A CAUSAL COMPARATIVE STUDY OF PERFORMANCE PAY FOR TEACHERS IN OHIO: DOES PERFORMANCE PAY AFFECT STUDENT AND TEACHER PERFORMANCE?

Joseph Thomas Hoelzle

A Dissertation

Submitted to the Graduate College of Bowling Green State University in partial fulfillment of the requirements for the degree of

DOCTOR OF EDUCATION

August 2018

Committee:

Paul Johnson, Committee Chair

Pavel Anzenbacher
Graduate Faculty Representative

Rachel Reinhart

Chris Willis

Kyle Ramey
ABSTRACT

Paul Johnson, Committee Chair

The majority of teacher contracts in the State of Ohio are based on the traditional pay scale, rewarding teachers for educational level attained and years of experience. Performance-based pay is an emerging trend with 11 school districts in Ohio identified as having a performance-pay system in their negotiated agreement. This study utilizes Ohio’s similar district methodology to identify the two most similar districts without performance-pay for each performance-pay district. This causal-comparative study compares these two groups by their student (performance index scores, value-added scores) and teacher (evaluation ratings) performance obtained from their local school report card. Data was analyzed utilizing $t$-tests of independent samples and $t$-tests of paired samples, with no significant differences between the two groups. The type of performance pay, as identified by Willis and Ingle (2016), were analyzed with regards to student and teacher performance differences. Performance rates are the type of performance pay that appears to have the most positive effect on student and teacher performance.
This dissertation is dedicated to my family who has been a source of encouragement, inspiration, and motivation throughout the doctoral process.

Thank You, Chelsea, Hannah, and Eden.
ACKNOWLEDGMENTS

My life has changed so much since I began the doctoral process by enrolling in classes in the fall of 2014. Since then, I have changed jobs, moved to a new town, and experienced the joy of fatherhood with the birth of my two daughters. None of this would have been possible without the encouragement and support of my family, Bowling Green Faculty, cohort-mates, and colleagues.

Family

Chelsea, thank you for encouraging me from the start, from agreeing to go on this journey with me, applying to the program, helping to edit papers, and telling me “you can do it” when I doubted myself. You have made things work and picked up the slack on the home front as I have embarked on this process. Hannah and Eden, you have grown so much, from being babies to toddlers, to now almost starting pre-school. Daddy loves you so much, and I am so proud of you. I am looking forward to seeing you continue to grow, and I can’t wait to see the women you become.

Mom and Dad, I would not have been able to achieve the educational levels I have if it wasn’t for the strong foundation you have given me. Words cannot express how grateful I am for the sacrifices you have made and education you provided for me from a young age. I know at times it was a challenge to raise me and I am lucky to have such patient and understanding parents. Rod and Lu, I am also fortunate to have in-laws who help in the raising of my children and support my family in any way that you can.

Dissertation Committee

Completing a study at this level has required much time and energy. Dr. Johnson, thank you for your encouragement and suggestions throughout this process. There have been many
times I have doubted myself, and your gentle encouragement and suggestions have helped me get to where I am today. Dr. Reinhart, thank you for teaching me everything I know about research and methodology. From taking baby stats with you many moons ago to working with you as the methodologist for this research, I have learned so much. Your helpful edits relating to APA format have also made this project better. Dr. Willis, your interest, and research on merit pay profiles influenced me to choose and design this research. Your suggestions have been constructive while I have dug more in-depth into the study. Dr. Ramey, thank you for agreeing to be a member of this committee. I admire the progress and achievements of the Oakwood school district, and your community compares very favorably to similar districts throughout the state. Dr. Anzenbacher, thank you for your participation and insight throughout this process.

Classmates and Colleagues

It would be a failure on my part if I did not mention the many other people who have supported me emotionally and have encouraged me these past few years as I have worked on this doctoral degree. I feel that I have been a part of the most exceptional leadership cohort in BGSU history. I have learned so much from all of you, appreciated the multiple perspectives brought into our classes, and miss the weekly Tuesday night discussions we had. I look forward to seeing our 100% graduation rate of doctors and seeing the many achievements yet to come out of this cohort.

During my doctoral studies, I have also worked full time as a middle / high school principal at Hardin Northern Schools and Tri-County North schools. The people that I have worked with make the lifestyle of an educator so enjoyable. Thank you to all the students I have worked with and seen mature. Thank you to the staff – teachers, guidance counselors, maintenance, secretaries, bus drivers, etc. Hopefully, you feel that I have supported and
positively influenced you, I know you have done for me. Thank you to my fellow school administrators, especially Doug, Sally, Jeff, Brett, Bill, Kristen, Joe, and Beth. I have learned much from your leadership, and am grateful for the many discussions we have had.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Problem Statement</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Purpose of Study</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Research Questions</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Theoretical Perspective</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Significance of the Study</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Limitations</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Delimitations</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>Assumptions</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>Definitions</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>Organization of Chapters</td>
<td>13</td>
</tr>
<tr>
<td>II</td>
<td>REVIEW OF THE LITERATURE</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>Theoretical Leadership</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>Goal Setting and Contingency Theories</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>Social Dilemma and Participative Management Theories</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>Expectancy Theory</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>Other Motivation Research</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>Federal Reforms</td>
<td>21</td>
</tr>
<tr>
<td></td>
<td>A Nation at Risk</td>
<td>21</td>
</tr>
<tr>
<td></td>
<td>Race to the Top</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>Ohio Reforms Leading to Increased Accountability</td>
<td>25</td>
</tr>
</tbody>
</table>
Data Analysis ............................................................................................................ 52

Variables ........................................................................................................ 52

Descriptive Statistics.......................................................................................... 54

Inferential Statistics .......................................................................................... 54

Limitations ............................................................................................................. 55

CHAPTER IV: ANALYSIS AND PRESENTATION OF DATA........................................ 57

Results ............................................................................................................. 57

Research Question 1 ...................................................................................... 57

Research Question 2 ...................................................................................... 59

Student Performance ..................................................................................... 60

Teacher Performance ..................................................................................... 63

Summary .......................................................................................................... 66

CHAPTER V: DISCUSSION AND CONCLUSION..................................................... 67

Overview of Study .......................................................................................... 67

Summary and Discussion of Results ............................................................... 69

Conclusions ...................................................................................................... 75

Recommendations ........................................................................................... 78

REFERENCES.................................................................................................. 86

APPENDIX A: IRB NO APPROVAL NEEDED LETTER................................. 97
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Graph of performance index scores from 2016-17 school year by district with comparison group</td>
</tr>
<tr>
<td></td>
<td>.......................................................... 61</td>
</tr>
<tr>
<td>2</td>
<td>Graph of value-added scores from 2016-17 school year by district with comparison group</td>
</tr>
<tr>
<td></td>
<td>.......................................................... 62</td>
</tr>
<tr>
<td>3</td>
<td>Graph of percent of teachers rated proficient or accomplished from 2016-17 school year by district with comparison group</td>
</tr>
<tr>
<td></td>
<td>.......................................................... 64</td>
</tr>
<tr>
<td>4</td>
<td>Graph of percent of teachers rated accomplished from 2016-17 school year by district with comparison group</td>
</tr>
<tr>
<td></td>
<td>.......................................................... 65</td>
</tr>
</tbody>
</table>
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Calculation of Performance Index Scores</td>
<td>30</td>
</tr>
<tr>
<td>2</td>
<td>Springer’s (2009) Types of Compensation Reforms</td>
<td>41</td>
</tr>
<tr>
<td>3</td>
<td>Performance Pay Districts and Participating Comparison Districts</td>
<td>47</td>
</tr>
<tr>
<td>4</td>
<td>Data Analysis Procedures by Research Question</td>
<td>52</td>
</tr>
<tr>
<td>5</td>
<td>$t$-Tests of Independent Samples Comparing Performance-Based Districts to Control Districts</td>
<td>58</td>
</tr>
<tr>
<td>6</td>
<td>$t$-Tests of Paired Samples Comparing Performance-Based Districts to Control Districts</td>
<td>58</td>
</tr>
<tr>
<td>7</td>
<td>Performance-Pay Districts by Type</td>
<td>59</td>
</tr>
<tr>
<td>8</td>
<td>Performance-Pay District Student Performance Comparison to Control Districts</td>
<td>60</td>
</tr>
<tr>
<td>9</td>
<td>Performance-Pay District Teacher Performance Comparison to Control Districts</td>
<td>63</td>
</tr>
</tbody>
</table>
CHAPTER I: INTRODUCTION

In 1983, A Nation at Risk was published by the National Commission on Excellence in Education, and ever since, reformers and policymakers have put forth educational reforms in an attempt to improve the American Education System. Many educational reform efforts have been attempted to improve America's schools, including standardization of content, new teacher evaluations systems, standardized testing, and performance pay for teachers. Moreover, results on international tests such as the Trends in International Mathematics and Science Study (TIMMS) and Program for International Student Assessment (PISA) have shown that American students are falling behind their international peers (Zhao, 2012).

An essential responsibility of those leading educational organizations is to foster the continuous improvement of those organizations. In order to quantify continuous improvement efforts, many educational reforms have advocated for the use of data to guide decision-making (Boudett, City & Murnane, 2013; Dufour & Marzano, 2011; Goldring & Berends, 2009). In Ohio, data has been collected using a new teacher evaluation system, staff and student attendance, student discipline, and student test results. Data from state-wide assessments have been aggregated to show student performance and average student growth in school districts.

With data from test scores now available, programs and reform efforts can be evaluated to see if they affect student performance. Merit or performance pay as a reform effort is becoming more prevalent (Willis & Ingle, 2016). As more school districts implement performance pay provisions in their teacher contracts, the provisions need to be evaluated to show the effectiveness of these reform efforts.

While some educational reformers have advocated for the implementation of performance pay policies (Hanushek & Lindseth, 2009), the research on the effectiveness of
performance pay to improve performance is mixed (Fullan, 2011). As such, some scholars have studied the impact of performance pay on student performance and found no relationship (Razo, 2014), and others have found a positive impact on student performance in other states (Pham, Nguyen, & Springer, 2017). While districts in other states have previously tried various compensation reforms, performance pay in Ohio is a relatively new phenomenon, with only 16 school districts identified as having some performance pay provision in their contract (Willis & Ingle, 2016). This study used available state-wide data to evaluate whether performance pay provisions in Ohio teacher contracts affected student and teacher performance.

In Chapter 1, the problem statement, purpose, research questions, and theoretical perspective are presented. This chapter also includes a statement of significance, delimitations and limitations, assumptions, definitions, and a brief description of the organization of the rest of the study.

**Problem Statement**

The traditional teacher pay scale rewards teachers based on years of experience and the education level they have attained. Historically, this has been the most common way to compensate teachers and was created to fight inequities in pay between male and female teachers (Guthrie, Springer, Rolle, and Houck, 2007). Ninety-five percent of public school districts in the United States currently utilize the traditional pay scale (Pham, et al, 2017). However, some districts in Ohio have begun to implement alternative compensation systems based on performance (Willis & Ingle, 2016). Fullan (2011) researched the topic of performance pay for teachers and pointed out that the research is mixed on whether performance pay for teachers is effective at increasing achievement for students, with some researchers advocating for performance pay, and other researchers concluding that performance pay does more harm than
good for schools. Compensation of teachers affects more than just the individual teacher because the performance of that teacher can also affect the performance of other teachers in the building and student achievement. If a teacher’s motivation increases because of this new form of compensation, the implementation of such a system may have positive impacts on student achievement.

In previous studies on performance pay, some researchers focused on its impact on student test scores. Researchers have also focused on performance pay’s effect on school funding, and others researchers have concentrated on teacher attitudes about switching to such a system. Razo (2014) compared student test results on Arizona’s standardized tests to teacher performance pay and teacher evaluation ratings to performance pay but did not find strong correlations between grade levels and subject areas. Hanushek and Lindseth (2009) advocated for implementing a performance-pay system for teachers because they believe that such a system will help solve school funding problems nationwide. Teacher attitudes about switching to alternative pay systems have also been studied and have shown that teacher perceptions are mixed about such systems (Goldhaber, DeArmond & Deburgomaster, 2011; Lundstrom, 2012; Tenhiala & Lount, 2013; Viscardi 2014).

While researchers have studied the impact of performance pay in other locations and professions, performance pay for teachers in Ohio is a relatively new phenomenon (Willis & Ingle, 2016). Ohio implemented the Ohio Teacher Evaluation System (OTES) starting in 2012. This system creates a consistent, state-wide model to evaluate teachers, created from research-based teacher performance criteria. Along with OTES, the state has also implemented many other changes since 2010 including the following: switching to the common core state standards (CCSS) for English language arts and mathematics, converting to Ohio’s new revised standards
for science and social studies, changing to Partnership for Assessment of Readiness for College and Careers (PARCC) assessments during the 2014-15 school year, changing to American Institute of Research (AIR) assessments during the 2015-16 school year, and implementing new state report cards during the 2013-14 school year to compare districts statewide. The new state-issued report cards offer a wealth of data by which to compare districts that have implemented performance-pay systems with other districts that have not.

**Purpose of Study**

The purpose of this study was to investigate performance-based pay systems for teachers in Ohio and to determine if these systems had a significant impact on student achievement and teacher performance during the 2016–17 school year. This study also attempts to discover if different types of performance-based pay systems have are more or less successful in impacting student achievement and teacher performance.

The intent was to discover if districts identified as having performance-based pay structures have significantly different performance index scores, value-added ratings, and teacher evaluation ratings than similar school districts who have traditional teacher pay structures. Data was collected on these school districts from the Ohio Department of Educations' interactive local report card. The data was then transferred to SPSS with t-scores and p-values being reported.

**Research Questions**

1. Do Ohio school districts that have implemented a performance-pay system differ significantly from Ohio school districts that have not implemented a performance-based pay system with regard to their performance index scores, value-added ratings, the percent of teachers rated proficient or higher, and the percent of teachers rated
accomplished when using data from the Ohio School district report cards from the 2016-2017 school year?

2. Does the type of performance-pay system in Ohio school districts have different results with regard to a district's performance index score, value-added ratings, the percent of teachers rated proficient or higher, and the percent of teachers rated accomplished when using data from the Ohio School district report cards from the 2016-2017 school year?

**Theoretical Perspective**

The theoretical base that this study is grounded in is motivation theory. The overarching theory is that if a teacher's motivation improves, then the performance of this teacher and his or her students should also improve. Odden and Kelley (2002) identify five theories that combined can help explain worker motivation and the role incentive programs can play in motivating workers. These five theories include contingency theory, goal-setting theory, social dilemma theory, participative management theory, and expectancy theory.

Contingency theory explains that a compensation program will be useful if the program fits with the goals and vision of the organization. The more closely the incentive plan is aligned with the goals, mission, and vision of the organization; the more efficient the plan will be in motivating employees and increasing productivity (Lawler, 1990; Welbourne & Mejia, 1995). Locke (1968) proposed a goal-setting theory where employees are motivated by goals when these goals are specific, challenging, and accepted by employees as worthwhile and achievable. Motivation will be most significant when goals and rewards are combined (Heneman, 1992). The social dilemma theory postulates that individuals may become free-riders if group performance awards are given. However, Richards, Fishbein, and Melville (1993) asserted that the structure of a work team could influence worker behavior. The free-rider effect can be
minimized when group members communicate their expectations about fair individual contributions to the group effort (Runge, 1984). The participative management theory suggests that when individual workers contribute to organizational decision making, the individual will be more motivated and committed to the organization (Odden & Kelley, 2002). The most commonly cited theory with relation to performance or merit pay is the expectancy theory. Expectancy theory postulates that an employee will respond favorably to a compensation plan if the employee expects to be able to accomplish the goal being rewarded, perceives a connection between his or her efforts and the reward being offered, and values the reward being offered.

