachers' own preconceptions and belief systems can influence what actions the teachers take to respond to the data. The authors reflect that sense-making can cause beliefs to be formed and shaped or reshaped. For example, students may have performed poorly on an assessment that measures reading comprehension and the teacher identifies strategies to address comprehension; however, without proper analysis of the cause for the poor performance, a teacher may miss other possibilities. Possibilities in this example could include students not having contextual background information needed to understand the text or students having poor decoding skills that would enable them to understand the text.

McKown and Weinstein (2008) found that as the diversity of their classes increased, teacher bias also increased by as much as 1.5 standard deviations. The bias was determined by teachers ranking expected student performance by the end of the school year at the onset of the school year, and then comparing their rankings to data regarding parent-reported race. These researchers also found that when teachers were in classrooms with lower levels of diversity, teacher bias was decreased.

Bertrand and Marsh (2015) state that students determine attributes of the data in a sense-making process. For example, students may first consider the locus of causality. In this instance, students consider whether the outcome was caused by what the individual did, or was it caused by an external person. Second, students may consider the stability of the data. They may ask if the data represent enduring trend, or if a particular data point is a rarity. Third, Bertrand and

Marsh (2015) suggest that students consider the controllability of the data point. In other words, they may consider if it was within the students' control to change the outcome. Depending on how students perceive these three factors, students' motivation may increase or decrease when interpreting data about their performance.

Teaching School Systems to Use Data Effectively

As previously stated, data collection and use are now required aspects of education. Therefore, it is imperative to teach school systems, and in particular teachers, how to use data appropriately to measure instructional effectiveness. Choppin (2002) reviewed the processes of several schools in Michigan that were at various phases of data implementation. The researchers outlined some key points regarding the implementation of processes for data collection and use. These points included schools ensuring that they can provide the technical training as well as time to implement the data analysis process with integrity. Secondly, Choppin states that schools need to understand that a cultural shift will need to occur regarding data use and its consequences. He suggests that school systems plan for regular intervals of data collection and analysis. In doing so, school administrators need to create and develop a non-defensive learning culture so that teachers feel comfortable learning and adapting to the new processes and feedback they receive.

Training in assessment should be based on research-validated practice. Sometimes ideas that initially appear to be helpful do not show benefits when studied. For example, Horn, Kane, and Wilson (2015) conducted a study in which they compared two schools which each received instructional management training. During the training, one of the two schools received training on conversational reviews of data to make instructional decisions. At the conclusion of the study,

the researchers determined that there was no significant difference in the outcomes between the two schools.

Classroom Considerations

Turner (2014) highlights five essential elements of assessment-centered classroom instruction. The first step is pre-assessment so that a baseline can be established and the teacher has a detailed understanding of what students' current levels of knowledge are, their common misconceptions, if they have the necessary foundational skills, if they already have mastered the topic, and lastly, what level of interest students have in the topic. Step two is to clearly identify the students' strengths and areas of need. Step three is to plan to differentiate instruction based on results from the pre-assessment, while ensuring alignment with standards. The fourth step is to administer a variety of formative assessments throughout the instructional process. Turner also suggests that students should have a clear and explicit understanding of the learning objectives and how their understanding will be measured. Finally, Turner states that step five is providing a summative assessment which should be grounded in standards and administered after ample time has been provided for instruction.

Collaborative data collection can be done through teachers' anecdotal notetaking, student work portfolios, and recording of frequency, rate, duration or latency of particular events. Lingo, Barton-Arwood, and Jolivette (2011) argue that teachers must be aware that there are several ways to collect and analyze data, and these should be aligned with the needs the information is to address. More specifically, the data collection and analysis framework they provide should be applied to the context of general education teachers and special educators working together to ensure that students with special needs are getting their instructional needs satisfied by all teachers. Kallemeyn (2014) also conducted a study that reviewed a school effectively

implementing a new data cycle, and cited several reasons for its successful implementation. First, the school ensured there were effective collaborative teams, and provided a clearly defined process for inquiry for these teams. The collaborative teams met three times weekly, and once a week included a team member who was an instructional coach. The instructional coach was charged with extracting data, as well as providing some basic analysis of the data to both school leadership and teachers. The instructional coach also provided tasks to be completed by the following meeting. While Kallemnyn could not conclude that this process resulted in greater achievement for the students, the author suggests that the routine practice of meeting with the instructional coach and discussing outcomes in a safe environment supported student achievement.

