

Changes in perspective regarding Washington State
Teacher Principal Evaluation Protocol (TPEP).

by

Michael Stephen Amend

B.S., Eastern Washington University, 2000
M.S., Washburn University, 2012

AN ABSTRACT OF A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF EDUCATION

Department of Educational Leadership
College of Education

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2023

Abstract

In this quantitative study, I surveyed building-level administrators in Washington State to see whether their perspectives regarding the Teacher-Principal Evaluation Program (TPEP) that was current in 2022 had changed since they had moved from classroom teaching to building administration. The research consisted of a survey with three primary parts. The first part consisted of questions to collect demographic information. The second part consisted of questions to collect information on current administrators' perspectives regarding TPEP. The final section asked administrators to remember their experience being evaluated as a teacher and to answer the same questions from a teacher's perspective. Specific perspectives examined included usefulness of the instrument for professional growth, impact on workload, impact on perception of career choice, and effectiveness of TPEP supports provided. Surveys were sent to all principals, assistant principals, and principal's assistants in Washington State. Exclusion questions were included at the beginning of the survey to eliminate participants who did not have experience with TPEP as both an administrator and a teacher.

I collected and analyzed data from each of the surveys to find descriptive statistics for each of the questions. Paired questions were then analyzed using a *t*-test to see whether any item had a statistically significant difference, using $p < .05$. Of the 24 questions analyzed, nine of those questions were found to have a statistically significant difference. Furthermore, of the nine questions showing a statistically significant difference, six of them were found in group H3, perception of workload. Of those six questions, all of them showed a more negative perspective from the administrator perspective than they did from the teacher perspective.

Changes in perspective regarding Washington State
Teacher Principal Evaluation Protocol (TPEP).

by

Michael Stephen Amend

B.S., Eastern Washington University, 2000
M.S., Washburn University, 2012

A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF EDUCATION

Department of Educational Leadership
College of Education

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2023

Approved by:

Major Professor
Dr. Jeff Zacharakis

Copyright

© Michael Stephen Amend 2023.

Abstract

In this quantitative study, I surveyed building-level administrators in Washington State to see whether their perspectives regarding the Teacher-Principal Evaluation Program (TPEP) that was current in 2022 had changed since they had moved from classroom teaching to building administration. The research consisted of a survey with three primary parts. The first part consisted of questions to collect demographic information. The second part consisted of questions to collect information on current administrators' perspectives regarding TPEP. The final section asked administrators to remember their experience being evaluated as a teacher and to answer the same questions from a teacher's perspective. Specific perspectives examined included usefulness of the instrument for professional growth, impact on workload, impact on perception of career choice, and effectiveness of TPEP supports provided. Surveys were sent to all principals, assistant principals, and principal's assistants in Washington State. Exclusion questions were included at the beginning of the survey to eliminate participants who did not have experience with TPEP as both an administrator and a teacher.

I collected and analyzed data from each of the surveys to find descriptive statistics for each of the questions. Paired questions were then analyzed using a t-test to see whether any item had a statistically significant difference, using $p < .05$. Of the 24 questions analyzed, nine of those questions were found to have a statistically significant difference. Furthermore, of the nine questions showing a statistically significant difference, six of them were found in group H3, perception of workload. Of those six questions, all of them showed a more negative perspective from the administrator perspective than they did from the teacher perspective.

Table of Contents

List of Figures	x
List of Tables	xi
Acknowledgements	xii
Dedication	xiii
Chapter 1 - Introduction.....	1
Background.....	3
Statement of the Problem.....	11
Purpose of the Study	13
Research Methodology	13
The Role of the Researcher.....	15
Assumptions.....	15
Limitations/Delimitations	16
Significance of the Study.....	16
Summary	17
Definition of Terms	17
Chapter 2 - Literature Review.....	21
The Theoretical Foundation of Teacher Evaluation	21
History of Teacher Evaluation	21
Purpose of Teacher Evaluation	25
History of the Washington State TPEP Initiative	26
Washington State ESSB (Engrossed Secondary Substitute Bill) 6696	27
TPEP Steering Committee	27
TPEP Pilot Project	28
Four-Tiered Evaluation System	29
Washington State ESSB 5895.....	30
Washington Teacher Evaluation Frameworks Research Base and Structure	31
The CEL5D+ Model	31
Marzano Framework for Teacher Effectiveness.....	42
Charlotte Danielson Framework	46

Impact of TPEP Initiative	48
Impact on Teacher Shortage	48
Impact on Administrator Shortage.....	50
Impact on Workload and Supports Provided.....	50
TPEP Usefulness for Professional Development.....	55
Summary.....	58
Chapter 3 - Research Methodology	59
Research Questions.....	59
Research Design	60
Sample Population	61
Procedures.....	62
IRB Approval and Assurance of Human Rights Protection	62
Pilot Test	63
Dissemination of the Survey	64
Data Collection and Analysis.....	65
Instrumentation	66
Qualifying Questions	67
Demographic Section.....	68
Administrative Perspective Section	68
Teacher Perspective Section	68
Summary.....	68
Chapter 4 - Data, Results, and Analysis	69
TPEP Survey Results	69
Demographic Information.....	69
Findings	73
Statistically Significant Differences.....	73
Group H1 – Support Provided	74
Group H2 – Opinion of Career Choice	75
Group H3 – Impact on Perceived Workload.....	78
Group H4-Usefulness for Professional Development.....	86
Power Analysis for Paired <i>t</i> -Tests.....	89

Other Findings – Administrator Perspective.....	90
Item 18 – Division of Labor Among Administrators.....	91
Item 20 – Types of Supports Provided for TPEP	93
Item 21 – Usefulness of TPEP Supports.....	94
Item 23 – Ability to Carry Out Various Aspects of Administrative Duties.....	95
Item 24 – Strategies for Workload Management.....	97
Item 29 – Desired Areas for Training	98
Item 30 – Helpfulness of Feedback for Performing Specific Duties	98
Other Findings – Teacher Perspective	100
Item 38 – Type of Supports Provided	100
Item 39 – Usefulness of TPEP Supports.....	101
Item 41 – Ability to Carry Out Aspects of TPEP	103
Item 42 – Strategies for Managing Workload, Teacher Perspective	105
Item 45 – Training Desired, Teacher Perspective.....	106
Summary	107
Chapter 5 - Discussion of Findings.....	108
Implications Based on Findings.....	109
Demographics –	109
H1 – Usefulness of Supports.....	110
Usefulness and Availability of Administrator Supports	111
Usefulness and Availability of Teacher Supports.....	112
H2 – Perception of Career Choice	114
H3 – TPEP’s Impact on Perceived Workload.....	117
H4 – TPEP’s Usefulness for Professional Development.....	122
Other Findings – Administrator Perspective.....	128
Item 18 – Division of Labor Among Administrators.....	128
Items 20 and 21 – Types of Support Provided for TPEP and Their Perceived Effectiveness	131
Item 23 – Ability to Carry Out Various Administrative Duties.....	138
Item 24 – Strategies for Workload Management	139
Item 29 – Desired Areas for Training	141

Item 30 – Helpfulness of TPEP Feedback for Performing Specific Duties.....	142
Other Findings – Teacher Perspective	145
Items 38 and 39 – Types of Support Provided for TPEP and Their Perceived Effectiveness	145
Item 41 – Ability to Carry Out Aspects of TPEP	148
Item 42 – Strategies for Managing Workload, Teacher Perspective	151
Item 45 – Training Desired, Teacher Perspective	152
Limitations of the Study	154
Recommendations for Further Studies	156
Summary	158
References	159
Appendix A - Survey Instrument.....	175
Appendix B - CEL 5D+ Framework.....	199
Appendix C - Marzano Framework	201
Appendix D - Danielson Framework.....	202
Appendix E - IRB Authorization	203