All of these theories can be applied in an educational setting and are discussed further in chapter two. If teachers see value in the goals of the school, performance may improve through linking these goals to a compensation or incentive program. The goals may be created in collaboration with teachers, and if the teachers work in groups to accomplish these goals, the overarching theory is that the compensation program will be successful in motivating employees and increasing district performance. This preceding section has explained how the topic of motivation serves as the theoretical framework of this study. The next section of this chapter will be the significance of this study.

**Significance of the Study**

Willis and Ingle (2016) discussed that performance pay systems in the Ohio school districts could be a growing trend due to the state-wide implementation of a new teacher evaluation system and the federal Race to the Top (RttT) grant money. As school districts contemplate implementing alternative pay structures, they may inevitably face resistance from stakeholders such as veteran teachers who benefit from the traditional pay scale.
A principal justification for implementing a performance pay scale, or any change in education, is that the change may increase the academic achievement of students. This study adds to the body of literature that already demonstrates the effects, or lack thereof, that performance pay for teachers can have on districts' academic achievement or growth. If this study shows a positive impact on student teacher performance and growth, then there will be further justification for implementing performance pay for teachers in other school districts. If this study demonstrates no effect on teacher performance or student performance and growth, then there will be further justification for not implementing performance pay for teachers in other school districts.

While researchers in other states have investigated the impact of performance pay on academic achievement (Pham, et. al., 2017; Razo, 2014), this study focuses on performance pay for teachers in Ohio school districts. With the recent identification of Ohio districts that have performance pay provisions in teacher contracts (Willis & Ingle, 2016), it is now time to determine if these performance pay provisions have any impact on district-wide student achievement data. School leaders in Ohio districts are very familiar with performance index scores, value-added ratings, and OTES ratings. If a change in the way teachers are remunerated will have a positive impact on available district data, the Ohio school leaders may be very willing to work to implement new pay structures in their school districts.

Additionally, this study will also be beneficial for school districts that have implemented a performance pay system while they are evaluating the effectiveness of their system. If this study displays that a significant positive difference exists between performance pay districts and their peers, then the school districts should be more likely to keep their performance pay provisions in their contract. If a significant difference does not exist, then a district may want to
focus on other initiatives instead of performance pay to work on improving the student performance and growth in their district.

The effect performance pay systems have on the academic achievement of students is an important question that needs answering for policymakers, the board of education members, superintendents, and educators, in general. If policymakers such as state representatives, senators, the governor, and leaders at the Ohio Department of Education are provided with evidence that districts with performance pay have higher levels of student achievement, they may be more likely to pursue policies encouraging more schools to implement performance pay. Boards of Education may see that other Ohio districts have had favorable results from implementing performance pay for their teachers and may want to negotiate similar provisions into their local contracts. Superintendents and teacher unions may also be interested in the results of this study as they look for ways to improve their districts.

The results of this study could further justify to educational leaders the merit of implementing performance-based pay provisions in teacher contracts throughout the state of Ohio, and even nationwide. Results of this study give clues as to which type of performance-based compensation systems may be most successful in impacting student and teacher performance. Alternatively, the results of this study may show a lack of effect that performance-based pay provisions have on student achievement, encouraging leaders to focus their efforts on other areas to improve the school district.

**Limitations**

This study is only a snapshot of results of the 2016–17 school year, which is the most recent data available from the Ohio Department of Education’s interactive local report card. With the relatively new phenomenon of merit pay in Ohio, it is not logical to conduct a longitudinal
study of student performance at this time. The 2015–16 school year was the third consecutive year that Ohio had a different vendor and format of their standardized test, and with these changes, a longitudinal study is not possible because the standardized results would not be comparing results from the same assessments. With this study only focusing on results from one school year, the research could have different results in different school years.

A second limitation of this study is that causal-comparative research does not infer causality; it is only suggestive. This is a threat to external validity because the characteristics of districts in this study cannot be generalized to other districts that do not have the same characteristics or different time periods. This study is not be able to state whether implementing a performance pay system has caused a district to have better student performance. The study may be able to suggest that such a link exists.

This study is not a qualitative study. Previous research has looked at teacher perceptions of performance-based pay (Goldhaber, DeArmond & Deburgomaster, 2011; Lundstrom, 2012; Tenhiala & Lount, 2013; Viscardi, 2014). This study will be quantitative, using data to analyze the performance-based pay system's impact on teacher and student performance, and therefore, the thoughts, opinions, and viewpoints of people affected by performance-based pay provisions are not analyzed.

Another limitation discussed here for this study is that it does not find significant, statistical differences between the different types of performance-based pay in Ohio. Each district has a teacher contract individualized to their locality, and no teacher contract is the same. Willis and Ingle (2016) adapted Springer’s (2009) typologies to classify 16 Ohio districts into different categories. These 16 were screened down to 11 after contract confirmation. For the first research question, all 11 districts will be in the category of being identified as having a
performance-based pay provision in their contract in answering the first research question. The second research question attempts to classify these districts by typology, but the number of districts is so small that statistical significance will not be analyzed. As more Ohio school districts have performance-based pay provisions added to their teacher contracts, the different categories may be able to be analyzed again individually, but currently, the sample size would be too small for finding statistical significance.

**Delimitations**

By using the methodology of comparing similar districts, this study has attempted to control for extraneous variables. High wealth, low poverty districts tend to achieve at higher levels than high poverty districts on performance index scores. The study would be severely flawed if it compared the performance index scores from a suburban district like Oakwood to an urban district such as Dayton. By comparing similar districts, the researcher hopes that the differences in the makeup of districts do not have a significant effect on the differences in the performance of these districts. This threat to internal validity would be an example of selection because the districts may have characteristics that predispose them to particular outcomes.

While the performance index scores across the state correlate with the socio-economic make-up of their districts (Caldas & Bankston, 1997; Pettigrew, 2009), value-added scores do not. Value-added scores measure the average academic growth students make in a school year compared to the academic growth of other students. There is no correlation between the value-added scores and socio-economic ratings of a district (Sanders & Horn, 1994). This is another threat to internal validity. Value-added scores are designed to measure the average growth students make in one school year, with negative scores indicating that students did not grow as much as the expected average of their peers. There are high wealth districts with low value-
added scores, and there are low wealth districts with high value-added scores. Conversely, there are also low socio-economic status (SES) districts with low value-added scores and high SES districts with high value-added scores. By comparing both the value-added scores and performance index scores of a district, this study uses two different ways to analyze student test data to conclude the effectiveness of a school district.

Assumptions

At least three underlying assumptions exist for this study. First, 16 districts have been identified by Willis and Ingle in their 2016 publication that has some form of the performance pay compensation system for their teachers. It is possible that these districts have renegotiated their contract since the Willis and Ingle publication. For this dissertation, 5 districts from the Willis and Ingle study were identified as no longer having performance-based pay provisions in their contracts. This identification occurred after the research project was approved. The State Employment Relations Board (SERB) website and individual district websites were accessed to confirm that each district still has a performance pay provision in their contract.

Willis and Ingle (2016) initially accessed contracts from the Ohio State Employment Relations Board's (SERB) website to identify performance pay districts, then identified additional districts by sending an e-mail to superintendents throughout Ohio and asking them to self-identify. Therefore, the Willis and Ingle study is partially reliant upon the SERB Ohio keeping its contract database updated. It is possible that a school district has implemented some form of performance pay compensation system and not had the contract updated correctly in the SERB Ohio database. The second assumption for this study is that the identified similar districts have not implemented a performance pay compensation system for their teachers. The State
Employment Relations Board (SERB) website and individual district website was accessed to confirm that each district still does not have a performance pay provision in their contract.

Ohio utilizes an elaborate system to identify similar school districts known as its similar district methodology. This methodology will be explained in chapter two. This study is reliant upon the similar district methodology to identify districts that are closely related to the performance pay districts. The third assumption is that the designated similar districts are as identical as possible to the performance pay districts, including socio-economic make-up, except for the variable of whether or not they use a performance-based compensation system.

**Definitions**

**Performance-Pay** - Performance-based pay refers to a pay structure where teachers are compensated in some way based upon a pre-set criteria rewarding performance and are not based on the traditional pay scale that rewards years of experience and educational units attained.

**Similar District Methodology** - A method the Ohio Department of Education uses to identify up to 20 of the most similar districts to a specific district is based on criteria such as number of students, percent of students in poverty, socioeconomic status of district, race/ethnicity percentage of students, rural/urban continuum, and the tax capacity of a district.

**Performance Index Score** - The performance index score of a district is obtained from the Ohio Department of Education's interactive local report card and measures the level of performance a district achieved on the state-wide assessments. The maximum score a district could achieve would be 120 if every student in the district achieved the highest score on their assessment. Performance index scores have been found to be correlated to the socio-economic make-up of the community the school serves.
**Value Added Ratings** - Value-added ratings are obtained from the Ohio Department of Education's interactive local report card and measure the growth that students make from one year to the next. A value-added score of 0 for a district would signify that the students in that district on average made one year of expected growth in one year. Positive value-added scores signify that a district's students made more than a year's worth of growth in a year, and negative value-added ratings mean that the district's students on average made less than a year's worth of growth in a school year. Value-added scores are not correlated to the socio-economic makeup of the community a school serves.

**Ohio Teacher Evaluation System (OTES)** - All public school teachers in Ohio are evaluated under OTES. A teacher receives a rating of: ineffective, developing, skilled, or accomplished, and these ratings are submitted to the Ohio Department of Education using the electronic teacher principal evaluation system (ETPES). The overall score is a combination of ratings from the principal evaluation and student growth measures. The Ohio Department of Education provides data showing how many teachers in each district are rated at each level.

**Organization of Chapters**

This study is divided into four remaining chapters. Chapter 2 provides a review of the literature, starting with the theoretical framework, then reforms at the national and state levels, and finally taking a closer look at previous research and examples of performance-based pay. Chapter 3 describes the research design and methodology used in this quantitative study. Chapter 4 provides an analysis of the results of this study. In Chapter 5, conclusions are made while connecting the results of this study to previous studies and suggestions for future research will be offered.
CHAPTER II: REVIEW OF THE LITERATURE

This chapter provides a review of the literature previously published relating to this study. The review begins by examining the theoretical framework and relating motivational theory to teacher compensation. After discussing these different theories, the review will start with general educational reforms before focusing in on the reforms of merit pay. Various reforms efforts at the federal and state levels (focusing on Ohio) are discussed. Finally, research on merit pay is presented, focusing on: the history of paying teachers, forms of merit pay, reasons for having merit pay, teacher perceptions of merit pay, arguments against merit pay, merit pay typologies, and the Ingle and Willis (2016) study.

Theoretical Framework

As stated in Chapter 1, the theory behind this study is teacher motivation. If teacher motivation increases from a performance pay program, then the performance of teachers and academic performance of a district's students should also improve. Odden and Kelley (2002) identify five motivational theories that impact employee motivation including goal setting theory, contingency theory, social dilemma theory, participative management theory, and expectancy theory. After discussing these theories, other research on employee and teacher motivation is covered. A central argument for a school district to implement performance pay is the belief that paying a teacher based on performance criteria will motivate the teacher to improve his or her performance.

Goal Setting and Contingency Theories

Goal-setting and tying rewards to the achievement of those goals have been known to motivate people. According to the goal-setting theory (Locke, 1968), goals motivate employee behavior if those goals are specific, challenging, and accepted as worthwhile and achievable.
Ohio teachers set S.M.A.R.T. goals (that are specific, measurable, attainable, realistic, and timely) every school year to be compliant with the state-wide teacher evaluation system. Setting goals has been demonstrated to motivate employees to higher levels of performance (Mento, Steel, & Karen, 1987; Mohrman & Lawler, 1996; Rowan, 1996). By embedding pay with the attainment of goals, the motivation is even greater than just having a goal itself without compensation for goal attainment (Heneman, 1992).

The contingency theory is another theory that is related to employee compensation. This theory postulates that compensations systems work well when tied to the characteristics and strategies employed by the human resources department. These compensation systems work best when compensation and rewards are tied to the organization's mission, vision, and objectives, resulting in high employee motivation and productivity. (Lawler, 1990; Welbourne & Meija, 1995). For example, if a school has an objective of increasing parental contacts from the classroom teacher, the district could design a system to track and reward teachers who increased the quantity of parental contacts. Cummings (1994) emphasized the importance of organizational support when incentives are tied to goals. For example, a school district could provide professional development before implementing a form of contingency pay.

Both contingency and goal setting theories relate to this study because school districts can set goals within their organization and tie compensation to the attainment of these goals. School districts may also design their teacher compensation programs around group attainment of goals. The culture of a school and the norms of a teacher team can also affect a teacher's motivation.
Social Dilemma and Participative Management Theories

Some school districts provide an award to all of the teachers or a team of teachers if a certain objective is met. For example, a district may have profit sharing amongst the teachers if the district has positive open-enrollment, or a bonus may be awarded to all teachers if the district reaches a specific performance index score. Social dilemma theory proposes that with group awards, employees may become free riders and shirk their responsibilities if they will still receive the same award as their co-workers who put in extra effort (Odden & Kelley, 2002). However, this free-rider is more theoretical than actual, with work teams having an enormous influence on individual behavior (Richards, Fishbein and Melville, 1993). When the group communicates the expectations for individual behavior, the free-rider effect is minimized (Runge, 1984). When a school district provides teacher teams with an opportunity to earn bonuses and awards, some teachers may earn the awards off of the labor of their teammates, but more likely, the group peer pressure will elevate everyone's performance.

Another theory related to employee motivation and compensation is the participative-management theory. When employees are involved in decision-making, the stakes are higher, and the employee has a greater stake in making the decision successful (Vroom, 1976). When teachers have a voice in the creation of a performance pay system, those individual teachers will be more likely to help the implementation of that system. Participative management works when group norms support employee participation in the organization's decision-making process (Lawler, 1973; Welbourne & Mejia, 1995).

Both social dilemma and participative management theories suggest the importance of groups in employee motivation. School districts may design their performance pay structures
with group successes in mind as Odden and Kelley (2002) suggest. The next employee motivation theory discussed in this literature review will be expectancy theory.

**Expectancy theory**

While many have discussed motivation in their research (Maslow, 1970; Lewin, 1946; McClelland, 1961; Herzberg, 1960), the theory most closely related to this study because of the relationship between motivation, compensation, and performance is Vroom's (1976) expectancy theory. As opposed to intrinsic motivation, where one is motivated because of their interests or passions, expectancy theory falls more along the lines of extrinsic motivation, where one is motivated by external rewards or fear of punishments.

Expectancy theory is a popular theory used to explain how the design of a compensation plan could affect employee motivation. This theory is based on the belief that employees will respond favorably to a new compensation system when three criteria are met (Cumming, 1994; Lawler, 1990; Welbourne & Mejia, 1995). The first criterion is known as expectancy and is related to the employee's effort. Individuals must believe that they have the skills, training, and aptitude to accomplish the task rewarded to put effort towards achieving the goal or task. In the profession of education, teachers must believe that the job they are rewarded for has clear expectations that have been communicated and are attainable.

The second criterion for employees to respond favorably to a new compensation system is known as line of sight, instrumentality, or performance. Employees must see a clear connection between their efforts and the rewards received. If employees do not understand that their efforts are rewarded, they may lose faith in the reward system. Teachers may be motivated by a performance pay system if the connection between their performance and the rewards that their satisfactory performance brings is transparent.
The third criterion for employees to respond favorably to a new compensation system is the value that an employee sees in a reward. If an employee does not see value in the reward offered, or they do not understand enough value based off of the perceived effort to attain the award, the employee will not be enticed to work towards achieving the reward. If a teacher does not believe that the performance pay will be a significant increase in salary, the teacher may not put forth additional effort. However, the teacher may be motivated by performance pay if the pay is substantial and attainable.

According to Vroom's (1976) expectancy theory, changing a contract to pay teachers based on their performance should have a positive impact on the performance of teachers. If the performance of teachers in a school district improves, so should student test scores and the performance of a school district as a whole. If changing to a performance pay system for teachers helps the school district improve its performance, then it is the responsibility of a school leader to help foster this change.