While Kallemeyn (2014) found that teachers reported that they did not feel as though they were able to conduct a data cycle themselves, Huguet, Marsh and Farrell (2014) considered or reviewed ideas about how instructional coaches can build teacher capacity. Huguet suggests that coaches first assess teacher needs and provide modeling, observation, feedback and sharing of ideas, dialogue and questioning as effective tools for building teacher capacity.

Conclusion

In an age of tightening budgets, poorly performing schools, the need for accountability and definitive answers about effective instructional practice, data use is a requirement of every aspect and at all levels of education. However, data can be misused, overly relied upon, and cause users to make false assumptions that can cause as many negative implications as positive ones. Educational leaders and teachers can use data to inform practice and make informed decisions; however, if they are not careful, they may find that they are not providing enough

time, conducting appropriate analysis of data, or that they are overlooking small changes that could affect learning and other important outcomes.

Research such as that cited in this review of literature suggests that data collection yields important evidence about schools' effectiveness. However, it is important to note that user input, knowledge, and expertise still are necessary to make appropriate, well-grounded decisions about how to collect and use data appropriately and effectively. How to best implement data processes and train the current generation of educators and leaders on appropriate data applications still is being determined. The end goal is to help schools use data as a reliable tool to provide a strong foundation when making decisions that affect children's educational growth.

CHAPTER III

METHODS

The purpose of this study is to examine how teachers' use of data to inform instruction in an urban school system related to student outcomes on a standardized state assessment. The study focused on consistency and effectiveness of data use across schools.

Design

The study used a causal-comparative design and a one-way analysis of variance to determine whether there were significant differences between PARCC scores of schools with varied levels of data use and quantified the relationship with a correlation coefficient.

Participants

Twenty public schools in a large urban district were included in this study. Each participated in a School Effectiveness Review (SER) during the 2015/2016 school year, and enrolled students in third through eighth grades. In 2015/2016, there were 35 SERs conducted in schools that were due to have a review as prescribed by a pre-determined three-year cycle or in charter schools that would be involved with a renewal process during the following year. Ten of these schools were excluded as they were high schools, and five were excluded as they were special education separate public day schools.

Instrument

Data were collected as part of a qualitative review process (see operational definition of School Effectiveness Review) and by state mandated standardized testing, specifically, the Partnership for Assessment of Readiness for College and Careers (PARCC).

SER Data Review

SER data utilized for this study were based on two indicators under one key action related to

data usage. For both indicators, the evidence used to determine whether the school had met the criteria included information from focus groups of school leaders and teachers, combined with a review of documentary evidence. For the first indicator 1.3 a "Teachers analyze students' progress towards goals," school leadership and teachers were asked in focus groups about the process used to analyze data, and what requirements, if any, there were for doing so. Both focus groups also were asked what types of data points they review. The SER team reviewed what was stated, in addition to documentation provided by the school, to determine whether the school had begun to analyze data using informal data points such as anecdotal information, or whether it required a rich analysis process that includes components such as a review of data, identifying areas of an individual student's strengths and weaknesses, identifying goals and actions related to student improvement, and an assessment component to determine if the intervention was successful.