List of Figures

Figure 4.1. Distribution of Item H2-1	77
Figure 4.2. Distribution of Item H2-4.....	78
Figure 4.3. Distribution of Item H3-2 (TPEP’s Impact on My Ability to Address Issues of Student Discipline).....	80
Figure 4.4. Distribution of Item H3-3.....	81
Figure 4.5. Distribution of Item H3-4).....	82
Figure 4.6. Distribution of Item H3-5.....	83
Figure 4.7. Distribution of Item H3-6.....	84
Figure 4.8. Distribution of Item H3-7	85
Figure 4.9. Distribution of Item H4-2.....	89
Figure 4.10. Division of Evaluation Duties	91
Figure 4.11. Type and Number of TPEP Supports Provided.....	93
Figure 4.12. Usefulness of TPEP Supports Provided	94
Figure 4.13 Ability to Carry Out Aspects of TPEP	95
Figure 4.14. Strategies for Managing Workload	97
Figure 4.15. Usefulness of Feedback From TPEP	100
Figure 4.16. Type and Number of TPEP Supports Provided – Teacher Perspective	101
Figure 4.17. Usefulness of TPEP Supports Provided	102
Figure 4.18. <i>Ability to Carry Out Aspects of TPEP</i>	103
Figure 4.19. Strategies to Manage Workload, Teacher Perspective	105

No table of figures entries found.

List of Tables

Table 2.1. Views of School Leaders Who Report Receiving Adequate Supports to Work With Teachers on TPEP Compared With Those Who Did Not (Supported, n = 290; Not Supported, n = 152	52
Table 2.2. School Leaders' Perceptions of the Extent to Which Feedback From TPEP Informed Other Activities (n = 444)	56
Table 4.1. Survey Participants by Gender	71
Table 4.2. Survey Participants by Role.....	71
Table 4.3. Survey Participants by Ethnicity.....	71
Table 4.4. Survey Participants by Gender, Role, and Ethnicity	72
Table 4.5. Likert Scale for Paired Questions 19/37 and 33/48	73
Table 4.6. Likert Scale for Paired Question 25/43.....	74
Table 4.7. Likert Scale for Paired Questions 31/46, 32/47, 33/48.....	74
Table 4.8. Group H1 – Perceptions About Usefulness of TPEP Supports <i>Group H1 – Perceptions About Usefulness of TPEP Supports</i>	75
Table 4.9. Group H2 – Perceptions About Career Choice.....	76
Table 4.10. Group H3 – Perceptions About TPEP Impact on Perceived Workload	79
Table 4.11. Group H4 – Perceptions on TPEP's Usefulness for Professional Development.....	88
Table 4.12. Power and Effect Sizes for Items Rejecting the Null Hypothesis	90
Table 4.13. Questions 18.1–18.6	91
Table 5.1. Overall Mean and Standard Deviation of H1–H4	109

No table of figures entries found.

Acknowledgements

Any time a significant undertaking is completed, there are acknowledgements that need to be made. This one is no exception. First and foremost, I would like to acknowledge that none of this would have been possible without the grace and mercy of Jesus Christ. It is in Him and through Him that all things are accomplished. I would also like to thank my wife, Megan, and my son, Easton, for their love and patience through this entire process. I would also like to thank my parents for their wisdom, guidance, and encouragement through this process. When I have wanted to quit, they have offered help and been there cheering me on, and I am grateful.

I would also like to thank Dr. Jeff Zacharakis for his patience, direction, and consistency through the entire process. He picked me up on his doctoral student load when a colleague retired, took the time to learn my unique situation, and helped me navigate all the necessary pieces of the doctoral program. He has been encouraging, supportive, creative, and resourceful throughout the process and I would certainly not have finished without his help and guidance. I would also like to thank the members of my committee: Dr. Robert Hachiya, Dr. Cynthia Shuman, and Dr. Jerry Johnson. Thank you for taking on this project with me. I appreciate your feedback and am thankful to have had the opportunity to work with you. I would also like to thank Dr. Aaron Hill from the University of Florida for his insight into my statistical problems. Hobus, you were a lifesaver. Also, to my editor, Joanie Eppinga, thank you for encouraging me and bringing out the best in my writing!

Finally, I would like to thank all my administrative colleagues who have contributed to my practice, all the teachers who have made me a better administrator, and all the students who have taught me so many lessons. You all have blessed me with joy throughout the process, and I am thankful to have had the opportunity to engage with each of you.

Dedication

This work is dedicated to my family. Megan and E, this is what it looks like to take on a big project and finish. It's what we do. To the Amend and Coco clans, your love and support have helped me persevere when I have not wanted to. This work is also dedicated to all the educators who have given their time, talents, and treasures to the staff and students they are charged to serve.

Chapter 1 - Introduction

Rhetoric from current, and former, policy makers and educational leaders suggests that American education used to be the envy of every nation (Strauss, 2012). A deeper look at the data suggests this is no longer the case. According to Dr. David Drew (2011), historical data dating back to the beginning of the use of computers post World War II suggest that the United States has not, in modern times, been a leader in academic achievement, especially in the areas of math and science. Although data indicate that eighth grade students fare comparatively well—ninth out of 45 nations studied—older students “struggled to apply what they had learned in higher grades” (Drew, 2011, p. 28). At the time of publication, 2011, American students had never finished higher than 12th in math and ninth in science (Drew, 2011). Whereas U.S. News and World Report ranks the U.S. education system first among nations (“Best Countries for Education,” n.d.), the World Top 20 Project (which rates nations on factors ranging from early child education to college and career readiness) did not even have the United States rated in the top 20 at the end of 2020 (World Top 20 Project, 2020).

Since the inception of our democracy in 1776, an educated voting populace has been essential to the proper functioning of our government (Jefferson, 1816). According to former NEA president Dennis Van Roekel, the publication of *A Nation at Risk* in 1983 catapulted public education back to the top of the national agenda and became a catalyst for today’s so-called “reform” movement (Graham, 2013). School reform movements have taken many forms, including The No Child Left Behind Act (2002) and the Every Student Succeeds Act (2015). Within the framework of these acts of legislation, states have been given latitude to effect change locally. In this vein, Washington State adopted Engrossed Secondary Senate Bill (ESSB) 6696 (2010). This legislation established the Teacher Principal Evaluation Protocol (TPEP) evaluation

framework and steering committee to oversee the development and oversight of Washington's TPEP program (TPEP, 2013).

Since its inception in 2010, the TPEP steering committee has been charged with the establishment and maintenance of a statewide evaluation protocol. This protocol is intended to aid educators in continuing their professional development, continually building their capacity, and increasing the quality of education taking place in schools (S. Anderson, personal communication, November 25, 2018).

The TPEP framework places different demands on teachers and on administrators. As a result, teachers and administrators have different perspectives regarding the effectiveness and sustainability of TPEP as it currently exists (Elfers & Plecki, 2017). In this study, I explored how educators' perspectives change as they transition from classroom teacher to administrator. My goal in conducting this study was to gain insight into shifts in administrator perspective, including why 14% of principals and 7% of assistant principals strongly agree that TPEP has made them consider leaving their positions as school leaders. I also sought answers to questions about these administrators' perceived levels of support for TPEP implementation, the usefulness of TPEP for professional development, and TPEP's impact on educator workload (Elfers & Plecki, 2017). With this knowledge, administrator preparation programs, regional educational service districts (ESDs), and local school districts can seek to develop programs to help administrators transition more smoothly to the role of evaluator and address concerns in the above-listed areas in efforts to promote longevity in the profession.

This chapter begins with an overview of the context and background of this study. Following the background are a problem statement, a statement of purpose, and relevant research questions. This introductory chapter also includes explanations of the research methodology, the

role of the researcher, assumptions made by the researcher, and limitations/delimitations of the study. The chapter concludes with a discussion of the rationale behind, and the significance of, this research, along with some key educational terms relevant to this study.