Other Motivation Research

Proponents of a performance pay system often advocate for such a system because they claim that the opportunity to earn a higher salary is motivating and those hard-working teachers should be rewarded for their efforts. However, teachers say material rewards do not influence their behavior; they are not in it for the money (Goldhaber, et. al., 2011). A better approach would be to raise the base pay of teachers (Pink, 2009). Merit pay recommendations are consistent with bureaucratic assumptions- assumptions that include the perception that for the most part, employees are more likely to be motivated by rewards extrinsic to their work than rewards inherent in their work. They are not, however, consistent with what it takes to gain from
knowledge workers, the commitments to shared standards of excellence that are required to produce high-quality performance in all aspects of organizational life (Schlechty, 2009).

Daniel Pink (2009) has reported on what motivates people and claims that autonomy, mastery, and purpose are three areas that facilitate intrinsic motivation. Autonomy refers to people having independence and not being micro-managed. Mastery is the process of attempting to improve one's practice continuously. Purpose gives people a reason why their work is essential. According to Pink (2009), it is not how much money we make that ultimately makes us happy; it is whether our work fulfills us. Most teachers do not have an issue with finding purpose in their work; however, changing the structure of schools to facilitate better autonomy and mastery may have a more significant impact on employee motivation than changing the pay structure to a performance-based compensation system.

Understanding the research of Abraham Maslow can also enhance one’s ability to motivate employees. Maslow is credited with creating his hierarchy of needs. Included in this hierarchy are physiological needs, safety needs, love and belongingness needs, esteem needs, and self-actualization. Lower level needs must be satisfied to meet higher level needs. In relating Maslow's research to motivating teachers, teachers can achieve self-actualization and perform their greatest when these needs are met. The culture of the workplace should make teachers feel valued and loved to fulfill their love, belongingness, and esteem needs. Compensation would most relate to the lowest level need, physiological needs, because teachers need to be compensated enough to provide for the physiological needs of themselves and their families. Frederick Herzberg (1960) takes the research of Maslow (1970) to another level with his motivation-hygiene theory. Maslow’s theory is helpful in identifying needs or motives, and Herzberg’s theory provides us with goals and incentives that tend to satisfy those needs (Hersey,
Blanchard, & Johnson, 2001). Money tends to fulfill needs at the physiological level. Once these basic physiological needs are met, schools will still need to help motivate their employees by finding ways to meet their esteem and self-actualization needs. In reviewing general motivational theory, Hersey et al. (2001) conclude that people seek security, social systems, and personal growth.

Blackburn and Williamson (2016) conclude that external pressure can help get things done, but the success in schools come when people see value in what they are doing and believe they will be successful. When teachers experience hands-on-learning, are involved in the planning of professional development and feel valued by their supervisors, the school is more likely to improve its culture.

Education research on performance pay and whether it affects teacher motivation is still in its infancy. Hulleman and Barron (2010) are unclear "whether performance pay will have an energizing effect on teacher motivation and quality or a negative effect on the public-service ethos of education" (p. 30). Gagne and Forrest (2008) assert that monetary incentives may increase quantity in manufacturing fields but may not work in a classroom, which is a more complex task requiring quality work. Expected rewards, based on performing a task at a specific level, may undermine intrinsic motivation for an employee (Deci, Koestner, and Ryan, 1999). While private sector employees are more likely to report being motivated by financial rewards (Rothstein, 2009), teachers may highlight the intrinsic motivations in teaching to help them stay motivated since their salaries are already so low.

The overarching theory of this study is that if a teacher becomes more motivated, then the performance of that teacher and his or her students should also improve. With traditional teacher salary schedules out of favor with many policymakers, suggestions have been made to change to
more of a performance-based pay system. These same policy makers have been attempting to improve our educational system through the passage of law, research, and grant programs, often suggesting movement towards merit or performance-based pay systems for teachers.

**Federal Reforms**

One of the first ways the federal government got involved with educational reforms was through the passage of the Elementary and Secondary Education Act (ESEA) of 1965. This law has attached funding to specific objectives the federal government has for educational reforms and has had notable revisions with the passage of No Child Left Behind (2002), and the Every Student Succeeds Act (2015). These revisions have increased accountability and also led to more stringent requirements for teacher licensing and student assessments. The most notable reform efforts encouraged by the federal government include the publishing of A Nation at Risk and the passage of the Race to the Top stimulus package.

**A Nation at Risk**

A Nation at Risk (1983) was published by the National Commission on Excellence in Education by the request of the United State Secretary of Education, T.H. Bell. This report on the quality of education in America assessed the quality of teaching in America during this period, compared America's schools to other advanced nations, and defined problems that must be overcome for America's schools to pursue excellence. After assessing the state of America's schools in the 1970s and early 1980's, the report gives five recommendations for America's schools to improve, including improving content taught in America's high schools, developing standards and increasing expectations, increasing the amount of time students are in the classroom, making teaching a more respected and rewarding profession, and improving school leadership and fiscal support.
With the publication of a Nation at Risk, critiques of the report soon followed. Ohanian (1985) points out that none of the committee members of the national committee on excellence were teachers; if one wants to improve education in America, then teachers should be consulted on when deciding which reforms to emphasize. Ohanian (1985) also suggests that politicians should pay their own salaries based off of performance first, before experimenting with merit pay for teachers.

The second recommendation, which focused on making teaching a more rewarding and respected profession, states:

"Salaries for the teaching profession should be increased and made professionally competitive, market sensitive, and performance-based. Salary promotion, tenure, and retention decisions should be tied to an effective evaluation system that included peer review so that superior teachers can be rewarded, average ones encouraged, and poor ones either improved or terminated" (p. 38).

As a result of a Nation at Risk (1983), some school districts began looking at ways to implement a performance-based pay system for their teachers.

The United States Secretary of Education at the time a Nation at Risk was published was Terrel H. Bell. According to Bell (1993), most of the top-down reform efforts of the 1980s were ineffective at improving student achievement. In Bell's (1993) reflections a decade after a Nation at Risk, he suggests eight areas to be addressed to realize America's educational goals: technology, staffing, a more significant national role, increasing enrollment at private schools, parental involvement, smaller schools, a focus on leadership, and utilization of brain research. Again, with a focus on staffing, Bell (1993) states:
"The tired and worn-out school staffing and teacher personnel practices...must be scrapped. The single salary schedule cries out for enhancements that will compensate our great teachers on a level commensurate with their worth...our current practices hamstring our efforts to build a truly great teaching professional. We would do well to take a page out of the higher education book" (p. 595).

Twenty-five years after A Nation at Risk was published, A Nation Accountable was published by the United States Department of Education (2008). This document reflects on the reforms suggested decades earlier and concludes that America is even more at risk because of increasing global competition. Data is provided showing in increase in graduation and literacy rates over the previous twenty-five years. Teacher quality is still an issue with the fair allocation of resources a critical problem to be addressed. The report suggests that programs such as the teacher incentive fund may be an effective way to address teacher quality, along with national board certification of teachers and the increase in standards and accountability.

Many other reform efforts suggested in this report have also been attempted. Not only have individual districts attempted these reforms, but state and federal education laws, grants, and budgets have also suggested and mandated reforms suggested in A Nation at Risk (1983). A federal education stimulus package that also encourages performance pay is Race to the Top (2009).

**Race to the Top**

In the late 2000s, America suffered its greatest economic downturn since the great depression of the 1930s. As a result of this recession, Congress passed the American Recovery and Reinvestment Act of 2009 in an attempt to stimulate the economy. Nearly $100 billion was spent to restore education budgets, reward innovation, and advance reforms (United States
At $4.35 billion, the Race to the Top (RttT) – a competitive grant program to spur K-12 improvements – constituted just less than 5 percent of the total stimulus package of ARRA (United States Department of Education [USDOE], 2009).

The RttT grant awarded funds to states in three rounds of applications. In each round, applications were evaluated off of 6 criteria, including state success factors, standards and assessments, data systems to support instruction, great teachers and leaders, turning around the lowest-achieving schools, and general selection criteria (USDOE, 2009).

The application was evaluated with states able to receive up to 485 points divided amongst the different criteria. A total of 138 points could be earned by the state focusing on great teachers and leaders. Of these 138 points, 58 of these could be earned through improving teacher and principal effectiveness. In order to earn these points, states were encouraged to "use evaluations, at a minimum, to inform decisions regarding...compensating, promoting, and retaining teachers and principals, including by providing opportunities for highly effective teachers and principals to obtain additional compensation and be given additional responsibilities"(USDOE, 2009, p. 9). In response the RttT grant, Austin, Texas, and Jefferson County, Colorado developed performance-based compensation systems. (The White House Office of the Press Secretary, 2009). Nine states, including Ohio, developed state-led pay for performance proposals, while 18 states had pilot programs testing pay for performance in individual districts, and nine states also granted local districts the authority to implement their performance pay initiatives (Rose, 2010). The point system used to evaluate RttT applications encouraged performance based pays schemes for teachers, along with many other innovations in education.
Although federal law and grants from the federal government have significantly impacted the current state of Education in American, often state and local reform efforts and initiatives have a greater impact on individual schools. The next section of this chapter will focus on reforms that have occurred in Ohio in the past few years. Many of these reforms were part of Ohio's application for Race to the Top funds and came to fruition as a result of Ohio being rewarded 400 million dollars from this grant (Starzyk & McNeil, 2011).

Ohio Reforms Leading to Increased Accountability

The ancient Greek philosopher Heraclitus is quoted to have said: "the only thing constant in life is change." This quote could summarize the feeling of many educators in the state of Ohio since being awarded RttT funds. Ohioans have seen changes to their local school district report cards, testing systems, and the ways that teachers have been evaluated since 2012. These reforms have increased the accountability teachers and school districts have to local taxpayers and parents, while also providing data on teacher and student performance.

Teacher Evaluation

Danielson (2007) created a framework for teaching that divides the complex activity of teaching into 22 different components. These components encompass different aspects of a teacher’s responsibilities as documented through research promoting improved student learning (Danielson, 2007). These components are categorized into four domains of teaching responsibility: planning and preparation, classroom environment, instruction, and professional responsibilities. The different components and domains each have an accompanying rubric that places teachers at an unsatisfactory, basic, proficient, or distinguished level, broken into specific elements. Marshall (2013) used the work of Danielson (2007) and others when he developed his recommendations for teacher evaluation. Marshall (2013) uses rubrics based on a four-point
scale to evaluate teachers; the levels of performance are: highly effective, effective, improvement necessary, and does not meet standards. Ten criteria have been identified by Marshall (2013) for each of six domains: planning and preparation for learning; classroom management; delivery of instruction; monitoring, assessment, and follow-up; family and community outreach; and professional responsibilities (p.127). In current evaluation research, the emphasis is on the use of rubrics to help educators identify effective practices and areas for improvement.

Along with using rubrics, Marshall (2013) has other recommendations for improving the practice of teacher evaluations. Instead of conducting a few full-length lesson observations that can be rehearsed and prepared for, an evaluator may get a more holistic picture of what is happening in a teacher's classroom by conducting numerous short walkthroughs with follow-up conversations. Principals should conduct two or three five-minute walkthroughs daily (Marshall, 2013). Along with walkthroughs and rubrics, principals should have a positive instructional impact by participating with teachers in their unit planning and working with teacher-based teams (Dufour & Dufour, 2012; Marshall, 2013).

With this research available to policymakers, the 129th Ohio general assembly added to the Ohio Revised Code (ORC 3319.112) a law requiring the adoption of a standards-based framework for the evaluation of teachers. With this law, the Ohio Teacher Evaluation System (OTES) was created, with every school district in the state of Ohio eventually adopting this framework. Teachers in Ohio will be rated as either: accomplished, skilled, developing, or ineffective. These final ratings are calculated as a percentage based on student growth and another percentage based on the evaluators rating. By May 10 of each school year, the evaluator must submit the results of each teacher's evaluation with his/her rating online to the Electronic Teacher Principal Evaluation System (ETPES).
To evaluate a teacher formally in Ohio, one must be OTES credentialed. To initially be credentialed, the evaluator must attend three-day training and pass an online credentialing test. Bi-annually, evaluators must also receive updated training and pass another test. This credentialing process is designed to ensure that all evaluators have similar training in evaluating teachers, and therefore, teacher evaluations should be consistent throughout the state.

Once trained and credentialed, the evaluator collects evidence on the individual teacher's performance from classroom observations, classroom walkthroughs, and pre-conferences and post-conferences. This evidence is then put into and categorized in the OTES with the rubric, which is tied to the standards for Ohio educators. This rubric has ten domains, including focus for learning, assessment data, prior content knowledge/sequence/connections, knowledge of students, lesson delivery, resources, differentiation, classroom environment, assessment of student learning, and professional responsibilities (Ohio Department of Education, 2015). Many of the descriptions in the domains of this rubric are consistent with the recommendations from previous research (Danielson, 2007; Marshall, 2013).

The standards-based evaluation of teachers is one of the reform efforts that Ohio recently implemented to improve teacher and student performance. Another change that has occurred is the how students are assessed in grades 3-8 and high school every year.

Testing

Standardized testing has been administered to Ohio students since the 1990s when for graduation, proficiency tests were required. Ohio revised Code 3301.0710, which mandates that students must be tested annually in English language arts and mathematics in grades 3 through 8 along with science tests in grades 5 and 8. Ohio students are also tested on end-of-course tests at the high school level after completing English 1, English 2, Algebra, Geometry (or Integrated
Math 1 and 2), Biology, United States History, and United States Government. Students are rated at a different level of skill depending on their performance on each test, including advanced plus, advanced, accelerated, proficient, basic, or limited.

In previous years, Ohio switched the type of assessments given to students in grades 3–8 and in high school. Until the 2013–2014 school year, the Ohio Achievement Assessments (OAAs) were administered to students in grades 3–8 in reading and math, and sophomores took the Ohio Graduation Tests (OGTs). Content standards introduced in 2002 were the basis for the OAAs and OGTs. For the 2014–2015 school year, Ohio students were administered national assessments developed by the Partnership for Assessment of Readiness of College and Careers (PARCC). Students were administered Ohio assessments developed by the American Institutes for Research (AIR) beginning in the 2015–2016 school year. The PARCC and AIR tests gave school districts the option of administering the tests online or through paper and pencil during the 2014–2015 and 2015–2016 school years. All tests were administered online during the 2016–2017 school year. Discrepancies became evident between the two administrations of the PARCC test, with online administrative results scoring lower (Molnar, 2016).

**District Report Card**

The results of these tests are reported in individual school districts and are published annually in a report card issued to the school district. These reports provide detailed information on the district's performance on state-wide tests (achievement) and an average of student improvement on state-wide tests (value-added or progress), K-3 literacy, gap closing, student attendance rates, graduation rates, a prepared-for-success measure, and more. Currently, the state issues a letter grade for how well the district is performing in these various areas, but more information and data is available with additional research. These report cards provide families,
the community, and educators the information they need to understand how a district is performing (ODE Website, 2017).

Beginning in 1998 with Senate Bill 55 by the 122nd Ohio General Assembly, the Ohio Department of Education has issued an annual report card for each school district, in line with Ohio Revised Code 3302. Starting in 2002, with the passage of No Child Left Behind at the federal level, the local report cards became more detailed with requirements to contain additional information relating to adequate yearly progress (AYP) and disaggregation by subgroup. More recently, in 2015, Ohio transitioned to an interactive local report card that increases its online availability, issues letter grades, and changes the term of the "value-added" rating to "progress."