The second indicator assessed was 1.3 b "Teachers modify instruction in response to data." To assess this indicator, the site team asked school leadership and teachers to explain how they modified instruction and what data points were being utilized. To determine effectiveness, the SER team reviewed documentation that the schools provided to confirm what the focus groups identified as means of modification. The documentation can be related to a data analysis process in which they identify steps to analyze and develop action steps, or could be related to lesson plans in which teachers have identified instructional groups based on assessment data. In general, the SER team determines ratings based upon the level of implementation of the indicator, in addition to a preponderance of documentation to support the level of implementation across the school. The team, consisting of at least three team members comes to consensus on the rating, and lastly, an external, independent editor, reviews the reports and

ensures consistency of ratings.

PARCC Data Review

The PARCC assessment is an annual, end of year assessment that is administered to third through twelfth grade students in Maryland to assess the areas of English Language Arts (ELA) and mathematics. The PARCC assessment is aligned with the Common Core State Standards. As of the 15/16 school year, the PARCC had been administered in Baltimore City Public Schools for two consecutive years. Students can earn scores on the PARCC ranging from one to five (1 – does not meet expectations, 2 – partially meets expectations, 3- approaching expectations, 4- met expectations, and 5 – exceeded expectations) for both ELA and mathematics. Both ELA and mathematics scores were averaged for the students in grades three through eight in each of the 20 schools included in the study to yield one weighted average PARCC score per school. To calculate the weighted average PARCC scores, the researcher summed the number of students earning each score (1-5), then multiplied each of those sums by the scores they received. Lastly, the sum of the weighted averages was divided by the number of students tested to yield a weighted average PARCC score for each school.

Procedure

Schools that participated in SERs during the 2015-2016 school year were selected based on the previously noted criteria. Then, based on the evidence presented regarding each of the indicators, 1.3 a "Teachers analyze student progress towards goals." and 1.3 b "Teachers modify instruction in response to data." the team modified the wording in the statement to reflect accurately what was happening at the school. For example, a school recently may have trained staff on a data cycle, and at the end of the school year there may have been only one data cycle completed. In this case, the statement would read "Teachers are beginning to analyze student

progress towards goals." In addition, if not all teachers had adopted a data cycle, the statement would be modified to state, "Some teachers…" or "A few teachers…" For the purposes of this study, the researcher assigned scores to schools based on the SER ratings of the indicators using the following five-point rating scale.

Indicator One (1.3 a): Teachers analyze student progress towards goals.

5- The school requires all teachers to analyze student progress on a routine and regular basis, and evidence supports that all teachers are completing this process.

4- The school requires all teachers to analyze student progress on a routine and regular basis, and evidence supports that most teachers are completing this process.

3 – The school has unclear expectations regarding data analysis, or the expectations are that teachers only review data at the beginning, middle and end of year, and/or teachers are inconsistently completing the process.

2 – The school has unclear expectation regarding data analysis, and only a few teachers are completing an analysis process.

1- The school has no expectations regarding data analysis and there is no evidence that teachers are analyzing data.

Indicator Two (1.3b): Teachers modify instruction in response to data.

5 – All teachers consistently modify instruction in response to data.

4 – Most teachers consistently modify instruction in response to data.

3 – Some teachers modify instruction in response to data.

2 – Few teachers modify instruction in response to data.

1 – No teachers modify instruction in response to data.

Once the two ratings were applied to each of the schools, the researcher summed the

two scores to yield a DATA USE score to compare to PARCC scores. Each school received a rating with a maximum of ten points possible.

A one-way ANOVA then was run to compare schools' mean PARCC scores (which ranged from one to five) across low, medium and high levels of data use as established by rating the two SER key indicators. School's data use ratings were able to range from one to 10, and the corresponding cutoffs for classification were determined by the researcher based on the spread of scores attained. These were low=4-6 (n=4), medium=7-9 (n=7) and high=10 (n=9).

CHAPTER IV

RESULTS

The purpose of this study was to examine how teachers' use of data to inform instruction in an urban school system related to student outcomes on a standardized state assessment. The study focused on consistency and effectiveness of data use across schools.