Background

In August of 1981, Secretary of Education T. H. Bell created the National Commission on Excellence in Education (NCEE) with the express purpose of offering “a report on the quality of education in America ... to the American people by April of 1983” (Gardner, 1983, p. i). This commission was created in response to “the widespread public perception that something is seriously remiss in our educational system” (Gardner, 1983, p. 1). The NCEE was tasked with:

- assessing the quality of teaching and learning in our Nation’s public and private schools, colleges, and universities
- comparing American schools and colleges with those of other advanced nations
- studying the relationship between college admissions requirements and student achievement in high school
- identifying educational programs which result in notable student success in college
- assessing the degree to which major social and educational changes in the last quarter century have affected student achievement
- defining problems which must be faced and overcome if we are to successfully pursue the course of excellence in education. (p. 1)

When the NCEE presented its report to then Secretary of Education Bell, the report’s authors “lambasted the state of America’s schools and called for a host of much needed reforms to right the alarming direction that public education was seen to be headed” (Graham, 2013, p. 1). A Nation at Risk made clear the urgency of attaining a world-class education. The report

did this by underscoring the need to be able to learn and apply skilled intelligence as the new “raw materials” of the “information age” (Gardner, 1983, p. 2). It also gave many indicators of high risk, including international comparisons of student achievement, declining literacy rates in both adults and teens, declining standardized test achievement, lack of higher order intellectual skills, declines in science achievement, increase in remedial college courses, and complaints about lack of basic skills in the incoming labor force made by business and military leaders (Gardner, 1983). The final conclusion of the report’s authors when faced with all of these pieces of evidence was: “For the first time in the history of our country, the educational skills of one generation will not surpass, will not equal, will not even approach those of their parents” (Gardner, 1983, p 6).

The causes of this decline were clearly outlined in the report and broken down into four categories: content, expectations, time, and teaching. Deficiencies in content sprung from ineffective curricular choices and offerings; curriculum was watered down and purposeless. Students elected to take courses that were not rigorous enough to adequately prepare them for future opportunities rather than more rigorous options (Gardner, 1983).

Deficiencies in expectations were cited as a major source of educational decline as well (Gardner, 1983). These deficiencies took the form of decreases in homework and lack of rigorous course requirements for many students. At the middle school and high school levels, the expectation deficiencies included students’ taking more elective classes and competency exams that were not rigorous enough. At the college level, lackadaisical admission standards and a “writing down” of textbooks that failed to challenge students were identified as major culprits (Gardner, 1983, pp. 12–13).

Lack of instructional time was also cited as a cause for the decline in American education. In its findings, The Commission noted discrepancies in length of school year and day between the United States and other industrialized countries. For example, in England, students spend 8 hours a day and 220 days a year in school, whereas the United States requires only 6 hours a day and 180 days a year. In addition, the Commission noted that school time devoted to tasks that were not academic, such as cooking and driving, offered the same credits toward graduation as academic tasks such as reading, writing, math, and science. Lack of instruction in the area of study skills was identified as an issue as well. The Commission also noted that time on task was an issue for students. Time on task decreased because of poor classroom management, resulting in even greater decreases in academic learning time (Gardner, 1983).

The final area that the Commission targeted as a source of the steady decline of education in America was teaching. The Commission found that teachers were being drawn from the “bottom of the academic barrel” and were “not well versed enough in key content areas” (Gardner, 1983, p. 14). The Commission also cited low salaries, perceived teacher surpluses, and science and math teacher shortages as teacher-related contributions to the decline in education (Gardner, 1983, p. 14).

Since the release of *A Nation at Risk* (1983) there has been much debate regarding the validity of the report and its usefulness for educational reform. The debate surrounding best practices in education has led to many states taking significant steps to attempt to address the report’s mandate to increase the rigor of the standards for learning (Graham, 2013). These steps have taken myriad forms. All states have adopted some sort of state standards, and 45 states have adopted the Common Core State Standards (CCSS), helping to create common expectations for students at all different levels of education to help ensure their future success (Graham,

2013). In addition, partnerships have been formed between private industry, educational experts, and legislators to identify key areas for reform and determine plans of action to improve these identified areas (Graham, 2013).

In addition to the creation of standards, many other educational improvement movements have come about. These movements included, but are not limited to; the effective schools movement; the accelerated schools movement, which began in 1986 and attempted to bring all students into the academic mainstream by offering special education students instruction originally intended for gifted and talented students (Bloom et al., 2001); and the school-within-a-school movement, which was born out of the book *Big School, Small School* (Barker & Gump, 1964) and attempted to blend the benefits of a big school and those of a small school to get the best of both worlds.

The debate around best practices in education also spurred on several acts of legislation including; the establishment of national goals (termed “Goals 2000”) in 1994, the Improving America’s Schools Act (1994), the No Child Left Behind Act (2002), and the Every Student Succeeds Act (ESSA, 2015). Although some of these acts saw some measure of success, the main tenets of success have not been steadily addressed. Former NEA president Dennis Von Roedel summed it up well as quoted in a 2013 article written by Edward Graham called “A Nation at Risk Turns 30”:

In some ways A Nation at Risk was helpful, but what we’ve seen in the past 30 years are too many misguided efforts—most notably No Child Left Behind—that have only compounded the problems the report identified ... Educators across the country work hard to give their students the great education they deserve, but lawmakers cannot keep pulling the rug out from under them with bad ideas. We need to do what we know works. We need to fully fund our

schools, invest in early childhood education, increase parental involvement and keep our class sizes small, especially in high-poverty schools at the lower grades. This is how we'll make our schools work for every student. (p. 4)

In addition to nationwide legislation, a nationwide test (the National Assessment for Educational Progress or NAEP) was developed in 1969. The NAEP assessment was mandated for schools receiving any Title I funding beginning in 2001 (NAEP, 2021). Data from these tests are used to create the nation's report card that has been published yearly since 1996 as a measure of educational progress (NAEP, 2021; Smith, 2005). The NAEP is administered through the Department of Education (DOE) and has shown mixed indicators of success on a nation-wide scale. Dr. George Bohrnstedt (2013), from the American Institute for Research (AIR), stated:

The NAEP long-term data trend for 17-year-olds shows that both their scores in reading and math have been essentially flat since the mid 1980s. The reading scores of 13-year-olds grew only slightly during these decades, but their mathematics scores grew more impressively. The greatest growth has been in 4th graders' mathematics performance. Gains in their reading results have been more modest, though better than those for 13- or 17- year olds. (p. 2)

Since Borhnstedt's 2013 article was published, there have been continued yearly rounds of NAEP testing that show more of the same unimpressive growth. Bohrnstedt offered an interesting explanation for this flat NAEP data. Bohrnstedt contended that these lines are flat because of changes in demographics and the achievement/opportunity gaps that exist between those demographics. For example, students who have traditionally performed worse on standardized assessments, including minority groups and English language learners (ELLS), have increased in population and now make up a larger portion of the NAEP data collected.

These minority groups are making gains, but their gains, as well as the gains of nearly all other demographics, are offset by the increased number of students in these traditionally lower-performing demographics. In terms of overall NAEP data, students making up the minority demographics continue to account for an increasing percentage of the student population, so the overall NAEP scores stay relatively flat while all groups are actually showing growth (Bohrnstedt, 2013).