The results of these report cards may affect how a community perceives its school district. Wheatley (2012) found that although there are community pressures for school improvement, local report card components do not serve as a good predictor of school district tax levy passage. It has also been found that poor school report card data does not have an effect on superintendent tenure or school board tenure (Kogan, Lavertu, & Peskowitz, 2016). School report card grades do have a strong correlation with the cost of housing in the community (Figlio & Lucas, 2004). With open enrollment policies in many Ohio districts, school choice for parents and students is also an issue. Students are more likely to leave their local school district of residence and enroll in another school district when public information reveals poor academic performance in the district of residence (Friesen, Javdani, Smith, & Woodcock, 2012). District report cards have correlations with the local housing market and school enrollment. Performance Index scores and Value-added ratings are two components of the local report card that will be investigated further in this study.
Performance Index Scores

According to ORC 3302.01, “performance index score means the average of the totals derived from calculations, for each subject area, of the weighted proportion of untested students and students scoring at each level…on the state achievement assessments.” These tests refer to all state assessments administered in grades 3 through 8 and the English language arts and mathematics assessments administered in high school. The performance index score may range from 120 if all the students in the district test at an advanced level to 0 if none of the students in the district are tested. The table below shows the weights assigned to each skill level with each weighted score multiplied by the percentage of students at that level (Ohio Department of Education, 2017).

Table 1
Calculation of Performance Index Scores

<table>
<thead>
<tr>
<th>Proficiency Level</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Advanced Plus</td>
<td>1.3</td>
</tr>
<tr>
<td>Advanced</td>
<td>1.2</td>
</tr>
<tr>
<td>Accelerated</td>
<td>1.1</td>
</tr>
<tr>
<td>Proficient</td>
<td>1.0</td>
</tr>
<tr>
<td>Basic</td>
<td>0.6</td>
</tr>
<tr>
<td>Limited</td>
<td>0.3</td>
</tr>
<tr>
<td>Tests Not Taken</td>
<td>0.0</td>
</tr>
</tbody>
</table>

The performance index score is a way to display the academic achievement of a school district’s students numerically. School district demographics such as poverty levels are strong predictors of a school’s performance index score (Bennet, 2010; Fleeter, 2016). Since there is such a strong correlation between the percentage of economically disadvantaged students in a school district and the district’s performance index score, this study also investigates value-added scores, which Ohio refers to as its progress measure.
Value-Added

Ohio Revised Code 3302.021 defines "value-added" as a measure of academic gain for a student or group of students over a specific period of time. For example, if a student scores in the forty-fifth percentile as compared to his peers across the state on a mathematics test at the end of fourth grade and then scores in the fifty-fifth percentile at the end of fifth grade, this student will have shown positive growth. In Ohio, value-added scores are not calculated through the use of percentiles; instead, they are calculated using national curve equivalents (NCE's) which are similar to percentiles but calculated differently. If the averages of students in a school district display positive growth, then the school district will have a positive value-added score. Conversely, if the averages of students in a district demonstrate negative growth with decreases in their percentile scores year to year, then the district will have a negative growth score. Value-added scores are reported as district progress on the Ohio School District’s report cards.

As stated previously, student growth is a component of teacher evaluations. If a teacher receives a value-added score, then that score is used as part of the evaluation rating. Hanushek and Rivken (2010) advocate for the use of value-added ratings as part of teacher evaluations along with supervisor observations because of the objective nature of the measure. Harris and Sass (2014) have found that value-added ratings of teachers are a more accurate predictor of a student’s future success than a principal’s evaluation of teachers. Thomas (2014) found statistically significant negative teacher attitudes towards the use of value-added scores in Ohio teacher evaluations. Darling-Hammond (2015) utilized teacher quotes to argue against using value-added scores in teacher evaluations and suggested the use of multiple measures of teacher effectiveness, including student-learning objectives (SLOs), teacher-designed rubrics, and
principal observations. Some issues with using value-added ratings in teacher evaluation include the failure to take measurement error into account, the possibility of roster verification error, and the difficulty of revising a teacher’s rating as more information becomes available (Ballou & Springer, 2015). The use of value-added ratings in teacher evaluations has been both recommended and criticized.

This study utilizes value-added models to compare the success of various school districts for various reasons. This modeling is based on student growth, not proficiency or achievement. Therefore, the modeling is fair regardless of whether students are high-achieving, low-achieving, highly mobile, in special education, or from specific socioeconomic or demographic groups (Ohio Department of Education, 2016a). Value-added ratings are among the most reliable, least biased, and most suitable models showing student growth (McCaffrey & Lockwood, 2008). Selvage (2013) demonstrated that there is no relationship between a district’s value-added scores and community demographics.

While the Ohio Department of Education utilizes student test scores to report each school district's achievement and growth, the department also compiles other data to compare districts. School districts can identify their closest neighbors socioeconomically utilizing a tool on the interactive district report cards known as the similar district methodology.

**Similar District Methodology**

An additional feature of the local report cards in Ohio is the ability to compare similar districts. Ohio is a very diverse state with 610 school districts. These school districts vary in size, location, demographics, and the communities they serve. Some serve under 500 students while others serve over 10,000 students. Ohio has school districts that are urban while other districts are in rural, suburban, or small-town locations. Some school districts serve a very high-
wealth population while other districts serve a majority of students who come from very low incomes. The Ohio Department of Education compiles a list for each school district of up to 20 of their most similar "neighbors" in the state. In some cases, there may be fewer than 20 similar districts identified such as the identification of large, urban school districts.

"In order to evaluate performance for a given district, it is often useful to consider how similar districts compare on the same data" (Ohio Department of Education [ODE], 2016b, p. 1). A similar district methodology has been created that takes into account district size, poverty levels, socioeconomic status, a rural/urban continuum, the race/ethnic makeup of the district, and the non-agricultural and non-residential tax capacity of the district. The district size is computed by the measure of average daily membership (ADM) or the number of students served by each school district. Poverty levels are calculated by using the percentage of students who are considered economically disadvantaged, as reported to ODE by district education management information system (EMIS) coordinators. The socioeconomic status of a district is calculated by using information from the 2010 federal census using a composite score of median income, the percentage of the population with a college degree, and percentage of the population with an administrative/professional occupation. The rural/urban continuum distinguishes school districts that have an urban feel from those that have a rural feel, combining measures of population density, the percent of the agricultural property, population data from the census, and the incorporation of a city of over 40,000 people. The measure of race/ethnicity compares the diversity of the student population in districts utilizing data reported through EMIS (education management information system). The tax capacity of a district measures the community's ability to generate revenue from its property tax base. These measures are then combined to identify districts that are closely related to other districts in the state. Comparisons can then be
easily made using achievements and other data to see how one district is performing relative to its peers.

The Ohio Department of Education does recognize three limitations of using this similar district methodology. First, the method does not include a geographical dimension, which is often more useful in comparing financial data such as salaries. Second, the method selects up to the nearest 20 neighbors. Sometimes, many other similar districts could be used for a comparison, but sometimes, there are not 20 districts that are closely related. Lastly, this is not a grouping of 20 districts; instead, the state has identified the 20 most similar districts for each district (ODE, 2016b).

The State of Ohio has had many educational reforms in the past five years, including changing the way teachers are evaluated, changing the format and style of assessments students take, and changing the way these test results are reported to the general public. School districts in Ohio who have performance pay provisions in their teacher contracts have also begun to appear (Willis & Ingle, 2016). The next section of this literature review is merit or performance pay.

**Performance Pay**

As previously mentioned, the traditional single-salary schedule is currently out of favor with many policymakers as a way to remunerate teachers. Along with reforms in the ways teachers are evaluated, students are tested, and the results of these reported, school districts have begun to look at merit or performance pay for teachers. This section discusses performance pay, starting with a brief history of how teachers have been compensated, categorizes different forms or types of performance pay, analyzes arguments for and against performance pay, explains
different types of performance pay, and identifies Ohio districts who have a form of performance-pay.

**History of Paying Teachers**

There have been three main ways that teachers in the United States have been compensated since the early 1800s (Prostik, 1996). With the county being mainly a rural agricultural society in the nineteenth century, the boarding-around method of paying teachers was popular in that era. In this era, nearly eighty percent of the nation's population lived in rural areas, and half of the working citizens were farmers (Prostik, 1995). The school calendar often revolved around harvest and planting seasons so that workers were available on the farms. One-room schoolhouses were the norm with teachers often boarding with a different family in the community each week. These teachers were often young, yet-to-be married females, and the communities' evaluation of them was often a judgment of the teacher's morals based on how she acted while staying with different families. Much of the work by students were done using teacher-proof curricular materials such as the McGuffey Reader (Fuller, 1982).

As our country became industrialized, the boarding-around system became ineffective. More students needed higher levels of education, increasing the need for more teachers and a grade-based system (Tyack, 1974). As students began attending secondary schools in the late 1800s and early 1900s, further educator training was needed to teach the higher grades. Many of the one-room schoolhouses consolidated into county systems led by a professional superintendent of schools. The responsibility for evaluating teachers shifted from the community's perception of the teacher to the administrator's evaluation of teaching method or skill. Minimum salary requirements were passed in state law, with many local communities adopting salary schedules where the secondary teachers were paid more than the elementary
teachers. Teaching at the secondary level was more accessible to white males, resulting in
gender- and race-based inequalities in the grade-based system (Guthrie et. al., 2007). The graded
pay system was overtly racist, and many people in this era began to believe that teaching was
teaching, regardless of grade level being taught (Odden and Kelley, 2002).

During the late 1800s and early 1900s, labor unions such as the American Federation of
Labor and Congress of Industrial Congress of Workers began to organize. In education,
organizations such as the Interborough Association of Women Teachers in New York pushed for
the collective bargaining of teacher contracts. With the rise of labor unions and the push for
equal pay for equal work, in 1921, Denver, Colorado, and Des Moines, Iowa, became the first
cities to introduce the single-salary schedule for teachers (Sharpes, 1987). Initially called the
"position automatic schedule," the single-salary schedule is a chart where teachers are paid based
on their years of experience and education levels attained (Moelman, 1927). By 1950, ninety-
seven percent of all school districts in the United States had adopted the single-salary schedule
(Sharpes, 1987). Although there have been attempts to move away from the single-salary
schedule, in 2006, Podgursky estimated that ninety-six percent of public school districts still
compensated teachers under a single-salary schedule.

Shortly after the single-salary schedule had come into existence, it was criticized. As
found in Willis and Ingle's (2016) study, in 1932, Douglas Scates stated the following:

The task of answering the question, for what are teachers paid? Carries a number of
implications... No one will say that teachers keep on improving indefinitely; and there are
few who believe that most teachers keep on improving significantly in the quality of their
service for a full ten years. Yet, we have many schedules granting increases for eleven,
twelve, thirteen, or more years (p.337).
Odden and Kelly (2002) indicate that treating teachers as equals does not account for the variance in teacher performance.

**Arguments for Performance Pay**

There are many advantages to moving towards a performance-based pay system for teachers. There has been a push from the business community to pay teachers for their performance. The involvement of business leaders in school reform efforts often results in a recommendation for some form of merit pay based on rationalized performance standards that are measurable by some objective means (Schlechty, 2009).

People have advocated for performance pay for teachers for various reasons. There is a link between performance-pay and increased student performance (Figlio & Kenny, 2007). Researchers believe that such a system will not only lead to better student achievement but also make more efficient use of limited education dollars (Hanushek & Lindseth, 2009). They also argue that a school finance system could play a part in motivating schools, teachers, and administrators (Hanushek & Lindseth, 2009). Hanushek and Lindseth (2009) argue the following:

The current manner in which teachers are compensated not only fails to meet the goal of motivating teachers to improve student performance, it actually works against it because as long as teachers maintain their jobs, even if just barely, they are paid the same as every other teacher in the system with the same number of years teaching under their belts and the same college degrees.

The recruitment of teachers is another reason to switch to a performance-based pay system. School districts will be able to hire and retain a more qualified pool of teachers as candidates realize that they can both teach and enjoy the economic reward for a good job
The higher marginal incentives associated with merit pay should, theoretically, attract high-ability and less risk-averse people to teaching (Goldhaber, et. al., 2011). Performance-pay would enable districts to pay more to teachers in shortage areas such as math and science (Hanushek & Lindseth, 2009).

Introducing performance pay to younger teachers may be an excellent way to start since “high school teachers early in their careers are about twice as likely as high school teachers near the end of their careers to support subject area pay, combat pay, and incentives” (Goldhaber, et. al., 2011). It is important to understand teachers’ attitudes related to performance-pay when implementing such a system because the system will most likely fail if teachers are not supporting this change.

Teacher Attitude

Understanding teacher attitudes about performance-pay should not be underestimated because “teacher compensation reforms can founder, for example, when policymakers pay little attention to teachers’ acceptance of the reform” (Goldhaber, et. al., 2011). Evidence on teachers’ opinions about compensation reform is mixed (Goldhaber, et. al., 2011). The traditions of egalitarianism and service in teaching might lead teachers to resist wage differentiation (Goldhaber, et. al., 2011). High school teachers, in contrast to elementary teachers, are more supportive of performance pay. Teachers who have a higher sense of trust and respect regarding their fellow teachers are less supportive of performance pay, whereas those who have a higher sense of trust and respect regarding their principal are more supportive of performance pay (Goldhaber, et. al., 2011).
Tenhiälä and Lount (2013) provided an overview of how professors reacted to pay system reform in Finland. Their study observed that professors had both negative and positive affective reactions to a policy-mandated pay system reform in their university system.

The impact of a pay reform on employee behavior would be underestimated by ignoring the role of affective reactions. Findings suggest that employees’ negative affective responses to pay system reform, but no negative trait effect was related to high levels of helping and voice. Negative mood has been found to promote helping behaviors, which explains why employees experiencing negative affect can sometimes perform especially well. We did not find evidence that employees experiencing higher levels of negative affect were more likely to leave the organization in the next two years after the pay system was introduced.

This study demonstrates that even when employees do not like a change in their pay structure, positive outcomes can occur. In this situation, employees who were not happy with the change increased their communication with each other and often helped fellow employees understand the change.

A qualitative interview study of teachers at four schools in Sweden that had implemented a performance pay system was conducted. “The majority of these teachers perceived the system as arbitrary, unfair, unclear, and felt that it fosters an awkward working environment” (Tenhiälä & Lount, 2013). The teachers in this study did not feel that the system motivated them or promoted an increased commitment to their teaching (Tenhiälä & Lount, 2013). Some teachers such as vocational teachers did not believe that principals had the knowledge or understanding of their work to evaluate them adequately and tie raises to their performances (Tenhiälä & Lount, 2013). Teachers were also critical of a performance pay system because “it
can be used to punish or silence teachers who are critical or have disparaging opinions” (Tenhiälä & Lount, 2013). Such a system has caused an adverse organizational climate, and teachers mention that it undermines cooperation (Tenhiälä & Lount, 2013).

Research has demonstrated that teachers’ opinions and attitudes about performance pay are mixed with some teachers supporting it and other teachers seeing many problems with such a system. In the following section, we will explore some of the arguments against implementing such a system.

**Arguments against Performance Pay**

Reasons for changing from the traditional teacher pay scale to a performance pay system have been discussed, and teacher attitudes toward these changes have been accounted for. Others have demonstrated that performance pay for teachers may not be a good idea. It is just not feasible to compensate teachers by performance evaluations. There is too little agreement on the goals of education; the relationship between the actions of teachers and the learning of students is too complicated and challenging to trace (Ballou, 2001). The potential for demoralization and breakdown in cooperation is too high (Ballou, 2001). Teachers are wary of a pay system that raises the specter of teachers competing with one another (Goldhaber, et. al., 2011). “Performance pay systems have displayed many instances of cheating, and a program in Nashville had zero impact on student performance” (Pink, 2009, p.191). Performance pay does not influence overall career satisfaction or the sense of effort rewarded, and it has a negative influence on satisfaction with salary (Belfield & Heywood, 2008). Based on this research, it would appear that changing to a performance pay system for teachers could lead to negative consequences.
Forms of Performance Pay

Moving away from a single salary schedule towards performance pay has come in many forms, with individual districts attempting different approaches. Odden and Kelly (2002) describe performance pay plans and career ladders as alternative approaches to the single salary schedule introduced after *A Nation at Risk*. Odden and Kelly (2002) describe new compensation approaches such as skill and competency pay, compensation for outcomes or behaviors, group performance awards, and benefits as part compensation. Podgursky and Springer (2006) classify 20th-century compensation experiments as either merit-based pay or knowledge and skill-based pay. Springer (2009) identifies six different types of compensation reform including pay-for-performance, knowledge and skill-based pay, career ladders, hard to staff subjects, hard to staff schools, and recruitment and retention awards [see table 2].