Twenty schools were selected to participate in this study. Schools selected were required to have students in the tested grades. High school students were not included in the study, as in the state of Maryland high school assessments are content-based. The sample included four schools with grades three through five, 10 schools with grades three through eight, and six schools with grades six through eight.

To rate schools' data use as high, medium, or low, assigned scores from indicator 1.3a and 1.3b were summed. Schools receiving a score of 10 were determined to be high, schools receiving a score between seven and nine were designated as medium, and schools receiving a score of four to six were placed into the low group. (No schools received ratings under four.)

To determine the weighted mean PARCC score, 20 schools' test scores in both math and English Language Arts (ELA) were weighted (for example, a score of five was given five points for each student achieving that score, four points for achieving a four) and then the sum of the weighted scores was divided by the total number of students for whom scores were reported. The resulting measure was the weighted mean PARCC score for each of the 20 schools.

Table 1 offers descriptive statistics regarding the ratings earned on the Standards related to Data Use, Data Use ratings assigned and aggregated PARCC scores for the 20 schools. Table 2 presents the descriptive statistics for the schools broken down by data use levels. As can be seen, the mean PARCC scores did ascend across data use conditions, from 1.6576 to 1.9122.

Table 1

Descriptive Statistics for Sample Data Use and PARCC Scores

	N	Minimum	Maximum	Mean	Std. Deviation
Standard 1.3a	20	2.00	5.00	3.750	1.251
Standard 1.3b	20	2.00	5.00	4.500	.946
DATA USE TOTAL	20	4.00	10.00	8.250	1.997
PARCC mean	20	1.32	2.48	1.837	.327

Table 2

Descriptive Statistics for Data Use and PARCC Scores disaggregated by data use level

LOW DATA USE	N	Minimum	Maximum	Mean	Std. Deviation
Standard 1.3a	4	2.00	3.00	2.250	.500
Standard 1.3b	4	2.00	3.00	2.750	.500
DATA USE TOTAL	4	4.00	6.00	5.000	.817
PARCC mean	4	1.38	2.24	1.658	.405
MEDIUM DATA USE					
Standard 1.3a	7	2.00	4.00	3.000	.577
Standard 1.3b	7	4.00	5.00	4.857	.378
DATA USE TOTAL	7	7.00	9.00	7.857	.690
PARCC mean	7	1.35	2.11	1.843	.252
HIGH DATA USAGE					
Standard 1.3a	9	5.00	5.00	5.000	.000
Standard 1.3b	9	5.00	5.00	5.000	.000
DATA USE TOTAL	9	10.00	10.00	10.000	.000
PARCC mean	9	1.32	2.48	1.912	.351

Hypothesis 1

Table 3, below, shows the results of the one-way ANOVA which compared mean PARCC scores across data usage levels (low, medium and high). The results indicate that the F value (.823) was not statistically significant ($p < .456$), meaning none of the three groups'

PARCC scores were statistically significantly different, therefore, hypothesis 1 was retained. Post-hoc comparisons (Scheffe tests) confirmed these results, as seen in Table 4

Table 3

Results of One-Way ANOVA comparing PARCC Scores Across Schools' Level of Data Use

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	.180	2	.090	.823	.456
Within Groups	1.857	17	.109		
Total	2.037	19			

Table 4

Post Hoc (Scheffe) Comparisons of PARCC Scores Across Schools' Level of Data Use

Group (DATA USE)	**Group**	**Mean Difference**	**Std. Error**	**Sig.**	**95% Confidence Interval**	
					Lower Bound	**Upper Bound**
LOW (4-6)	Medium	-.18503	.20718	.677	-.7403	.3702
	High	-.25462	.19863	.456	-.7870	.2777
MEDUM (7-9)	Low	.18503	.20718	.677	-.3702	.7403
	High	-.06960	.16658	.917	-.5160	.3769

Hypothesis 2

The researcher also tested the significance of correlation between the total Data Use level and the mean PARCC scores for all 20 schools in order to see if there was a significant linear relationship between those variables. The results, which follow in Table 5, ($r = .300$, $p > .199$) indicated that the relationship was not statistically significant. Therefore, hypothesis 2 also was retained, suggesting Data Use levels, as assessed, were not significant predictors of PARCC scores.