Given these achievement/opportunity gaps, Washington State has taken specific measures to help ensure equity of opportunity for every student through the development of a comprehensive four-tiered teacher and principal evaluation program (TPEP). TPEP was born out of Washington State ESSB 6696, passed in 2010, which requires schools to use one of three established frameworks to evaluate teachers and administrators. Each educator evaluated using the TPEP framework will receive a rating of *unsatisfactory*, *basic*, *proficient*, or *distinguished*, along with detailed feedback on their performance based on specific criteria and elements. The criteria and elements will be described further in the literature review section. After the passage of ESSB 6696, the state commissioned the TPEP steering committee, made up of representatives from OSPI, the Washington State Governor's Office, the Washington State Education Association (WEA), the Association of Washington School Principals (AWSP), the Washington State Parent Teacher Association (WPTA), and the Washington State School Directors Association (Washington State Teacher/Principal Evaluation Program, n.d.). The TPEP steering committee then established core principles, key evaluation criteria, summative evaluation ratings, student growth impact ratings, approved evaluation frameworks for both teachers and administrators, and an implementation calendar (TPEP, 2013). The pilot for the TPEP program began during the 2011–12 school year and culminated with all certificated teachers and

principals being evaluated by the system during the 2015–16 school year. The state of Washington is no longer in the pilot stage and has now fully implemented the three adopted TPEP frameworks (TPEP, 2013).

Since full implementation, Washington State has produced a report called the *Analysis of School Employees Evaluation Survey* (SEES) that investigates state trends in evaluation scores for teachers. Demographics are broken down by district size, type of evaluation (comprehensive or focused), teacher rating received, percentage of students of color, and free and reduced lunch qualification. Districts are also asked to identify the electronic platform they use to collect the evaluation data and how they use the evaluation results. These data are analyzed and submitted to the OSPI yearly as part of state and federal reporting requirements (Elfers & Archberger, 2017). These data can be used to inform educational decisions at the local, state, and federal levels and to appropriately allocate resources to best serve Washington’s students.

Currently, over 90% of teachers in the state of Washington receive TPEP scores of *proficient*, level 3, or *distinguished*, level 4 (Elfers & Plecki, 2017). This result indicates that over 90% of teachers in Washington State are quality teachers who can adequately provide ample learning opportunities for their students. Data from the beginning teacher retention survey by Plecki, Elfers, and Van Windekens (2017) indicates that the number of first-year teachers has increased nearly 80% from 2010 through 2016, leaving almost 12% of teachers nationally in either their first or second year of teaching. This assessment corresponds with needs to accommodate increased numbers of students and smaller class sizes (Plecki, Elfers, & Van Windekens, 2017). Of these beginning teachers, 12% leave the profession each year. For beginning teachers, the TPEP evaluation protocol is critical, as it can provide the support needed for these teachers to experience success in their first years in the profession. Programs such as

the Beginning Educator Support Team (BEST) help teachers integrate and have been shown to be effective in stemming the tide of beginning teachers leaving the profession (Plecki, Elfers, & Van Windekens, 2017). These efforts, coupled with effective implementation of the TPEP framework, can help retain quality teachers for the life of their professional careers. Similar trends can be seen with administrators. As Plecki, Elfers, and Wills (2017) documented in their study on principal retention and mobility, administrator attrition rates in Washington are, at 6%, lower than the nationwide trends of 12%. That being said, 6% of Washington principals still exit the profession each year. Survey results indicate that the increased evaluation demands of TPEP may contribute to this phenomenon (Plecki, Elfers, & Wills, 2017).

Implementing a new and complex initiative like TPEP is an expensive proposition, and it requires extensive training. The legislature appropriated \$10 million in 2013 and then \$5 million each of the next 2 years, through grants, to provide training for teachers and administrators throughout the state. This happened through ESDs, district-level training, and in-building trainers (Elfers & Plecki, 2017). School leaders who have a deep knowledge of the framework can use it more efficiently and can better serve beginning teachers, helping both administrators and teachers to be successful, which can lead to greater retention (Elfers & Plecki, 2017).

Because administrators play a role in retaining teaching staff, it is critical for schools to retain quality administrators. In conducting this study, my purpose was to identify how administrators' perspectives on TPEP shift as they move from the classroom to administration. In addition, I sought to find ways to better support those administrators who are in difficult evaluation circumstances—in particular, those in small schools where they are the sole evaluator for an entire building (Elfers & Plecki, 2017). Administrators' perceptions of the current TPEP

framework and of its effectiveness as an evaluation instrument will be expanded upon further in the literature review in Chapter 2.

Statement of the Problem

Since *A Nation at Risk* (1984) identified four main deficiencies in the American education system, many potential causes have been given significant research attention. These include school size (Baker & Gump, 1967; Cotton, 1996); school day format (Gruber & Onwuegbuzie, 2001; Hurley, 1997; Matazzaro, 1999; McGorry & McGorry, 1998; Mistretta & Polansky, 1997); curriculum and content (Kanter & Konstantopoulos, 2010; Tarr et al., 2010); staff accountability (Ballard & Bates, 2008; Bullough et al., 2003; Linn, 2003,); and course requirements (Lee et al., 1997; Lillard & DeCicca, 2001; Sipple & Killeen, 2004). The dual themes of curriculum and content and staff accountability were the focus of this investigation, as they are the themes most directly addressed by the TPEP evaluation framework.

The debate around the perceived decreasing performance of American schools has caused educators and legislators alike to begin the work of trying to remedy what was deemed a national tragedy (Gardner, 1983). To help address the concerns identified by those concerned about the performance of the American Education System, Washington State adopted legislation intended to increase the rigor of teacher evaluation. The TPEP Evaluation Initiative was meant to to create a collaborative evaluation framework that moved away from the typical two-tiered system (satisfactory/unsatisfactory) to a four-tiered evaluation system that rates teachers on a scale from Not Using, scored numerically as a 1, to Innovating, scored numerically as a 4. In the TPEP frameworks, specific elements address each of the four deficiencies—content, expectations, time, and scheduling—called out by *A Nation at Risk*. Concrete examples include the following:

- Domain 1-Curriculum and Pedagogy: Addresses content and expectations deficiency (UWCEL, 2012).
- Domain 1-Classroom environment: Addresses expectations and time management deficiency (Danielson, 2013).
- Domain 5-Student Impact: Addresses the aggregate of all four deficiencies by looking directly at student performance (Marzano, 2013).

Each of the elements in the framework can be used by evaluators to offer specific feedback to improve instructional practice. Large-scale issues of scheduling cannot be addressed by classroom-level observation, but they can be effectively addressed at the building level. To this end, the AWSP framework, which is the administrative equivalent of the TPEP framework, addresses issues of scheduling with Criterion Six, which targets the management of all school resources (AWSP, 2014).

The frameworks effectively target the issues identified as negatively impacting student learning for improvement. The issue that I sought to address with this study is not whether the frameworks address the identified issues that will improve student achievement. I am investigating changes in educator perspective regarding TPEP's impact on their career choice, the type and effectiveness of supports provided, their perceived workload, and TPEP's usefulness for professional development.

Since the TPEP frameworks have been implemented, Washington State OSPI produces an Analysis of School Employees Evaluation Survey (SEES; Elfers & Archberger, 2017). This report identifies different trends in evaluation and functional uses for TPEP evaluation data. The report is also used by the state to gauge the impact of the TPEP initiative. In addition, Elfers and Plecki (2017) released a report that outlined teacher and administrator perceptions of TPEP

(Elfers & Plecki, 2017). This report indicated that teachers and administrators had many similar feelings and perceptions about TPEP. Elfers and Plecki's report also indicated that there were significant differences between how teachers and administrators perceived, and experienced, the TPEP frameworks (Elfers & Plecki, 2017). Currently, these differences in perception may seem minor; but projected into the future, these differences in perception may pose significant barriers to the continued implementation of the TPEP evaluation system. According to the survey, barriers such as lack of support, need for continuous professional development, and management of workload for both teachers and administrators all need to be addressed as the program moves forward (Elfers & Plecki, 2017). Without resolution, these issues could result in greater attrition of teachers and administrators, causing teacher and administrator shortages.