Table 2

Springer’s (2009) Types of Compensation Reforms

<table>
<thead>
<tr>
<th>Type</th>
<th>Description</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pay for Performance</td>
<td>Rewards based on predetermined tasks or outcomes, or both, related to teacher and student behavior.</td>
<td>Teacher collaboration, professional development, and student behaviors are input examples. Output examples are student test scores, graduation rates, dropout rates, and attendance rates.</td>
</tr>
<tr>
<td>Knowledge and Skill-Based</td>
<td>Rewards based on completion of teacher activities that are related to student outcomes, as well as a demonstration of classroom mastery.</td>
<td>Portfolio completion, dual certification, graduate degree in subject taught, teacher evaluation ratings, National board certification</td>
</tr>
<tr>
<td>Career Ladders</td>
<td>Provides new roles for teachers with additional pay and responsibilities as they increase their knowledge and skills. Plans typically involve vertical steps with multiple</td>
<td>Additional training and professional development, earning advanced degrees, assuming higher levels of instructional responsibility, mentoring new teachers.</td>
</tr>
</tbody>
</table>
objectives within each step.

<table>
<thead>
<tr>
<th>Hard to Staff Subjects</th>
<th>Incentives are targeted to teachers in subject areas where there are shortages, which are based on need in the school, district, or state level.</th>
<th>Teachers are compensated for pursuing subject area endorsements in high needs areas, teaching in a high need subject area.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hard to Staff Schools</td>
<td>Incentives are offered for teaching in high needs schools or districts, typically either high-poverty, low-performing, or geographically remote schools. Like hard to staff subjects, these incentives are designed specifically to address market factor influences.</td>
<td>Teachers are awarded a bonus to work in a high-needs, hard-to-staff school.</td>
</tr>
<tr>
<td>Recruitment and retention Awards</td>
<td>Rewards are offered for signing a contract to working a specific school or district. Annual bonuses are offered each year of continued service in the school or district.</td>
<td>Continuity awards, contract bonuses, etc.</td>
</tr>
</tbody>
</table>

**Ohio Districts with Performance Pay**

Willis and Ingle (2016) utilized Springer's (2009) taxonomy of teacher compensation reforms and identified 16 Ohio school districts that have some form of performance pay provision in their negotiated agreement. In this study, all collective bargaining agreements between teacher unions and boards of education going into effect after September 24, 2012, were analyzed, with close attention paid to compensation provisions. September 24, 2012, was chosen as a cut-off date because that is when Ohio Revised Code 3319.112 (E) became effective, requiring a standards-based evaluation system, such as OTES, for teachers. The 16 school districts identified as having some type of pay-for-performance provision include: Arlington, Centerburg, Cincinnati, Clear Fork Valley, Cleveland, Grandview Heights, Hudson City, Indian Hill, Liberty Benton, Licking Heights, Mentor, Oakwood, Perrysburg, Ripley-Union-Lewis-
Huntington, Vandalia Butler, and Vanlue (Willis & Ingle, 2016). These districts vary in location and typology.

Willis and Ingle (2016) also grouped the identified contracts into four sub-categories: performance stipends, performance schedules, performance rates, and hybrid systems. Centerburg, Licking Heights, and Ripley-Union-Lewis-Huntington have performance stipends linked to district performance or teacher attendance. Clear Fork Valley is the only district identified using performance stipends tightly linked to the Ohio Teacher Evaluation System, where monetary rewards are given to teachers in addition to their base salary if they are rated accomplished or skilled in their evaluation. Cleveland, Grandview Heights, Hudson City, Indian Hill, and Perrysburg use a performance schedule, where teachers must meet specific performance indicators to move on the salary schedule. Each district has performance indicators that must be met. Liberty Benton, Mentor, Oakwood, and Vanlue utilize performance rates, where a teachers' salary is recalculated annually based on specific performance criteria. Cincinnati’s performance pay system falls into a hybrid classification, where elements of multiple evaluation systems allow teachers to advance through the salary schedule. Even though Willis and Ingle (2016) identified 16 different school districts in Ohio that have some type of performance pay, each contract is different, and the merit pay provisions can be broken into four subcategories.

Anstadt (2017) conducted a mixed-method program evaluation of Perrysburg’s performance-based compensations system. She found statistically significant differences with attendance, mentoring participation and teacher evaluation ratings; with teachers electing to participate in Perrysburg’s performance-based compensation system receiving higher teacher evaluation ratings, more participation in mentoring activities, and better attendance rates.
Most research publications have a section recommending future research on the topic. Hulleman and Barron (2010) state that "little research has been done on performance pay in education" (p. 27). The analysis conducted by Willis and Ingle (2016) is no different. They state, "Future research also may want to examine...possible impacts of merit pay provisions to outcomes such as ...student achievement at the teacher, school, and district levels" (p. 34). This recommendation is justification for this study.

Summary

This chapter provides a review of previously published literature relating to this study. The chapter began with a discussion on motivational theory, including goal setting theory, contingency theory, social dilemma theory, participative management theory, and expectancy theory. Reforms at the federal and state level were also discussed, including federal acts and grants, as well as a description of relevant Ohio Revised Code (ORC) statutes and changes in testing and teacher evaluation procedures. The chapter concluded with a teacher compensation review, beginning with a history of paying teachers and concluding with a discussion of the Willis and Ingle (2016) study. Chapter 3 discusses the research methods used in this study.
CHAPTER III: METHODOLOGY

This study compares the performance index scores for school districts who have implemented some version of performance pay for teachers to a systematic sample of similar districts who have not implemented such pay structures. This chapter describes the methodology of this study, including the research design, participants, instrumentation, procedures, research questions, data analysis, including variables, and limitations.

Research Design

Causal-comparative research compares two or more groups regarding a cause (performance pay for teachers) that has already happened (Creswell, 2014). Sondergeld (2015) suggests that the purpose of causal-comparative research is to understand the relationship between a group membership and outcomes. In this study, there are two groups, one group having performance-pay language in a teacher contract, and the other group or not having performance-pay provisions in their contract. The outcomes are the performance index scores, value-added ratings, and teacher evaluation ratings of the different groups. The independent variables are naturally occurring, do not involve researcher intervention, and are made up of a comparison group.

This study utilizes a causal-comparative research design because the independent variable of whether a school district has implemented a performance-pay system is naturally occurring and cannot be manipulated. These two groups are compared without using researcher intervention or random assignment; instead, this study examines the phenomena of performance pay for teachers as it already exists in Ohio (Sondergeld, 2015). This study can also be considered ex-post facto or after the fact because districts using a performance pay system have already been identified in the Willis and Ingle (2016) study.
Two limitations with using causal comparative design are the inability to infer causality and the effect on extraneous variables (Sondergeld, 2015). Since participants are not randomly assigned to different groups in causal comparative designs, causality cannot be claimed and any relationship observed between the independent and dependent variable is only suggestive of causality. Other extraneous variables could affect performance index scores, value-added ratings, and teacher evaluation ratings. This study attempts to control for these extraneous variables through utilization of similar districts in the creation of the comparison group.

**Participants**

Willis and Ingle (2016) identified 16 school districts in Ohio who have implemented merit pay provisions in their collective bargaining agreements. These districts include Arlington, Centerburg, Cincinnati, Clear Fork Valley, Cleveland, Grandview Heights, Hudson City, Indian Hill, Liberty Benton, Licking Heights, Mentor, Oakwood, Perrysburg, Ripley-Union-Lewis-Huntington, Vandalia Butler, and Vanlue. After contract verification, these districts were identified as no longer having performance-based pay provisions and therefore dropped from the study: Vandalia Butler, Grandview Heights, Clear Fork Valley, Centerburg, and Ripley-Union-Lewis-Huntington. The Ohio Department of Education identifies up to 20 similar districts for each district in the state. There are only six large urban school districts in Ohio (Canton and Youngstown are grouped with other smaller city school districts), and therefore, Cleveland and Cincinatti only have four other peers in their group.

By comparing the districts that have been identified as having implemented performance pay for teachers to districts identified as similar, this study controls for confounding variables such as district size, geographic classification (rural, urban, small town, and suburban), and socio-economic make-up. This study utilizes a single stage sampling method because the
districts that have implemented a performance pay structure have already been identified, and similar districts are also identifiable, utilizing the similar district grouping tool available on the Ohio Department of Education website. This sampling was purposeful, selecting districts that meet the criteria for the study.

The two most similar districts to each performance pay district are used as a comparison group. This study utilized data from the fiscal year 2017 when determining similar districts. If the district has previously been identified as a member of the comparison group because of another similar district, then the 3rd, 4th, 5th, or sixth most similar district was utilized in the comparison group. Two comparison districts instead of one comparison district were utilized as a comparison group for each identified district to increase the number of districts in the study. With a more substantial number of participating districts, there should be less sampling error present (Roberts, 2010).

Table 3

*Performance-Pay Districts and Participating Comparison Districts*

<table>
<thead>
<tr>
<th>Performance Pay District</th>
<th>Comparison Districts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arlington</td>
<td>Botkins</td>
</tr>
<tr>
<td>Cincinnati</td>
<td>Toledo</td>
</tr>
<tr>
<td>Cleveland</td>
<td>Columbus</td>
</tr>
<tr>
<td>Hudson City</td>
<td>Upper Arlington</td>
</tr>
<tr>
<td>Indian Hill</td>
<td>Orange</td>
</tr>
<tr>
<td>Liberty Benton</td>
<td>Johnstown-Monroe</td>
</tr>
<tr>
<td>Licking Heights</td>
<td>Canal Winchester</td>
</tr>
<tr>
<td>Mentor</td>
<td>West Clermont</td>
</tr>
<tr>
<td>Oakwood</td>
<td>Wyoming</td>
</tr>
<tr>
<td>Perrysburg</td>
<td>Avon</td>
</tr>
<tr>
<td>Vanlue</td>
<td>Ridgemont</td>
</tr>
<tr>
<td></td>
<td>New Reigel</td>
</tr>
<tr>
<td></td>
<td>Akron</td>
</tr>
<tr>
<td></td>
<td>Dayton</td>
</tr>
<tr>
<td></td>
<td>New Albany-Plain</td>
</tr>
<tr>
<td></td>
<td>Revere</td>
</tr>
<tr>
<td></td>
<td>Rootstown</td>
</tr>
<tr>
<td></td>
<td>Northmont</td>
</tr>
<tr>
<td></td>
<td>Kettering</td>
</tr>
<tr>
<td></td>
<td>Bexley</td>
</tr>
<tr>
<td></td>
<td>Kings</td>
</tr>
<tr>
<td></td>
<td>Waynesfield-Goshen</td>
</tr>
</tbody>
</table>

*Instrumentation*

For this study, the performance index scores and value-added ratings from the 2016-2017 school year for each district were needed. Ohio teacher evaluation system (OTES) ratings of
teachers from the 2016-2017 school year for each school district were also needed. The performance index scores, value-added ratings, and teacher evaluation ratings were collected through Ohio Department of Education’s online interactive local report card. Since all of the data needed for this study is available on Ohio Department of Education’s website, no unique instrumentation was used. However, the performance index scores and value-added scores published in each district's interactive local report card are dependent upon the results of statewide assessments given annually.

Reliability produces the same results on repeated trials and remains stable and consistent over time (McMillan, 2008). Although there have been inconsistencies in performance index scores because of changes in the test format from PARCC assessments to AIR, how performance index scores are calculated has remained the same. The way in which value-added scores are calculated has remained consistent with previous years. Teachers have been evaluated under the same process and rubric since the 2012-2013 school year, with evaluators becoming credentialed through similar training and assessments to ensure inter-rater reliability (Roberts, 2010, p. 152).

Validity refers to measuring what one is trying to measure (McMillan, 2008). With this being a causal comparative research design, with participants identified through the Willis and Ingle (2016) study and similar district methodology, the participants were non-randomly selected. Inferential statistics are made for use with random samples, and the validity of this study is compromised with the use of purposeful participant selection (Reinhart, 2015).

Results of tests comparing these districts’ performance index scores, value-added ratings, and teacher evaluation ratings from different years are unknown at this time, but this study could easily be replicated using the same methods and districts. The three measures being utilized are further described in the following sections.
Performance Index Scores

Performance Index Scores are calculated by the results of how all students in a school district perform on state-wide testing. Although the poverty levels of a school district are a strong predictor of its performance index score (Bennet, 2010; Fleeter, 2016), utilizing similar districts as a comparison group alleviates concerns relating to the reliability and validity of this measure.

Value-Added Scores

For the third year in a row, Ohio has switched the type of assessments given to students in grades 3-8 and high school. In the 2013-14 school year, the Ohio Achievement Assessments were administered to students in grades 3-8 in reading and math, and the Ohio Graduation tests were given to sophomores. These tests were based on the content standards introduced in 2002. For the 2014-15 school year, Ohio students were administered national assessments developed by the Partnership for Assessment of Readiness of College and Careers (PARCC). Students were administered Ohio assessments developed by the American Institutes of Research (AIR) since the 2015-16 school year. The PARCC and AIR tests gave school districts the option of administering the tests online or using paper and pencil during the 2014-2016 school year. Discrepancies were discovered between the two administrations of the PARCC test, with online administrative results scoring lower (Molnar, 2016). As a result of these discrepancies, all Ohio school districts were required to administer the tests online during the 2016-2017 school year. With the constant changes in Ohio assessments, their reliability and validity of value-added scores have been thrown into question (Hoover, 2014).

Value-Added Scores are published annually by the Ohio Department of Education at the teacher, school, and district level. Each student receives a value-added score based upon their
expected growth on a test, teachers receive a score based upon the average of all their students, and the district receives a score based on the average growth of all the students receiving a value-added rating. A roster verification process is administered before value-added scores are assigned. This is a three-step process consisting of: school set-up, teacher review, and administrator verification. These scores are the most reliable and least biased in displaying student growth (McCaffrey & Lockwood, 2008).

**Teacher Evaluation Ratings**

Ohio teachers are evaluated under the OTES model and must be evaluated by a credentialed evaluator. The credentialed evaluator uses the OTES rubric to collect data from the teacher and assign a rating; which accounts for 50% of the teachers' evaluation rating. The other 50% of the evaluation rating is calculated via growth measures such as: value-added scores, growth scores from vendor assessments, or growth scores from student learning objectives. To become a credentialed evaluator, one must attend a three-day in-person training and pass an online test. These assessments require the potential evaluator to watch a video of a teacher leading a lesson, then rate the teacher using the OTES rubric. This process ensures that all evaluators throughout the state are consistent in their assignment of ratings. Every two years, a credentialed evaluator must complete on-line re-calibration training and pass an additional online test administered through the National Institute for Excellence in Teaching (NIET). Through the training and assessment of all credentialed OTES evaluators, evaluations should be consistent throughout the state.

**Procedures**

The first step in most research projects is obtaining Institutional Review Board (IRB) approval. The federal definition “human subject” is not met in this study and the IRB has
formally determined this through the exempt review process by submission of a "review
determination form" (See Appendix 1). Since the participants in this study are school districts
and not human subjects, full human subjects review board approval is not necessary for this
study.

Thirty-three Ohio school districts were identified to take part in this study. The contracts
of these districts were re-checked from the SERB Ohio website and individual district website to
ensure that the performance-based pay group still has performance pay provisions in their
contracts, and the comparison group does not have performance pay provisions in their contracts.
Their district-wide performance index scores, value-added ratings, and teacher evaluation ratings
from the 2016-17 school year were manually accessed from their interactive district report cards
through the Ohio Department of Education website (education.ohio.gov). These scores and
ratings were then entered into a spreadsheet and then transferred to SPSS for data analysis.