Table 5

Correlation between Data Use Total Score and Mean School PARCC Scores

		PARCC Scores
DATA USE TOTAL	**Pearson Correlation**	.300
	Sig. (2-tailed)	.199
	N	20

CHAPTER V

DISCUSSION

The purpose of the study was to determine whether schools' use of student data would correlate with outcomes for students on a standardized test, in this case, the PARCC ELA and Math tests. The review of literature and researcher's experience suggested such a relationship might exist. However, the results of the study did not yield data indicating a significant relationship between levels of student achievement data and use of these data and PARCC scores.

Implication of Results

The results of this study did not indicate that the level of use of data assessed was related to improving student outcomes on standardized tests, although the mean PARCC scores did increase slightly, but not significantly, as ratings of data use levels increased from low to high (see Table 2). While there may not be a significant relationship between data use and outcomes, this lack of a clear relationship between data use and outcomes found may be due to teachers' understanding of what students know and do not know, and making analytical decisions regarding student progress based on informal observations and anecdotal information rather than making decisions based on a data analysis process. Additionally, the results of the study may underestimate the actual relationship between the two constructs due to measurement errors and defining data collection or outcomes in too generally or inaccurately.

Theoretical Consequences

The results of this and related studies are significant as much time and many resources are provided for collecting and using data to inform and evaluate instruction and its outcomes. If this study were conducted on a larger scale and across more varied settings such as urban, rural,

and suburban areas and yielded similar results, it would be important to note that spending time reviewing results of assessments, and making modifications to instruction based on those data may not have a net positive effect. Alternatively, additional studies might help clarify which data are beneficial to collect and how they might best be used to improve educational practice and outcomes. Before determining that data are not useful, several limitations to the generalizability of these findings need to be considered.

Threats to Validity

The most notable limitation in this research study was that the determination of the schools' use of data was based upon a qualitative review process, and the level to which the data were analyzed by the schools studied may not have been fully or accurately captured by the process applied to rate schools' data use. A second limitation is that the study does not account for additional variations influencing the data, such as schools with populations of students having intensive social and emotional needs or high levels of poverty. Such schools may be disproportionally low achieving schools, regardless of their use of data. Therefore, the correlation between data use and outcomes may be obscured as a result of these or other confounding variables. Using more than 20 schools from a larger area may have reduced this threat to the validity of these findings. Also, clearly linking specific practices to particular learning outcomes might have enabled the researcher to make more definitive claims about the usefulness of particular data collection and usage practices for impacting particular skills in students. Disaggregating the sample by age and subject matter also may have yielded clearer findings.

Connections to Previous Studies

As stated by Bertrand and Marsh (2015), teachers may not have effective understanding of

how to use data appropriately. If the use of the data was inappropriate or interpretations were inaccurate, teachers may not have made sound instructional decisions for their students despite higher data use ratings, and this may have obscured the relationship between data use and PARCC scores.

Kallemeyn (2014) also found that teachers do not feel as though they have a full understanding of how to best use student data. During teacher focus groups, teachers may have seemed as though they were utilizing a useful analysis tool; however, their confidence in their ability to use this data cycle and find useful meaning was not measured.

Turner (2014) noted that a data cycle must include at minimum five steps. The SER review does not have prescriptive requirements for a data cycle, and therefore schools rated as having completed data cycles may not have implemented them as comprehensively as recommended by Turner. Clearer operational definition of the construct "data use" would be useful considerations for future studies.