Purpose of the Study

My purpose in conducting this study was to examine changes in perspective regarding the TPEP evaluation protocol among school district employees who have gone from being evaluated on TPEP as a teacher to doing evaluation of teachers as an administrator. Perspectives that will be investigated include the impact of TPEP on perception of career choice, the impact of TPEP on teachers'/administrators' perceived workload, the degree and type of support received, and TPEP's usefulness in aiding professional growth.

Research Methodology

A 2017 survey developed by researchers at the University of Washington that outlines district-level administration, building-level administration, and teacher perceptions of key elements of the TPEP evaluation framework was adapted to measure for changes in perception with job changes (See Appendix A). The primary research question for this investigation was:

Are there any significant statistical differences between educator perceptions of TPEP as teacher and as administrator? This question was further broken down into four sub questions:

- Is there a significant statistical difference between educator perspectives on TPEP regarding the support provided as a teacher and as an administrator?
- Is there a significant statistical difference between educators' perspective on TPEP regarding their opinion of their career choice and career mobility as a teacher and as an administrator?
- Is there a significant statistical difference between educators' perspective on TPEP regarding their perceived workload as a teacher and as an administrator?
- Is there a significant statistical difference between educators' perspective on TPEP regarding its usefulness for professional growth as a teacher and as an administrator?

With this research, I sought to discover whether there are any statistically significant differences in educator perspectives as they transition from teacher to administrator.

The survey asks current administrators, who were evaluated using the TPEP framework, to recall their perceptions of TPEP when they were classroom teachers and to share their current perceptions as administrators. A total population sampling method was used to survey all administrators in Washington State. Deans of students and athletic directors were not surveyed because they do not use the framework except in isolated circumstances. I analyzed survey data to see whether there were significant statistical differences in perception for individuals as they transitioned from teacher to administrator. This survey was distributed between August 1 and January 1st, 2022. Microsoft Excel was used to conduct an analysis of the survey data. Further information about this research study is contained in Chapter 3.

The Role of the Researcher

The state of Washington has done a great deal of work to measure the perception of TPEP from differing vantage points. These vantage points include those of district-level administration, building-level administration, and classroom teachers. Current research indicates that there are several key areas where teachers and administrators differ in their perspectives on the strengths and weaknesses of the TPEP framework (Elfers & Plecki, 2017). As a teacher, I have firsthand knowledge of the work needed to meet the rigorous requirements of the TPEP protocol. Having taught in the state of Washington from 2014–16, I was evaluated by building administration on the Danielson Evaluation Framework. I also have firsthand knowledge of evaluation using the TPEP framework and of the demands it places on administrators. I spent 4 years evaluating a variety of staff members in a large district with the Marzano Evaluation Framework as a principal's assistant. I also spent two years as the building principal of a small K–12 school again conducting evaluations using the Marzano Framework. Currently, there seems to be a void in the research regarding changes in perspective for educators who transition from the classroom to administration. My goal in conducting this study was to add to the body of research by surveying administrators to see whether, and in what ways, their perspective on TPEP had changed as their role changed.

Assumptions

The following assumptions were established for this study:

1. Participants would answer questions honestly.
2. Participants would be able to remember their perceptions of TPEP in a teacher capacity.
3. No other variables existed that would have a major outcome on this research.

4. The methodology would not adversely impact the outcome of the study.

Limitations/Delimitations

This study was subject to the following limitations:

1. This study was limited to administrators practicing in the state of Washington.

Because of the specific nature of the statewide evaluation protocols, caution should be taken in generalizing these results outside of Washington State without replicating the study locally.

2. Many administrators in Washington State had not both been evaluated using TPEP and evaluated teachers using TPEP. This deficiency limited the sample size and was the reason for using the total sampling method rather than a targeted sample method.

3. The results of this study may have been affected by the ability of participants to accurately remember their perceptions of TPEP when they were teachers. This effect could increase with participants' length of time out of the classroom.

Significance of the Study

Education is a rapidly changing field. The value of highly trained and highly competent administrators and teachers goes without saying. Michael Fullan (2014) underscored this value when he identified “professional capital” as the “first key” to someone’s maximizing their impact as a principal (p. 70). Fullan called professional capital a “function of the interaction of three components: human capital, social capital, and decisional capital” (p. 70). Professional capital, and all its components, takes time to build. In this study, I sought to identify the impact of our current evaluation system on the longevity of both teachers and administrators in their current positions. A survey of teachers and administrators published in July of 2017 indicates that 31% of Washington teachers agree that TPEP made them either somewhat or strongly

consider leaving the teaching profession (Elfers & Plecki, 2017). Survey data also indicates that 39% of principals and 16% of assistant principals shared the same sentiment (Elfers & Plecki, 2017). While this increased desire to leave the profession may be a TPEP related issue, it may also be related to the respondents' place in their career or an accumulation of other issues. Leithwood et al. (2004) asserted that "the principal is second only to the teacher in terms of impact on student learning" (p. 5). If we have an evaluation system that may be contributing to 31% of our teachers and 39% of our principals considering leaving the profession, there is certainly cause for further investigation into why respondents feel this way.

Summary

This first chapter has provided an overview of the study outlined above. I undertook this study because of personal experience with ineffective teacher evaluation practices and a lack of valid professional development feedback from evaluators using the standard two-tiered evaluation systems. The purpose behind this research was to examine how educator perspectives on the Washington State TPEP Evaluation Framework have changed as the educators transitioned from being evaluated to doing the evaluating. Data were analyzed using Microsoft Excel. Chapter 2 contains an investigation of aspects of the literature concerning the history of educational evaluation and reform as it pertains to Washington State's current evaluation system and the impacts that system is having on teacher and administrator longevity.

Definition of Terms

Achievement Gap: In education, achievement gap refers to the disparity in academic performance between groups of students of different demographics (Coleman et al., 1966).

ASCD (Association for Supervision and Curriculum Development): ASCD "empowers educators to achieve excellence in learning, teaching, and leading so that every child is healthy,

safe, engaged, supported, and challenged” (ASCD, 2020, para. 1). ASCD is made up of over 113,573 members from more than 129 countries and 64 affiliate organizations (ASCD, 2021).

AWSP (Association of Washington School Principals): AWSP exists to “support principals and the principalship in the education of all students” (AWSP, n.d., para. 1). The association does this by advocating for members, offering professional learning, offering coaching and mentoring, providing legal advice, promoting diversity and equity, and communicating relevant news to all administrators across the state (AWSP, 2021).

CCSS (Common Core State Standards): The CCSS are “a set of high-quality academic standards in mathematics and English language arts/literacy (ELA)” (DJUSD, n.d., para. 1) and they “outline what a student should know and be able to do at the end of each grade” (para. 1). Begun in 2009, the final draft was presented to states in June of 2010. The CCSS are now used by over 40 states (para. 1).

ESD (educational service district): Washington State is broken up into nine different regions, each served by an independent service district that provides services to all districts within the region (OSPI, n.d.b).

ESSB (Engrossed Substitute Senate Bill) 5895: This legislation was passed to mandate parameters around teacher and principal evaluation, including adopting preferred frameworks, requiring student growth goal data to be used in evaluations, establishing an implementation timeline, defining what is satisfactory and unsatisfactory, outlining annual requirements for focused and comprehensive evaluations, requiring the use of evaluation in HR decisions, establishment of professional development programs to support teacher growth, and requirements to move from provisional to continuing teacher contracts (ESSB 5895, 2012).

ESSB 6696: This legislation was passed to help enact education reform by mandating a common state evaluation framework for all teachers and administrators (ESSB 6696, 2010).

NAEP (National Assessment for Educational Progress): The NAEP “is the largest nationally representative and continuing assessment of what America’s students know and can do in various subject areas” and is required of all schools receiving any Title I funding (NAEP, 2021).