**Research Questions**

The following are the two research questions in this study.

1. Do Ohio school districts that have implemented a performance-pay system differ
   significantly from Ohio school districts that have not implemented a performance-based
   pay system with regard to their performance index scores, value-added ratings, the
   percent of teachers rated proficient or higher, and the percent of teachers rated
   accomplished when using data from the Ohio School district report cards from the 2016-
   2017 school year?

2. Does the type of performance-pay system in Ohio school districts have different results
   with regard to a districts performance index score, value-added ratings, the percent of
teachers rated proficient or higher, and the percent of teachers rated accomplished when using data from the Ohio School district report cards from the 2016-2017 school year?

**Data Analysis**

Both research questions compare the performance pay districts to the traditional pay districts. For the first research question a $t$-test of independent samples and a $t$-test of paired samples was utilized. In addition to conducting data screening procedures, the researcher will calculate group means and standard deviations. Further explanation of these analyses follow.

Table 4

*Data Analysis Procedures by Research Question*

<table>
<thead>
<tr>
<th>Research Questions</th>
<th>Independent Variable</th>
<th>Dependent Variable</th>
<th>Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Performance Pay for Teachers</td>
<td>Performance Index Score</td>
<td>$t$-test of independent samples</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Value Added Score</td>
<td>$t$-test of paired samples</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Teacher Evaluation Rating</td>
<td></td>
</tr>
<tr>
<td>2.</td>
<td>Type of Performance Pay for Teachers</td>
<td>Performance Index Score</td>
<td>Comparison of districts and Control Means utilizing graphs</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Value Added Score</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Teacher Evaluation Ratings</td>
<td></td>
</tr>
</tbody>
</table>

**Variables**

The independent variable is Performance Pay for Teachers, which is a dichotomous, categorical variable of the individual school district. Sixteen Ohio school districts have implemented a form of a performance pay system for teachers (Willis & Ingle, 2016), and these sixteen were narrowed down to eleven after re-checking their contracts. Contracts were verified to ensure that the twenty-two comparison districts do not have performance-pay provisions in them. This variable will be using a nominal scale because the variable will only show a difference between the two types of districts. The similar districts will be identified using the
Ohio Department of Education’s similar district methodology. The three dependent variables are: Performance Index Score, Value-Added Rating, and Teacher Evaluation Ratings.

**Performance Index Score.** Performance index scores are quantitative and can range from 0-120. The Ohio Department of Education has also calculated the performance index percentage by dividing the district's score by 120. The Ohio Department of Education has also compared districts across the state using an A-F scale for performance index; which would be categorical using an ordinal scale. For this study, the range of 57.10-108.8 is utilized.

**Value Added Rating.** Value-added ratings are quantitative and can range from -70 to 70. These ratings are published annually by the Ohio Department of Education and are available on ODE's website through each district's interactive local report card. For this study, a range of -65.08 to 23.66 is utilized.

**Teacher Evaluation Ratings.** The teacher evaluation rating is a quantitative measure, describing the percentage of teachers in each district rated at skilled or accomplished. A second sub-variable, being just the percentage rated accomplished is also utilized. There are four ratings that a teacher may be rated: ineffective, developing, skilled, or accomplished. The percentage of teachers rated skilled or accomplished can range from 0 to 100% and does range in this study from 70.9% to 100%. The percentage of teachers rated accomplished for each district in this study ranges from 6.9% to 94.4%. These rating percentages are available on ODE's website through each district's interactive local report card.

**Possible Extraneous Variables.** An extraneous variable is another variable that may affect the independent variable (Sondergeld, 2015). Many things can affect a student's test scores and performance of index scores, value-added ratings, and teacher evaluation ratings. Some of these include: district size, geographic classification (rural, urban, mall town, and
suburban), and socio-economic background. These are all possible extraneous variables in the study. The researcher attempted to control for these extraneous variables by selecting similar districts as a comparison group.

**Descriptive Statistics**

Descriptive statistics summarize a given data set and measure the central tendency and spread of the data. Group means, standard deviations, frequencies, and range were all utilized in describing the dependent variables of the group of districts that have implemented performance pay for teachers, and the group of districts who have not implemented such structures. A mean is the most common measure of central tendency and is the arithmetic average (Vannatta, 2008). Standard deviations are the most frequently used measure of variability and provides statistical expression for the average amount by which data differ from the mean (Sondergeld, 2015). Frequency is the number of participants in the study. Range is the easiest measure of variability, calculating the distance between the highest and lowest score in a distribution (Pryczak, 2006). All of the descriptive statistics mentioned above give the reader a better understanding of the central tendency and variability of the data set.

**Inferential Statistics**

Independent samples $t$-tests and paired samples $t$-tests were used to compare mean scores of two groups on one dependent variable (Sondergeld, 2015). The independent samples $t$-tests and paired samples $t$-tests were be used for the first research questions. This analysis technique examined if there are statistically significant differences between the performance-pay group and comparison group with regards to performance index scores, value-added ratings, and teacher evaluation ratings.
Prior to conducting these analyses, the data was screened for outliers and the test assumptions. The first parametric test assumption is that the dependent variable must be continuous at the interval or ratio level. The second assumption is that the scores of the dependent variable are normally distributed amongst a bell-curve. The third assumption is that the spread of the scores of the group are similar or have homogeneity of variance (Reinhart, 2016).

After data screening, $t$-tests of independent means and paired samples $t$-tests were conducted. A $t$-test of independent means and paired samples $t$-tests compare the significance between two different groups. When an independent variable is nominal, and the dependent variable is interval or ratio, a $t$-test is used. For the results of this study to be significant, the $t$-value should be greater than 1, and the $p$-value should be less than .05. The $p$-value represents the chances of finding significance out of error but does not indicate the size of difference or effect.

**Limitations**

Three limitations are discussed for this study. First, there are time restraints on this research. Value-added ratings are most reliable after three consecutive years of administering the same test. With changes in the landscape of assessments in Ohio, students have only been administered tests created by the same vendor two years in a row. This study uses the value-added data that is available to finish this dissertation, and future research can utilize future data sets.

Second, with regards to time restraints, the Willis and Ingle study was published in 2016. It is possible that since the publication date, more school districts have added performance pay provisions to their contracts. Identified school districts may also have deleted performance pay
provisions from their contracts since the Willis and Ingle (2016) study. With the information available at this time, this study assumes that the identified performance pay school districts are the only districts in Ohio that have such provisions in their contracts. The State Employment Relations Board (SERB) website and individual district website were accessed to confirm that each district’s contract is classified correctly and five of the identified districts were subsequently dropped from this study. The Ohio School Boards Association and Buckeye association of School Administrators were also contacted, inquiring if there was knowledge of additional school districts that had implemented performance pay, with no additional districts added.

The number of school districts in this study is relatively small. Only eleven school districts have been identified as having a performance pay provision in their contracts. Twenty-two districts make up the comparison group. A more substantial number of school districts is preferable for this study, but that is impossible because the small number identified as having performance pay.
CHAPTER IV: ANALYSIS AND PRESENTATION OF DATA

This study examined the differences in student and teacher performance between performance-pay districts, and those who do not have performance-pay provisions in their teacher contracts. The chapter is organized according to the two specific research questions as it first reports the differences in student and teacher performance through \( t \)-tests of independent samples, and \( t \)-tests of paired samples with regard to district performance index scores, value-added scores, the percent of teachers rated proficient or above, and the percent of teachers rated accomplished. It then reports differences in student and teacher performance by type of performance-pay system.

Results

Research Question 1

Do Ohio school districts that have implemented a performance-pay system differ significantly from Ohio school districts that have not implemented a performance-pay system with regard to their performance index scores, value-added ratings, the percent of teachers rated proficient or higher, and the percent of teachers rated accomplished when using data from the Ohio School district report cards from the 2016-2017 school year?

A \( t \)-test of independent samples was utilized to examine this question. Table 5 displays the results of the \( t \)-test of independent samples, along with the mean and standard deviation for both the performance pay districts and the similar control districts. The performance index scores of performance pay districts were higher than the comparison groups, however this difference was not significant; \( t(33) = 0.14, p = .888 \), two-tailed. The value-added rating of performance pay districts was less than the comparison group but not significant; \( t(33) = -0.45, p = .656 \), two-tailed. In contrast, the performance pay districts reported a lower percent of
teachers rated proficient and accomplished and a higher percent of teachers rated accomplished than the control districts. Again, these differences were not significant.

Table 5

*t-Tests of Independent Samples Comparing Performance-Based Districts to Control Districts*

<table>
<thead>
<tr>
<th></th>
<th>Performance-Based Districts (n=11)</th>
<th>Control Districts (n=22)</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Performance Index Score</td>
<td>91.28</td>
<td>15.83</td>
<td>90.48</td>
<td>15.12</td>
</tr>
<tr>
<td>Value Added</td>
<td>-5.10</td>
<td>28.92</td>
<td>-1.50</td>
<td>17.15</td>
</tr>
<tr>
<td>% Teachers Skilled and</td>
<td>93.71</td>
<td>9.21</td>
<td>96.36</td>
<td>3.81</td>
</tr>
<tr>
<td>Accomplished</td>
<td>61.06</td>
<td>26.73</td>
<td>54.33</td>
<td>22.74</td>
</tr>
</tbody>
</table>

A *t*-test of paired samples was also utilized to examine this question. Table 6 displays the results of the *t*-of paired samples, along with the mean of difference and standard deviation of difference for the performance index score, value added ratings, and teacher evaluation ratings. A positive difference between performance-pay districts and control districts exists for both performance index scores and the percent of teachers rated accomplished, while there are negative differences between the mean for value-added ratings and the percent of teachers rated skilled or accomplished. The performance index scores do not significantly differ \( t(11) = 0.625, \ p = .546, \) two-tailed. The value-added ratings do not significantly differ \( t(11) = -0.676, \ p = .514, \) two-tailed. The percent of teachers rated skilled or accomplished do not significantly differ \( t(11) = -0.676, \ p = .356, \) two-tailed. The percent of teachers rated accomplished in performance pay districts do not significantly differ from the similar comparison group \( t(11) = -0.825, \ p = .429, \) two-tailed. In utilizing a *t*-test of paired samples, there are no significant differences between performance pay districts and similar districts with regards to student or teacher performance.

Table 6

*t-Tests of Paired Samples Comparing Performance-Based Districts to Control Districts*
<table>
<thead>
<tr>
<th>Performance Index Score</th>
<th>0.80</th>
<th>4.27</th>
<th>.625</th>
<th>.546</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value Added</td>
<td>-3.60</td>
<td>17.66</td>
<td>-.676</td>
<td>.514</td>
</tr>
<tr>
<td>% Teachers Skilled and Accomplished</td>
<td>-2.63</td>
<td>9.01</td>
<td>-.676</td>
<td>.356</td>
</tr>
<tr>
<td>% Teachers Accomplished</td>
<td>6.76</td>
<td>27.07</td>
<td>.825</td>
<td>.429</td>
</tr>
</tbody>
</table>

**Research Question 2**

Does the type of performance-pay system in Ohio school districts have different results with regard to a district’s performance index score, value-added ratings, the percent of teachers rated proficient or higher, and the percent of teachers rated accomplished when using data from the Ohio School district report cards from the 2016-2017 school year?

The performance pay districts utilize various methods for performance pay. Performance stipends link district performance of teacher attendance to a stipend a teacher may receive. In a performance schedule, teachers must meet specific performance indicators determined by each district to move up on the salary schedule. In a performance rate district, the teacher’s salary is recalculated each year based on specific performance criteria. There are also hybrid performance-pay systems where elements of multiple evaluation systems allow teachers to advance on the salary schedule. Licking Heights is identified as having performance stipends as their type of performance pay. Cleveland, Hudson City, Indian Hill, and Perrysburg all utilize a performance schedule for it type of performance pay system. Liberty Benton, Mentor, Oakwood, and Vanlue all have performance rates for their performance pay system. The only district that has been classified as having a hybrid performance-pay system is Cincinnati. Table 7 displays the types of performance-pay each district utilizes.

Table 7

*Performance-Pay Districts by Type*

<table>
<thead>
<tr>
<th>Performance Stipends</th>
<th>Performance Schedule</th>
<th>Performance Rates</th>
<th>Hybrid</th>
</tr>
</thead>
<tbody>
<tr>
<td>Licking Heights</td>
<td>Cleveland</td>
<td>Liberty Benton</td>
<td>Cincinnati</td>
</tr>
</tbody>
</table>
Student Performance. Individual performance pay district results for performance index scores and value-added ratings (student performance data) are compared to control districts in table 8 and figures 1 and 2. The first column identifies the type of performance pay that each district has. The second and third columns identify the performance index score of each district along with the mean performance index score of the of the corresponding control districts. Columns four and five provide the value-added scores for each district and the control mean.

Table 8
Performance-Pay District Student Performance Comparison to Control Districts

<table>
<thead>
<tr>
<th>Performance Pay Type</th>
<th>Performance Index Score 16-17</th>
<th>Valued Added Score 16-17</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Treatment</td>
<td>Control M</td>
</tr>
<tr>
<td>No Type Identified</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arlington</td>
<td>97.20</td>
<td>98.40</td>
</tr>
<tr>
<td>Performance Stipend</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Licking Heights</td>
<td>85.60</td>
<td>91.35</td>
</tr>
<tr>
<td>Performance Schedule</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cleveland</td>
<td>59.10</td>
<td>60.00</td>
</tr>
<tr>
<td>Hudson City</td>
<td>104.00</td>
<td>101.10</td>
</tr>
<tr>
<td>Indian Hill</td>
<td>108.80</td>
<td>103.30</td>
</tr>
<tr>
<td>Perrysburg</td>
<td>103.00</td>
<td>102.30</td>
</tr>
<tr>
<td>Performance Rate</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Liberty Benton</td>
<td>96.00</td>
<td>93.95</td>
</tr>
<tr>
<td>Mentor</td>
<td>94.40</td>
<td>90.95</td>
</tr>
<tr>
<td>Oakwood</td>
<td>108.40</td>
<td>104.45</td>
</tr>
<tr>
<td>Vanlue</td>
<td>80.60</td>
<td>84.25</td>
</tr>
<tr>
<td>Hybrid</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cincinnati</td>
<td>71.50</td>
<td>65.25</td>
</tr>
</tbody>
</table>

Arlington has a lower performance index score than its control mean and a higher value-added score compared to its control group. Licking Heights, a performance stipend district has a lower performance index score and a lower value-added score than its compare group.
Cincinnati, with a hybrid performance-based compensation system has a higher performance index score but lower value added scores than its control group.

Three (Hudson, Indian Hill, and Perrysburg) of the four performance schedule districts have a higher performance index score than the control mean for their districts. Also, three (Liberty Benton, Mentor, Oakwood) of the four performance rate districts have a higher performance index score than the control mean for their districts. Performance index score differences are displayed in Figure 1, and value-added differences are presented in Figure 2.

![Figure 1](image.png)

*Figure 1. Graph of performance index scores from 2016-17 school year by district with comparison group.*

In Figure 1, Hudson City, Indian Hill, and Perrysburg (performance schedule districts) all have lower performance index scores than their control groups. Liberty Benton, Mentor, and Oakwood (performance rate districts) all have a higher performance index score than their comparison districts. Licking Heights (performance stipends) has a lower performance index score than its control mean. Cincinnati (hybrid) has a higher performance index score than its
control mean. Since three of the four performance rate districts have higher performance index scores than their control means, this type of performance pay appears to have the most success at increasing student performance if you are measuring student performance based on performance index scores.

Figure 2. Graph of value-added scores from 2016-17 school year by district with comparison group.