Implications for Future Research

Students and teachers would benefit from more in-depth analyses of the processes that schools implement regarding data collection and use and how they affect outcomes. Because the SER reviews a large amount of information and many data points, this process may not be the most effective or efficient way to determine the degree or quality of use of data by a school in improving student achievement. Additionally, future studies should ensure that schools studied vary in geographical and socioeconomic strata. Such studies also should control for other confounding factors that affect achievement scores and investigate how use of data may identified and described.

Conclusion

Despite the obvious importance of goals and accountability in education, which require data collection and use, this study did not find a significant relationship between the use of data and high stakes accountability achievement results. Limitations in the scope and depth of the study yielded only preliminary findings, and further exploration with a more rigorous scale of determining data use and its impact on specific outcomes is warranted.

School systems would benefit from defining empirically supported, appropriate data cycles, and once those cycles are implemented, ensuring that the processes are used consistently across schools for a sufficient adoption period. Standardized methods and requirements from a school system related to using data for the enhancement of student achievement need to be addressed, along with the implementation and effectiveness of a data cycle on school or student outcomes, although often times inconsistent and difficult to measure.

This study examined a general relationship between data usage and standardized test scores. Future studies are warranted to help educators understand and develop practices for collecting and using data that relate to and improve student outcomes in a cost-effective manner.

References

Bertrand, M., & Marsh, J. A. (2015). Teachers' sensemaking of data and implications for equity. *American Educational Research Journal, 52*(5), 861.

Celio, M. B., & Harvey, J. (2005). *Buried treasure: developing an effective management guide from mountains of educational data.* Retrieved from http://www.wallacefoundation.org/knowledge-center/Pages/Buried-Treasure-Developing-a-Management-Guide.aspx

Choppin, J. (2002). *Data use in practice: Examples from the school level. Center for the Study of Systemic Reform in Milwaukee Public Schools*. Retrieved from http://archive.wceruw.org/mps/AERA2002/data_use_in_practice.htm

Griswold, P. A. (2005). Relating academic data from the elementary grades to state test results in high school: Implications for school improvement through professional development. *Journal of Research in Childhood Education,* 20(2), 65-74.

Lingo, A. S., Barton-Arwood, S. M., & Jolivette, K. (2011). Teachers working together. *Teaching Exceptional Children,* 43(3), 6-13.

Huguet, A., Marsh, J. A., & Farrell, C. C. (2014). Building teachers' data-use capacity: Insights from strong and developing coaches. *Education Policy Analysis Archives,* 22(52).

Kallemeyn, L. (2014). School-level organizational routines for learning: Supporting data use. *Journal of Educational Administration,* 52(4), 529.

Marsh, J., Pane, J., & Hamilton, L. (2006). *Making sense of data-driven decision making in education.* Retrieved from http://www.rand.org/content/dam/rand/pubs/occasional_papers/2006/RAND_OP170.pdf

McCaffrey, D., Koretz, D., Lockwood, J. R. & Hamilton, L. (2003). *Evaluating Value-Added Models for Teacher Accountability.* Retrieved from http://www.rand.org/pubs/monographs/2004/RAND_MG158.pdf

McKown, C., & Weinstein, R. S. (2008). Teacher expectations, classroom context, and the achievement gap. *Journal of School Psychology*, *46*(3), 235-261. doi://dx.doi.org.goucher.idm.oclc.org/10.1016/j.jsp.2007.05.001

Rivkin, S., Hanushek, E., & Kain, J. (2005). Teachers Schools and Academic Achievement. *Econmetrica*, Vol 73, No 2, 417-458. Retrieved from https://econ.ucsb.edu/~jon/Econ230C/HanushekRivkin.pdf

Horn, I., Kane, B. & Wilson, J. (2015). Making sense of student performance data: Data use logics and mathematics teachers' learning opportunities. *American Educational Research Journal*, 52(2), 208. Retrieved from http://search.proquest.com.goucher.idm.oclc.org/docview/1670208369

Turner, S. L. (2014). Creating an assessment-centered classroom: Five essential assessment strategies to support middle grades student learning and achievement. *Middle School Journal*, 45(5), 3-16.