NCEE (National Commission for Excellence in Education): The NCEE was created by former Secretary of Education T. H. Bell on August 26, 1981. The purpose of the Commission was to present a report to the American people detailing education in America. The report was called *A Nation at Risk* and was released in April 1983 (Gardner, 1983).

OSPI (Office of Superintendent of Public Instruction): OSPI is the primary agency that oversees public K–12 education in the 295 public school districts and six state-tribal education compact schools in Washington State. It is also responsible for allocating funding and providing tools, resources, and technical assistance to facilitate the highest quality education for all students (OSPI, n.d.).

Race to the Top Grants: These grants are funded by the Department of Education to help states encourage school reform concerning:

- Adopting standards and assessments that prepare students to succeed in college and the workplace and to compete in the global economy;
- Building data systems that measure student growth and success, and inform teachers and principals about how they can improve instruction;
- Recruiting, developing, rewarding, and retaining effective teachers and principals, especially where they are needed most; and

- Turning around our lowest-achieving schools. (DOE, 2016, para. 2)

RIGS (Regional Implementation Grants): RIGS are grants that started in 2010 and were released to selected school districts to begin pilot TPEP programs (Wolf, 2013).

SEES (School Employees Evaluation Survey): SEES is a report released yearly by OSPI that describes teacher and administrator trends and disaggregates them by various demographics for grant-reporting purposes (OSPI, n.d.e).

Title 1: Educational legislation that “provides financial assistance to local educational agencies for children from low-income families to help ensure that all children meet challenging state academic standards” (NCES, n.d., para. 2).

TPEP (Teacher and Principal Evaluation Protocol): TPEP was born out of Engrossed Secondary Substitute Senate Bill 6696 during the 2010 legislative session. It was adopted as part of Washington’s Race to the Top Grant and helped provide a common framework for teacher and principal evaluation (TPEP, 2013).

VAM (value added model): VAM models are teacher evaluation systems based in some part on student test scores (NASSP, 2019).

WEA (Washington Education Association): WEA is the largest representation of public school educators in the state of Washington. It is made up of 370 local associations and represents all public school educators in the state (NEA, 2021).

WPTA (Washington Parent Teacher Association): In existence for over 100 years, the WPTA is “a network of millions of families, students, teachers, administrators, business, and community leaders devoted to the educational success of children and the promotion of family engagement in schools” (WPTA, 2021, para. 2).

Chapter 2 - Literature Review

The release of *A Nation at Risk* (1983) provided information to fuel a national debate on the many deficiencies in American education (Gardner, 1983). Shortly after the report's publication, Soar et al. (1983) found that the public perceived teacher evaluation to be a significant part of the problem outlined in the report. This chapter presents a review of academic literature surrounding the history and theoretical framework behind teacher evaluation, a history of the establishment of TPEP legislation in the state of Washington, and a review of each one of the approved evaluation frameworks. The structure, research base, framework similarities, and framework differences of each of the three frameworks will be described.

The Theoretical Foundation of Teacher Evaluation

Teacher evaluation is largely considered to be the collection and use of relevant information to make decisions regarding teacher performance (Darling-Hammond et al., 1983). These decisions fall into two different categories: formative and summative. Formative evaluations generally serve the purpose of developing teachers or programs (Popham, 2013; Scriven, 1966) while summative evaluations are used to determine professional assignment, retention, or termination (Centra, 1993; Millman, 1981; Popham, 1988; Scriven, 1966).

History of Teacher Evaluation

Since the inception of colonial schoolhouses, teacher evaluation has been a topic of local interest (Cremin, 1976). Throughout the 17th and 18th centuries, the needs of communities varied greatly. This variation necessitated the establishment of supervisory committees designed to manage school operations, including teaching staff (Lucio & McNeil, 1969; Robert et al.,

1985). These committees functioned to monitor teacher performance and student progress through classroom visits (Lucio & McNeil, 1969; Robert et al., 1985).

As communities developed and the institutional structures of educational systems continued to expand, those structures outgrew the supervisory committees tasked with their evaluation (Lucio & McNeil, 1969). By 1850, many schools had added new positions that closely resembled today's administrative positions of principals, assistant principals, and superintendents (Rury, 1991). In addition, educational requirements for those wanting to enter the field of education became more rigorous and required two-year teacher preparatory programs (Spring, 2010). By the end of the 19th century, the responsibility for teacher supervision and evaluation had fully shifted to building leaders, such as principals and assistant principals (Marks et al., 1985).

The way in which teachers were evaluated shifted significantly between the end of the 19th century and the beginning of the 20th century. At the end of the 19th century, H. G. Kratz surveyed students from Iowa to identify their perceptions of desirable characteristics in teachers (McNergney et al., 1984). Kratz's hope was to identify characteristics that were beneficial to positive student perception and to use them as a standard benchmark for teacher evaluation. It was discovered that the most desirable characteristics were helpfulness and favorable physical appearance (McNergney et al., 1984). Although the premise of isolating what makes teachers successful was helpful, it was not until the beginning of the 20th century that public education turned to more scientifically grounded principles to evaluate teachers more effectively.

At the beginning of the 20th century, Frederick Taylor outlined specific scientific management principles. These principles measured factory workers' behavior in efforts to increase productivity (Taylor, 1911). The primary principle proposed by Taylor (1911), "to

secure the maximum prosperity for the employer, coupled with the maximum prosperity for each employee,” was the driving force behind his work (p. 9). The translation of these industrial management practices to the education field brought both efficiency and accountability to teacher evaluation (Tanner & Tanner, 1987). Bolstered by research from Arvil Barr (1948) and by Simeon J. Domas and David Tiedeman (1950), the application of industrial management practices shifted the focus of evaluation from characteristics, as proposed by Kratz in 1896, to effective practices, so that teachers could be more effectively trained in the profession (Tanner & Tanner, 1987). The most commonly identified practices that were shown to be effective included student performance, pedagogical skill, and classroom management (Burke & Krey, 2005; Marks et al., 1985).

Although teaching is a scientific profession, it is also very human. As a result, the predictable pendulum shift from the cold practices of scientific management to a more humanistic approach took place in the mid 20th century (Shinkfield & Stufflebeam, 1995). These more humanistic practices, as outlined by Melchior (1950) and Ayer (1954), were seen as beneficial because they called for teachers to be treated as individuals. They also took teacher input into account when making decisions regarding curriculum and instruction (Ayer, 1954; Melchior, 1950). Another benefit of the humanistic approach was that administrators were encouraged to become resources to their teachers rather than simply evaluators (Marks et al., 1985). This shift led to the introduction of formative, rather than just summative, evaluations by the end of the 1950s (Millman & Darling-Hammond, 1990; Shinkfield & Stufflebeam, 1995). The downside of this pendulum swing, as identified by Lucio and McNeil (1969), was that teachers did not receive the support they needed to continually improve their practice if

evaluations were left to the hands-off approach implemented by many administrators (Lucio & McNeil, 1969).