Figure 2 is a graph of value-added scores, with the performance pay district scores represented by the bars, and their control mean represented by the line. According to Figure 2, two of the four performance rate districts (Mentor and Oakwood) have higher value-added ratings than their control groups. Two (Hudson and Perrysburg) of the performance schedule districts also have a higher value-added rating than their comparison districts, but two also have lower value added scores (Cleveland and Indian Hill). Cincinnati (hybrid classification), and Licking Heights (performance stipends) both have lower value-added ratings than their comparison districts. The hybrid and performance stipend districts have lower value-added ratings than their control groups. Both the performance rate and performance schedule district
types have an equal number of districts that higher and lower value-added ratings, showing no
differences when comparing based upon value-added scores.

**Teacher Performance.** Individual performance pay district results for the percent of
teachers rated skilled or above, and the percent of teachers rated accomplished (teacher
performance data) are compared to control districts in Table 9 and Figures 3 and 4.

Table 9

Performance-Pay District Teacher Performance Comparison to Control Districts

<table>
<thead>
<tr>
<th></th>
<th>% Teachers Proficient and Accomplished</th>
<th>% Teachers Accomplished</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Treatment</td>
<td>Control M</td>
</tr>
<tr>
<td>No Type Identified</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arlington</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Performance Stipend</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Licking Heights</td>
<td>70.90</td>
<td>97.60</td>
</tr>
<tr>
<td>Performance Schedule</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cleveland</td>
<td>86.50</td>
<td>89.00</td>
</tr>
<tr>
<td>Hudson City</td>
<td>99.60</td>
<td>97.80</td>
</tr>
<tr>
<td>Indian Hill</td>
<td>100.00</td>
<td>98.15</td>
</tr>
<tr>
<td>Perrysburg</td>
<td>99.00</td>
<td>95.85</td>
</tr>
<tr>
<td>Performance Rate</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Liberty Benton</td>
<td>98.90</td>
<td>94.70</td>
</tr>
<tr>
<td>Mentor</td>
<td>94.80</td>
<td>98.40</td>
</tr>
<tr>
<td>Oakwood</td>
<td>100.00</td>
<td>94.90</td>
</tr>
<tr>
<td>Vanlue</td>
<td>85.70</td>
<td>95.85</td>
</tr>
<tr>
<td>Hybrid</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cincinnati</td>
<td>95.40</td>
<td>94.15</td>
</tr>
</tbody>
</table>

Regarding the performance schedule districts, three (Hudson City, Indian Hill, and
Perrysburg) of the districts have a higher percentage of their teachers rated skilled or
accomplished than the control group, but only Hudson city has a higher percentage than their
control group for the percent of teachers rated accomplished, with Cleveland, Indian Hill and
Perrysburg having a lower percentage than their control group. All four performance rate
districts (Liberty Benton, Mentor, Oakwood, and Vanlue) have a higher percentage of their
teachers rated accomplished than their similar control group. Figure 3 displays the percent of
teachers rated skilled or accomplished when comparing performance pay districts to their control groups.

Figure 3. Graph of percent of teachers rated proficient or accomplished from 2016-17 school year by district with comparison group.

Figure 3 displays the percent of teachers rated proficient or accomplished with Performance pay district teacher percentages represented by the bars, and their control means are represented by the line. As displayed in figure 3, the performance schedule districts of Hudson City, Indian Hill, and Perrysburg all have a slightly higher percentage of their teachers rated skilled or above than the control group for their districts. Two of the four performance rate districts (Liberty Benton and Oakwood) have a higher percentage of teachers rated proficient or accomplished. Substantial differences are evident comparing Vanlue and Licking Heights to their control groups, with a lower percentage of teachers receiving at least a skilled rating. Except for Licking Heights and Vanlue, there is not wide variation between the performance pay
groups and their control means when comparing the percent of teachers rated skilled or accomplished.

Figure 4 compares the percentage of teachers rated accomplished in identified performance pay districts to their control groups. Performance pay districts are represented by the bars, and their control means are represented by the line. The performance rate districts of Liberty Benton, Mentor, Oakwood, and Vanlue all have a higher percentage of their teachers rated accomplished than their control group districts. Three (Cleveland, Indian Hill, and Perrysburg) of the four performance schedule districts have a lower percentage of their teachers rated accomplished than their comparison group. Hudson, Liberty Benton, Oakwood, and Cincinnati have a substantially higher percentage of their teachers rated accomplished than their control districts, while Licking Heights, Cleveland, and Perrysburg have a substantially lower percentage of their teachers rated accomplished than their control mean. There is much more
variation between districts and their control means when comparing just the percent of teachers rated accomplished than when one compares the two utilizing the percent of teachers rated skilled or accomplished.

**Summary**

This chapter presented the results of this study. The purpose the first research question was to determine if there are significant differences between the performance index scores, value-added ratings, and teacher evaluation ratings of districts that have performance based compensation systems, and those that do not. This study found no significant differences between performance pay districts and districts with a traditional pay scale. The purpose of the second research question was to determine if there are differences in district performance by the type of performance-based compensation system a district implements. A majority of performance schedule districts have a lower performance index score than comparison districts, and a majority of performance rate districts have a higher performance index score than their comparison districts. In a majority of performance schedule districts, a higher percentage of teachers are rated skilled or accomplished than the comparison group, but three of them also have a lower percentage of teachers rated accomplished. All four performance rate districts have a higher percentage of their teachers rated accomplished than the comparison group.

Chapter 5 summarizes and discusses these results further. Implications for action and recommendations for future research are also discussed.
CHAPTER V: DISCUSSION AND CONCLUSION

This chapter presents a discussion of results and conclusions for the study. The summary also presents an overview of the problem, restatement of the purpose statement and research questions, and review of the methodology. Major findings in this study are related to previous literature published. Unexpected findings that arose from this study are discussed. The chapter concludes with implications for action and recommendations for future research.

Overview of Study

About ninety-five percent of public school districts in the United States currently utilize a traditional pay scale (Pham et al., 2017) and some Ohio school districts have performance pay compensation systems (Willis & Ingle, 2016). Research has mixed results when evaluating the effects that these performance-based compensation systems have on student and teacher performance (Fullen, 2010). Ohio has implemented many changes in its education reporting system over the previous few years, including reporting student performance through performance index scores and value-added ratings; and reporting teacher performance through Ohio Teacher Evaluation System (OTES) ratings. With student and teacher performance data reported state-wide, school districts that have a performance-based compensation system were compared to similar districts that still utilize a traditional pay scale.

The purpose of this study was to determine if there are significant differences in student and teacher performance between Ohio districts that have performance-based pay systems and similar districts with a traditional pay scale. Further, the types of performance-based pay systems were investigated to determine if student and teacher performance differed by type of performance-based pay system. The research questions for this study are:
1. Do Ohio school districts that have implemented a performance-pay system differ significantly from Ohio school districts that have not implemented a performance-pay system with regard to their performance index scores, value-added ratings, the percent of teachers rated proficient or higher, and the percent of teachers rated accomplished when using data from the Ohio School district report cards from the 2016-2017 school year?

2. Does the type of performance-pay system in Ohio school districts have different results with regard to a district’s performance index score, value-added ratings, the percent of teachers rated proficient or higher, and the percent of teachers rated accomplished when using data from the Ohio School district report cards from the 2016-2017 school year?

This causal-comparative study compared performance-based pay districts to a purposeful sampling of similar districts with traditional pay scales. Willis and Ingle (2016) identified sixteen school districts in Ohio that have some form of performance pay in their teacher contract. After re-checking the teacher contracts on the State Employment Relations Board (SERB) website and accessing them on their websites, five districts were eliminated from this study, narrowing the group of performance pay districts down to eleven. Two comparison districts were identified for each performance pay district utilizing the similar district methodology tool on the Ohio Department of Education’s webpage. Performance index scores, value-added ratings, and teacher evaluation percentages were manually accessed through the interactive local report cards on the Ohio Department of Education’s website (education.ohio.gov) for the thirty-three districts. The scores and ratings were entered into a spreadsheet and then transferred to SPSS for data analysis. The results of this data analysis with a summary will be discussed in the following section.
Summary and Discussion of Results

No significant differences were found in student or teacher performance when analyzed using t-tests of independent and paired samples t-tests. However, noticeable differences were observed between types of performance pay. For example, three of the four performance schedule districts, and three of the four performance rate districts had higher performance index scores than their comparison group districts. Three of the four performance schedule districts have a higher percentage of teachers rated skilled or accomplished than the comparison group, but three of these districts also have a lower percentage of teachers rated accomplished. All four of the performance rate districts have a higher percentage of their teachers rated accomplished than the control group.

While Figlio and Kenny (2007) found a link between performance pay and increased student performance, this study was not able to discover a statistically significant relationship between the two. Hanushek and Lindseth (2009) advocate for performance pay systems because they may increase teacher motivation. Increased teacher motivation, in theory, should lead to improved performance index scores, higher value-added ratings, and better teacher evaluation ratings. This study did not display statistically significant differences between school districts that have performance pay and districts that do not have performance-based compensation with regard to student and teacher performance.

It is well documented that performance index scores are highly correlated with the socio-economic makeup the community of the school district (Bennet, 2010; Fleeter, 2016). With socioeconomic status being such a strong predictor of a school districts performance index score, the independent variable of having performance-pay policies does not have as strong of an impact on the performance index score. Based off of the results of research question two, some
types of performance pay may positively impact performance index scores with three of the four performance schedule districts having a higher performance index score than the mean of their two most similar peers. Improving student achievement scores is not something that occurs overnight, with it often taking more than three years after a reform effort is implemented see a noticeable improvement in student achievement. Other reform efforts may have a stronger effect on performance index scores, with will be further discussed in the recommendations section of this chapter.

The dependent variable of value-added scores was also used to measure student performance since there is no correlation between these scores and the demographics of the community (Selvage, 2013). Documentation exists displaying negative teacher attitudes towards the use of value-added ratings in teacher evaluations (Darling-Hammond, 2015; Thomas, 2014). Because of their objective nature, some still argue for the use of these scores in teacher evaluations (Hanushek & Rivken, 2010; Sass, 2014). Results from the first research question in this study found no significant differences with regards to value-added ratings between performance-pay districts and similar peers. Observable differences were also not displayed in value-added scores between types of performance pay.

Currently, in Ohio, a teacher’s evaluation rating calculated by assigning equal weight to the evaluators rating and the growth score. Teachers’ with a value-added score have the value-added score used in calculating their growth score. However, teachers do not get assigned a value-added rating if they did not have enough students taking a state test. An intervention specialist may not have enough students being tested to have a reliable value-added rating assigned. A physical education teacher is not teaching in a subject area that takes state tests, and also does not receive a value-added rating.
These teachers who do not receive a value-added rating still must have a growth score calculated into their final evaluation rating, and usually, these growth scores are determined locally through a student learning objective (SLO) process. In this process, teachers administer pre-tests, create growth targets for their students, then a district committee approves the growth targets, and at the end of the year, the teacher administers a post-test to determine if students met their growth target. While administering a pre-test on students to determine the knowledge students have entered a course is good practice, there are some issues with this process as it relates to using these scores to determine a teachers evaluation rating. Teachers have set low growth expectations for their students, with all students meeting these expectations, and having subpar teachers receive the highest growth ratings. Sometimes, the best teachers in a school may set high expectations for the growth of their students, and if these students just barely miss the growth target, then the teachers who set highest expectations for student learning may have the lowest growth scores, potentially leading to some of the lowest teacher evaluation ratings in a school.

Danielson (2007) and Marshall (2013) created rubrics to be used in teacher evaluation, and these rubrics were utilized in creating the Ohio Teacher Evaluation System. State law (ORC 3319.112) also requires the use of growth measures in the calculation of teacher evaluation ratings in Ohio. The results of this study display wide-variation of teacher evaluation ratings amongst different districts. Along with this wide-variation; districts with high value-added scores do not necessarily have a higher percentage of teachers rated accomplished since many teachers use student learning objectives and not value-added scores in the calculation of their growth scores that affect their evaluation rating. Value-added scores are not the same measure as SLO scores, and these differences lead to different results of teacher evaluation ratings. An
evaluator may be consistent in utilizing an evaluation rubric to assign a teacher evaluation rating, but this rating is not the same rating that appears on the district report card since different growth measures are also calculated into the final teacher evaluation rating.

Ballou (2001) points out that there is too little agreement on what the goals of education are, with the relationship between the actions of teachers and learning of students challenging to trace. The Ballou (2001) publication was written before OTES came into existence, and even research on teacher evaluation (Danielson, 2007; Marshall, 2013) was conducted. Perhaps the points made by Ballou (2001) are outdated. Many of the Ohio performance pay districts utilize OTES evaluations in determining teacher compensation. This study displays (albeit without statistical significance) that seven of the eight districts that have performance stipend or performance rate compensation systems display better teacher evaluation ratings than their comparison districts.

Performance-pay provisions have been classified into different typologies by both Springer (2009) and Willis and Ingle (2016). These classifications did not analyze the effectiveness of the different performance pay types. With such small number of participating districts, statistically significant differences were not analyzed in this research as well, but some interesting differences did appear when the performance index scores, value-added scores, and teacher evaluation ratings were graphed. For example, a majority of performance schedule districts have a lower performance index score than comparison districts, and a majority of performance rate districts have a higher performance index score than their comparison districts. In a majority of performance schedule districts, a higher percentage of teachers are rated skilled or accomplished than the comparison group, but three of them also have a lower percentage of teachers rated accomplished. All four performance rate districts have a higher percentage of
their teachers rated accomplished than the comparison group. With a salary increase calculated by a multiplying a percentage raise to one’s salary, performance rates appear to be the most motivating for employees of the types discussed. The smart employee, motivated by pay appears to be more motivated by rate increases than by stipends or moving steps on a schedule, as evidenced by all four performance rate districts having a higher percentage of teachers rated accomplished.

A few unexpected findings came as a result of conducting this study. First, while confirming that school districts that did have a performance-based compensation system, a noticeable majority of these contracts were affiliated with the American Federation of Teachers union, and not the National Education Association union. Why are their such differences between the stances of these two unions? It is documented that districts with performance-pay are spread geographically across the states, as well as socioeconomically (Willis & Ingle, 2016). The majority of Ohio school district teachers are represented by unions affiliated with the National Education Association, but the majority of teachers with performance pay are represented by unions affiliated with the American Federation of Teachers. Some of these performance-pay districts are in communities with high poverty (Cleveland, Cincinnati); other performance pay districts are in communities with many advantages (Hudson, Indian Hill, Oakwood). Because of the unique needs of both of these types of communities, are the teachers in these communities more likely to affiliate with the American Federation of Teachers? Are teachers more willing to try performance-pay in a high wealth district knowing that because of the socioeconomic makeup of the community, student and teacher performance will likely remain high, making performance-pay a less risk initiative? Are teachers in high poverty, urban
districts more willing to try performance pay since their achievement levels are already low, and any initiative that potentially leads to better performance is an improvement?

Second, although performance-based compensation is a relatively new phenomenon in Ohio (Willis & Ingle, 2016), an unexpected finding is that some school districts identified as having performance-based compensation systems appear to have scrapped their system and gone back to a traditional pay scale. Although there are many reasons why a district may have scrapped their performance-pay plans, this study did not investigate why Centerburg, Clear Fork Valley, Grandview Heights, Ripley Union Lewis Huntington, and Vandalia scrapped their performance-pay plans. Perhaps there was new leadership in these districts, or other efforts at school improvement were emphasized instead of performance pay provisions. Teachers may have never agreed with the performance-pay structures to begin with, and without teacher buy-in, any reform effort is doomed. The structure, type, or specific parts of their performance pay plans may have had errors or inconsistencies, leading to abandoning the performance pay structure in their contract.