From the middle to the end of the 20th century, a transition from the humanistic approach to a more clinical approach took place (Goldhammer, 1969). The clinical approach begins with a pre-observation conference between the teacher and the administrator, during which rapport is built. In this process, the administrator gains insight into the classroom and the lessons that will be observed (Reavis, 1976). The process continues with a formal observation where data is gathered and analyzed. This step is followed by a post conference in which areas of concern, identified before the observation, are discussed to develop a plan of action. In this analysis, the administrator reviews actions taken to see whether they met the goals of improved instruction and growth toward self-supervision (Reavis, 1976). Key to the clinical approach is a strong rapport and a collegial relationship between teacher and administrator (Tanner & Tanner, 1987). Although there are great advantages to the collegial nature of the clinical observations, they still used the satisfactory/unsatisfactory feedback system, making it difficult for teachers to see where they landed on a more specific performance spectrum. Since the release of *A Nation at Risk* in 1983, an increased sense of urgency regarding teacher evaluation has caused new evaluative trends to emerge (Zackoff et al., 2015). As mentioned in the introduction, *A Nation at Risk* called for large-scale improvements in schools, including increased time on task, increased rigor of courses, increased length of the school day, and increased competence of teaching staff (Gardner, 1983). This call for improvement began the development and the implementation of the current teacher growth model that has pervaded 49 of 50 states (Darling-Hammond & Berry, 1988). Many of these evaluation frameworks, including the system in Washington State, are based on research concerning best practices in both evaluation and instruction. These evaluation

frameworks include specific criteria that are scored on a four-tier system. Using these four-tiered systems, teachers can be rated on a professional continuum ranging from *beginning* to *distinguished*. In addition, the frameworks have specific high-leverage practices for evaluators to look for to increase the validity and reliability of their evaluative measures (TPEP, n.d.).

One of the more controversial components of teacher evaluation began to gain momentum after Race to the Top grants were introduced in 2009. Evaluation frameworks that included a component for student growth (also known as value-added models or VAMs) began appearing to encourage both teachers and evaluators to focus on student growth (Shavelson et al., 2010). These VAMs continue to be a source of debate because they reward teachers based on the performance of their students, and these rewards can include pay raises (Donaldson, 2012; Gordon et al., 2006). Opponents of VAMs argue that an overemphasis on student performance can lead to “academic misconduct ... teaching to the test, or cheating” (McNergney et al., n.d., para. 15). In addition, the American Statistical Association (ASA, 2014) released a position statement in 2014. In this position statement the ASA highlighted multiple other points that must be given consideration when considering VAMs including; VAMs are, many times, based on standardized test scores which do not take into account teacher contribution adequately, VAMs measure correlation, not causation (there could be many causes for a student's increase or decrease in test scores that are independent of the teacher), and that VAMS actually are better measures of system performance than they are of teacher performance (ASA, 2014).

Purpose of Teacher Evaluation

From the onset of teacher evaluation in colonial times, evaluation has served the dual purposes of formative and summative teacher assessment. Formative evaluation is evaluation that is meant to “shape, form, or improve teachers’ performance” (McNergney et al., n.d., para.

7). Summative evaluation is evaluation that is used to help make decisions regarding “initial licensure, hiring, promoting, rewarding, and terminating” (McNergney et al., n.d., para. 7).

Formative evaluation is used to help teachers grow in two primary ways. The first way formative evaluation helps teachers grow is by helping them develop and implement professional growth and development plans (PGDPs). These PGDPs are developed in collaboration with the evaluator and are monitored periodically throughout the year. The second primary way formative evaluation is used to help teachers grow is by pointing them toward targeted professional growth opportunities (Shavelson et al., 2010; Orphanos, 2014; Popham, 2013).

Summative evaluation is used for personnel decision-making of all kinds, including the decision to fire ineffective teachers (Beerens, 2000; Peterson, 2000; Popham, 2013), and for identifying outstanding teachers (Shavelson et al., 2010; Orphanos, 2014; Popham, 2013). One of the consistent purposes of summative evaluation has been to use it as grounds to dismiss substandard teachers (Castetter, 1976; Gordon et al., 2006; Tucker & Stronge, 2005). It was not until the middle of the 20th century that using formative evaluation for teacher development became more widely accepted (Millman & Darling-Hammond, 1990; Shinkfield & Stufflebeam, 1995). The groundwork for the three current evaluation frameworks accepted by Washington State was laid when researchers suggested that schools adopt an evaluation model that was more formative and encouraged teachers to improve, only moving toward dismissal if improvement did not occur (Ellett & Teddlie, 2003).

History of the Washington State TPEP Initiative

Since 2010, Washington State has been planning, legislating, funding, implementing, monitoring, and adjusting the TPEP initiative. Stakeholders from all parties have been included

in the process that has resulted in the state's current evaluation system. Various aspects of the TPEP initiative will be covered in greater detail in the following section.

Washington State ESSB (Engrossed Secondary Substitute Bill) 6696

Washington State ESSB was introduced in 2010. ESSB 6696 contained four main objectives: establishing a TPEP steering committee, establishing a TPEP pilot program, establishing the eight evaluation criteria to be used statewide, and establishing a four-tiered evaluation system (ESSB 6696, p. 18).

TPEP Steering Committee

As of 2023, the TPEP Steering Committee is made up of representation from OSPI, the Governor's office, the Washington State Education Association (WEA), The Association of Washington School Principals (AWSP), the Washington State Parent Teacher Association (WPTA), and the Washington State School Directors Association. The TPEP steering committee is still providing oversight for the TPEP program, which is in its seventh full year of implementation (OSPI, n.d.e). Sue Anderson, committee member, summed up the mission of the committee thus:

The mission statement for the Steering Committee is captured in the core principles of the program:

- High quality teaching and leading are key to student success.
- Growth in practice is developmental in nature.
- Growth occurs best when there are clear standards of practice supported by quality professional learning and learning-focused feedback.
- Evaluation systems should reflect and address the career continuum.

- The focus for teacher and principal growth should be driven by student learning needs.
- Ensuring the system reflects these core principles is the task of the Steering Committee. (Personal communication, November 25, 2018)

TPEP Pilot Project

ESSB 6696 also allowed for the establishment of the TPEP Pilot Project (ESSB 6696, p. 18). In the spring of 2010, over 50 districts applied to be a part of the TPEP Pilot Project. In June 2011, eight districts were chosen as pilot sites: Anacortes, Central Valley, Kennewick, North Mason, North Thurston, Othello, Snohomish, and Wenatchee. In addition to these eight districts, an eight-district consortium from ESD 101 was chosen that included the following districts: Almira, Davenport, Liberty, Medical Lake, Pullman, Reardon-Edwall, Wellpinit, and Wilber. Throughout the 2010–11 school year, these 16 school districts developed their pilot models. In the spring of 2011, a total of 65 districts began the TPEP project early through state-issued Regional Implementation Grants (RIGS) ranging from \$100,000 to \$180,000. By the fall of 2011, all pilot sites had implemented their newly developed evaluation models (TPEP, 2013).

Establishing Common Criteria. ESSB 6696 also established a common set of criteria that all principals and teachers would be evaluated on. These eight criteria for teachers are:

1. Centering on high expectations for student achievement
2. Effective teaching practices
3. Recognizing individual students' learning needs and developing strategies to address those needs
4. Providing clear and intentional focus on subject matter, content, and curriculum
5. Fostering and managing a safe and positive learning environment

6. Using multiple student data elements to modify instruction and improve student learning
7. Communicating and collaborating with families and school communities
8. Exhibiting collaborative and collegial practices focused on improving instruction.

(WAC 392-191A-060, 2013)

The eight criteria for principals are:

1. Creating a school culture that promotes the ongoing improvement of learning and teaching for students and staff
2. Demonstrating commitment to closing the achievement gap
3. Providing for school safety
4. Leading the development, implementation, and evaluation of a data-driven plan for increasing student achievement including the use of multiple student data elements
5. Assisting instructional staff with alignment of curriculum, instruction, and assessment with state and local district learning goals
6. Monitoring, assisting, and evaluating effective instruction and assessment practices
7. Managing staff and fiscal resources to support student achievement and legal responsibilities
8. Partnering with the school community to support learning. (WAC 392-191A-150, 2013)

These criteria provide a uniform platform for all districts across the state to use in evaluation. They are also the same, or very similar to, criteria used in other states, providing some measure for national comparison.