Lastly, while accessing teacher performance ratings on the interactive local report card, the variability in teacher rankings across districts was unexpected, with some districts having a very high percentage of teachers ranked accomplished; other districts have higher rates of teachers ranked in the developing range. As discussed previously, teacher evaluation ratings are not only reliant upon the evaluators’ assigned ratings based upon observations, but fifty percent of the ratings are also dependent upon a teacher’s growth score. A teachers’ growth score could be the value-added score of that teacher, but often the growth score is based upon the results of a teacher’s student learning objective. Some of the districts with higher rates of teachers in the developing range may have a very strict SLO approval committee or teachers that set high
expectations of student learning. The districts with higher percentages of teachers rated accomplished do not necessarily have better teachers; rather this data displays that either their evaluator does not evaluate them as strictly based upon the rubric, or teachers in these districts have high growth ratings. If these ratings were only based upon value-added scores, these ratings would be consistent throughout the state, but they are also based upon results from SLO’s, which are determined by many teachers setting their own growth targets.

**Conclusions**

This chapter began with a brief review of the methods and purpose of this study. This study compared the school districts that have performance-pay systems to districts that have traditional teacher pay scales by student and teacher performance. Performance index scores, value-added ratings, and teacher evaluation ratings were downloaded from the Ohio Department of Education report cards and then transferred to SPSS for data analysis. No statistical differences were found between the two types of districts, but trends did emerge when analyzing different types of performance-based compensation systems. Throughout this research, a few unexpected findings emerged, which may lead to future investigation, including differences in teacher unions and their willingness to negotiate performance pay, variation in OTES ratings between school districts, and some districts appearing to have already scrapped their performance-based compensation plans.

This causal-comparative study found no statistically significant differences between Ohio school districts that have performance pay for teachers and those that do not. A reason for this is the number of identified performance pay districts is minimal, making it difficult to find statistical significance. There are many factors, other than performance pay, which have a more significant impact on performance index scores and value-added scores that the similar district
methodology attempted to account for, but even when using this methodology, significance was not found. The variables of performance index scores, value-added scores, and teacher evaluation ratings are easy to measure to compare districts since they are in numerical form, but perhaps, other variables would be better indicators of school performance. One can judge a school district by its performance index score, but a better picture can be formed by visiting the school, talking to people and observing the culture of the building. Are people welcoming? Caring? Supportive? Some people walk into buildings and have feelings of “This is a good place. Learning is occurring here, and people want to be here.” One can also walk into schools, and it feels like pink slime is on the walls. Why is there such negativity? Why are people not working together to help students? School districts and students are more than just test scores. A teacher’s performance is more than just what their rating says.

The theoretical framework of this study was based upon motivation research. With no significant differences found, one could conclude that performance pay is not motivating for teachers. However, three of the four districts with performance rates have a higher performance index score than their similar districts, and all four performance rate districts have a higher percentage of their teachers rated accomplished than their similar districts. Performance rates appear to be the most motivating type of performance pay because of how pay is calculated. Not only is their opportunity for a teacher to receive a higher percentage raise with a better evaluation, but unlike stipends, salary increases for the following years will be calculated off of this new salary. Salary and raises may motivate for the short-term because the money will help people meet basic needs (Maslow, 1970), sustaining motivation over a school year, or multiple school years will require more than just performance pay policies. Improving the culture of a school district by creating positive school climate can be a more effective motivational tool.
Creating a school where the opinions and voice of teachers are valued will improve the teacher’s motivation (Lawler, 1973; Vroom, 1964; Welbourne & Meija, 1995). When teachers can participate in the management of their school and make decisions, possibly about performance pay policies, their motivation is likely to increase.

The logic of “when one’s motivation increases, their performance will also improve” is the basis of this study. Measuring the motivation of an employee or their performance is more difficult. The Ohio teacher evaluation system was created to help measure, identify, and improve teacher performance. However, good teaching is about much more than two observations, multiple walkthroughs, and student test results. Relationships are critical to the success of a teacher, but relationships are much more difficult to measure. A teacher may connect and develop great rapport and professional relationships with a particular group of students, but struggle to connect with other students. A student may have had a bad day when they were tested. This same student could tell stories of how a teacher cared for him or her, provided hope on a bad day, and created a sense of community. The energy and passion of teachers are not measured under the Ohio teacher evaluation system or by student test scores but are very evident to students who sit in these teacher’s classrooms.

As leaders, one should be aware of the results of student test scores and be compliant with the requirements of the Ohio teacher evaluation system. Knowing the good teaching is much more than test scores or an evaluation rating, educational leaders need to be concerned with the whole child more than the picture that test results paint. The recently passed Every Student Succeeds Act (United States Congress, 2015) and the strategic plan for education in Ohio (Ohio Department of Education, 2018) are examples of looking at more than test scores. Ohio’s strategic plan encourages the development of the whole child and explores numerous
pathways for graduation. Federally, ESSA allows for states to design a new indicator other than test results to measure a schools’ success.

There are many recommendations based of the results of this study. The next section will begin with suggestions for school leaders, and then give specific suggestions for future research.

**Recommendations**

School districts are complex organizations (Gagne & Forest, 2008), so complex in fact that one cannot wave a magic wand of compensation reform and expect to see gains in student and teacher performance instantly. Although performance-based compensation is one strategy that policymakers and critics may advocate for when attempting to improve schools, a safer bet may be to emphasize the importance of instructional leadership. Whitaker (2012) stresses that it is people, not programs that will ultimately improve schools. Numerous authors have stressed the importance of improving school culture to improve schools (Connors, 2014; Whitaker, 2012). Along with working to improve a school’s culture, making data-based decisions (Boudett, City & Murnane, 2013; Dufour & Marzano, 2011; Goldring & Berends, 2009) is also an essential part of school leader’s responsibility.

An analysis of performance index scores and value-added ratings was conducted and results display no significant differences between districts that have performance-pay and those that don’t. School leaders should continue to analyze this data, searching for strengths, school-wide trends and areas for improvement and comparing school results to other similar districts to get an overall picture of student performance in the school. After this analysis, school leaders should collaborate with their teachers and celebrate what students did well and make plans to address identified weaknesses.
No significant differences with regard to teacher evaluation ratings between school districts that have performance-pay systems, and similar districts that do not were found. With that being said, the Ohio Teacher Evaluation System is a research-based system that when used correctly can be used for instructional improvement. Part of the OTES process involves teachers setting professional goals, which is consistent with goal setting theory discussed in the theoretical framework in chapter two. Principals who are aware of what is occurring in classrooms, and give meaningful feedback tied to a research-based evaluation rubric should see improvements in student and teacher performance.

With a lack of districts identified as having performance-pay systems in Ohio, this study did not find statistically significant differences between districts that have a form of performance pay and similar districts that do not. Based on appearance, the type of performance-pay system known as performance rates appears to perform better than other performance pay types. Three of these four districts had a higher performance index score than their similar peers, and three of the four had a higher percentage of their teachers rated accomplished than their similar peers. Districts considering implementing a performance-pay system would be wise to consider utilizing performance rates.

Different motivational theories relating to performance pay were discussed in the theoretical framework section of chapter two. Even without statistical significance, creating compensation systems based upon performance may motivate employees in a school. Maslow (1970) would argue that along with compensation, employees have love, belonging, and esteem needs that schools can help satisfy through improvement of their culture. If a school leader is worried about the motivation of its teachers, a good place to start is to create a sense of community, where teachers feel valued and have sense of belonging in the school district.
Districts may attach a reward for individuals meeting specific goals (Locke, 1968), or connect rewards to supporting the district enact its stated vision and objectives (Lawler, 1990; Welbourne & Meija, 1995). By utilizing the goal setting and contingency theories (discussed in chapter 2) in designing performance-pay provisions, districts are able to create systems that can be motivating for employees. Aligning pay to certain objectives enables employees to see what teaching qualities a district values. Having a clearly stated objective attached to a pay provisions gives the employee a clear vision of what they need to do to reach valued performance indicators.

Perrysburg’s process is an example of how involving teachers in the design of performance-based compensation systems works (Anstadt, 2017) because the employee had a stake in making the system successful (Vroom, 1976). Not only were these systems designed with employee input, but employees in Perrysburg had a choice of opting into the new system or having their pay calculated through a traditional pay scale model. The employees in Perrysburg not only had a voice in designing their performance pay system, but also had the choice to participate in it. The documented success of Perrysburg’s performance-pay system (Anstadt, 2017) is consistent with the motivational theories categorized as participative management (Lawler, 1993; Welbourne & Meija, 1995). Districts looking to implement Performance-Pay provisions should look to the Anstadt (2017) study, along with Perrysburg’s contract to see examples of properly designing and implementing a performance-pay system.

Throughout the process of data collection and analysis, many opportunities became evident for future research related to this study. The majority of these recommendations are related to the field of performance pay, but there are also opportunities to compare school districts in other ways, and further analyze student and teacher performance.
As previously mentioned, as contracts were checked to ensure they were categorized correctly as a performance pay or non-performance pay district, it appeared that a majority of the performance pay school districts were districts whose teachers' unions were affiliated with the American Federation of Teachers and not the National Education Association. More research should be done on the influence of teacher unions’ political agendas on the negotiation of teacher contracts, and major differences between the majority of agreements negotiated by these two unions. Included in this research should be a comparison of performance pay provisions in these contracts.

As noted in an earlier section some school districts (Centerburg, Clear Fork Valley, Grandview Heights, Ripley Union Lewis Huntington, and Vandalia Butler) appear to have abandoned the performance-pay provisions in their teacher contracts. Often, we can learn as much from our failures as from our successes, and case-studies should be conducted to investigate why performance-based pay provisions were not successful in these districts. These case studies may be educational for other districts when considering or designing alternatives to the traditional pay scale. Qualitative research should also be conducted to investigate why some districts were able to maintain performance-based pay systems, and others were not. This research can be done by interview key stakeholders in these districts including the superintendent, board of education members, the school treasurer, and key teacher union representatives. These interviews can then be transcribed and analyzed to draw conclusions.

The variability in OTES rankings was noted in the surprise section of this chapter. The Ohio Department of Education attempted to ensure reliability and validity of OTES rankings through a consistent training and credentialing process for OTES evaluators, yet there appears to be much variation in the ratings that districts assign to their teachers. The results of this study
display that three of the performance schedule districts have a lower percentage of teachers rated accomplished. Could the design of the compensation system in performance schedule districts lead to fewer teachers being ranked accomplished in these districts? In the performance rate districts, all four have a higher percentage of their teachers ranked accomplished. More research should be conducted to investigate why there is such variability in OTES rankings across districts. Districts with significantly higher and lower percentages of teachers in different OTES categories can be identified and then further research should be conducted to see what accounted for the variability in rankings.

A frustration in the design of this study was the constant change in student testing in Ohio. In a three year period, Ohio changed from administering Ohio Graduation Tests and Ohio Achievement Assessments to tests created by the Partnership for the Assessment of Readiness of College and Careers (PARCC) to tests developed by the American Institute of Research. Ohio also transitioned from paper and pencil testing to an online testing system. These changes made longitudinal studies on performance pay districts impractical. As the testing systems have appeared to have stabilized, longitudinal studies comparing student and teacher performance before and after transitioning to performance-pay system can be conducted when new districts decide to implement performance-pay.

There are many limitations with the methodology of this research and there are other approaches that will also work when analyzing performance pay. With the low number of districts identified as having performance-pay provisions in their contracts, finding statistically significant differences when comparing school districts will be rare. With such low numbers of participating districts, a researcher may want to utilize qualitative research, interviewing the perceptions of key stakeholders instead of conducting quantitative analysis. Along with low
numbers of participating districts, this study is also limited by the constant changing of learning standards and assessments. If the testing system in Ohio stabilizes, with students taking the same test multiple years in a row, longitudinal studies may be able to be designed to analyze the effectiveness of performance pay provisions. There are many other factors that could impact the performance index scores, value-added scores, and teacher evaluation ratings in a school district, and even though the similar district methodology attempts to account for these confounding variables, these factors potentially limited the results of this study.

A school district could still analyze the effectiveness of their performance pay provision utilizing $t$-tests to compare their district to their twenty most similar peers. With this new approach, the researcher will only be able to conclude that a significant difference exists between their district, their performance pay provision, and similar peers.

If more school districts implement performance-pay systems, research should be conducted utilizing the similar district methodology to compare student and teacher performance. The type of performance-pay systems can also be further evaluated for student and teacher performance if more districts adopt performance-pay systems. While the similar district methodology could be adapted if with a larger number of districts participating, school districts can also utilized the identification of similar districts to compare their own available data to the results their 20 most similar peers. Even if a school district does not have a performance-pay provision in their contract, comparing one’s district to its similar peers is a good way to identify the strong and weak performance areas of your district.

Anstadt (2017) conducted a program evaluation of Perrysburg’s performance-pay system. With knowledge gained during her research, Perrysburg was able to make revisions to their compensation systems in the spirit of continuous improvement. Internal stakeholders should
continue to evaluate the effectiveness of all of their programs, with the mixed method research
design of program evaluation being a helpful way to evaluate such programs. In the Anstadt
(2017) study, data was collected comparing the student test results of teachers who elected to be
compensated through their performance pay system to teachers who chose to remain on the
traditional pay scale. Teachers who chose to go with the performance pay system were also
interviewed to understand their perceptions of the new system. Similar types of research should
be conducted in the other performance-pay districts with a quantitative approach of comparing
available data, and the qualitative approach of interviewing key stakeholders.

The last recommendation for future research is to conduct qualitative research on the
perceptions of teachers, superintendents, treasurers, and boards of education in the 11 Ohio
school districts that still have a performance-based pay system. While this was a quantitative
study, deeper understanding of issues can be gained through qualitative research. Quantitative
research tends to determine whether there is a significant difference in variables, but qualitative
research tends to answer questions of how variables affect one another and the process of that
connects variables (Maxwell, 2013, p. 31). The attitudes of teachers on performance pay
(Goldhaber et al., 2011; Tenhiala & Lount, 2013) have been discussed in chapter 2, but the
perceptions of Ohio teachers have not been addressed. Superintendents, other key
administrators, school treasurers, and Board of Education members may also be able to provide
valuable insight from their experiences regarding performance-pay schemes in their districts.
These key stakeholders should be interviewed so that their experiences with performance pay
can be documented and shared with others interested in performance pay. These interviews can
then be transcribed and analyzed to understand similarities and differences between the
stakeholders and specific contracts with performance pay.
The last section of this chapter discussed recommendations including future research that should be conducted. With the relative newness of performance-based compensation of teachers in Ohio, there are still many questions to be answered. What are the perceptions of different stakeholders on performance-based compensation in Ohio? Why did some districts abandon their performance-based pay plan? If the testing situation stabilizes in Ohio, a longitudinal study should be conducted to see long-term effects of switching to a performance-pay system. If more school districts adopt performance-pay plans, similar studies can be conducted to examine student and teacher performance, possibly leading to statistically significant differences between types of compensation plans.
REFERENCES


doi:10.1257/0002828041464489


Ohio Department of Education.(2016a). Common questions about Ohio's value-added student growth measure. Retrieved from

Columbus, OH:


Vroom, V. (1976). Leader. In M. Dunnette (Ed.), *Handbook of industrial and organizational psychology* (pp. 1527--1551)


DATE: January 10, 2018

TO: Joseph Hoelzle
FROM: Bowling Green State University Institutional Review Board


SUBMISSION TYPE: New Project

ACTION: IRB APPROVAL NOT NEEDED

DECISION DATE: January 10, 2018

Thank you for your submission of New Project materials for this project. The Bowling Green State University Institutional Review Board has determined this project does not meet the definition of human subject research under the purview of the IRB according to federal regulations.

We encourage you to continue to confirm with the IRB whether future projects of this nature require review.

We will retain a copy of this correspondence for our records.

If you have any questions, please contact the Office of Research Compliance at 419-372-7716 or orc@bgsu.edu. Please include your project title and reference number in all correspondence regarding this project.

This letter has been electronically signed in accordance with all applicable regulations, and a copy is retained within Bowling Green State University Institutional Review Board's records.