Four-Tiered Evaluation System

The final mandate connected to ESSB 6696 was to develop a four-tiered evaluation system (TPEP, 2013). This four-tiered evaluation system was developed by eight pilot districts, the pilot consortium from ESD 101, and the TPEP Steering Committee. The steering committee, drawing on the experience and voice of the pilot site practitioners, helped develop all aspects of the revised statewide evaluation system (TPEP, 2013). All three of the frameworks accepted by the state of Washington include some variety of the following four tiers:

1. Unsatisfactory
2. Basic
3. Proficient
4. Distinguished (TPEP, 2013).

Each of these criteria contains multiple individual elements within it that allow for targeted evaluation feedback centered on teacher development rather than on punishment or dismissal.

Washington State ESSB 5895

In March of 2012, the Washington State Senate passed ESSB 5895. The main objective of ESSB 5985 was to codify the work of the pilot program that began in 2010. To this end, ESSB added the three recommended frameworks into law, strengthened the proposed student growth goal language, established cut lines for teacher evaluation, and defined the implementation process and details (ESSB 5895, 2013). Washington state ESSB 5895 adds a requirement that OSPI “adopt up to three preferred instructional and leadership frameworks” and that it require “school districts to adopt a preferred framework” (ESSB 5895, 2013, para. E, lines 14–15). To this end, research was conducted by pilot districts and the TPEP Steering Committee to select appropriate frameworks. At the conclusion of the 2011–2012 school year, three

frameworks were chosen for teachers and two frameworks were chosen for administrators. The three chosen teacher evaluation frameworks were the focus of this study.

Washington Teacher Evaluation Frameworks Research Base and Structure

The frameworks chosen for teachers were the CEL 5D+ Teacher Evaluation Rubric 2.0, Charlotte Danielson’s Framework for Teaching (2011), and Robert J. Marzano’s Teacher Evaluation Model. All of these models address the state’s eight evaluation criteria and use a four-tiered evaluation model.

The CEL5D+ Model

The University of Washington Center for Educational Leadership (UWCEL) began working with school districts nationwide in 2001. The UWCEL currently has partnerships in 30 states to improve teaching and learning and to “eliminate educational inequities by creating cultures of rigorous teaching, learning and leading” (UWCEL, 2022). The UWCEL’s CEL5D+ evaluation framework was developed around a wide base of research that included case studies, academic papers, and local practitioners from throughout the nation. Through this research, five domains, each with multiple subdomains and elements, were established. See Appendix C.

CEL 5D+ Model Research Base. Much of the research concerning the CEL5D+ framework centers on coaching instructional leaders and building capacity at the district and building levels. Action research began in this area through partnerships between UWCEL and local school districts (Gallucci et al., 2006). During the 2003–2004 school year, UWCEL formed a partnership with the Highline School District in the Burien/Des Moines area south of Seattle, Washington (Gallucci et al., 2006). Through this partnership, UWCEL provided resources such as instructional leadership training, specific content knowledge aligned with school improvement goals, and increased openness in educational practice. It did so through a

two-year pilot program that included monthly leadership seminars, the development of an instructional leadership council, and targeted coaching (Gallucci et al, 2006). The leadership seminars were planned by UWCEL consultants and district leadership. The seminars brought together building principals and literacy coaches and helped grow their capacity in their content area—in this case, literacy—and as instructional leaders. This capacity growth took place through group observations of content-focused lessons and targeted teaching and learning activities related to the observed content (Gallucci et al., 2006).

The instructional leadership council developed as part of this partnership comprised district leaders, union representatives, principals from all grade levels, and the CEL program director. The council met monthly to build strategy around policy, practices, and district-level structures. Its main objective was to align policy with current improvement initiatives (Gallucci et al., 2006).

Finally, the partnership between UWCEL and the Highline School District involved intense coaching: teaching literacy coaches and teachers best practices, demonstrating those best practices at leadership seminars and in studio/model classrooms, allowing novice teachers chances to be observed attempting the new strategies, and providing feedback (Gallucci et al., 2006). At the conclusion of the two-year pilot program, which ended in the 2005–2006 school year, many pieces of evidence pointed to sustainable success for the Highline School District. This success included greater access and support for the professional development of principals, development of leadership voice, growth in the content area knowledge base, district level ownership of the work and vision, and significant shifts in the time district and building leaders spent in classrooms (Gallucci et al., 2006).

The UWCEL also went through a similar pilot cycle with the Marysville District in Marysville, Washington. Marysville is a district about 35 miles north of Seattle that serves as a bedroom community to Seattle because it offers easy access to public transit into the city and easy access to the interstate (Swanson, 2007, p. 1). At the time of the pilot study, the Marysville School District served approximately 52,000 students: 11,800 in 10 elementary schools, four middle level schools, one comprehensive high school, and three smaller alternative high schools (Swanson, 2007, p. 2). The district had experienced consistent growth resulting from its close proximity to both a military base and a Boeing plant, in addition to the growing commercial interests of the Tulalip Tribe (Swanson, 2007, p. 1). The district's demographics reflected those of the rest of the state: 34% free and reduced lunch, 14% special education, and 74% White, 9% Native American, 8% Latino, 7% Asian, and 2% Black (Swanson, 2007, p. 2). The partnership between Marysville and the UWCEL began in the spring of 2005, when Marysville District Superintendent Larry Nyland realized the work his district was doing in balanced literacy had outgrown their capacity to support the work (Swanson, 2007). Through carryover grant funds, they accessed UWCEL for assistance and formulated a plan to help move balanced literacy forward in Marysville schools (Swanson, 2007). The vision of the partnership centered on empowering the Marysville staff to carry out continued growth independently through gradual release of responsibility (Pearson & Gallagher, 1983) in which the district and UWCEL modeled a "to, with, and by" methodology (Swanson, 2007, pp. 6–7). In this practice, UWCEL consultants came and modeled balanced literacy to the students and teachers. UWCEL consultants then worked with the teachers to teach the balanced literacy practices together. Finally, the UWCEL consultants observed while the balanced literacy best practices were taught

by the teachers. The consultants then offered feedback to help further refine the practice of the teachers involved in the program (Swanson, 2007, pp. 6–7, 15).

During the summer before the official start of the UWCEL/Marysville partnership, in 2005, some structures were put in place to help the partnership begin well. These structures included intensive training connected to Marysville’s already existing summer school. The training included putting two teachers together with groups of 20 students so that they could develop collaborative partnerships, providing hands-on training two afternoons a week, and having teachers observe skilled practice by UWCEL consultant Lyn Reggett (Swanson, 2007, pp. 7–8). This endeavor developed a cohort of 50 teacher leaders that had experienced the authentic professional development and could be effective ambassadors for the program at their local buildings (Swanson, 2007, p. 8).

Over the first year of the pilot study, structures were put in place to help ensure transfer of professional development knowledge to educational practice. This measure was necessary because, as Stein and Nelson (2003) noted,

Teachers must be held accountable for continuously improving their instructional practice ... Principals must not only be capable of providing professional development for their teachers, but also have the knowledge, skills, and strength of character to hold teachers accountable for integrating what they have learned in their professional development into their ongoing practice. (p. 425)

To this end, the Marysville UWCEL partnership established multiple structures designed to support the successful transfer of professional development knowledge. The first structure put in place was called Instructional Practice Seminars. These seminars included teams consisting of district administrators, principals, literacy coaches, teacher leaders, and UWCEL staff. The teams participated with the expectation that they would help colleagues at their sites with “gentle

pressure and relentless support” throughout the process (Swanson, 2007, p. 14). During this time the teams dug into three main themes: using the implementation triangle, increasing district leadership in professional development to model “trying on the practice” (Swanson, 2007, pp. 15–16), and examining the impact of the professional development (PD) model on the culture of the Marysville district. Through these three avenues, the instructional practice sessions served to disseminate best practice to staff in each of the local buildings (Swanson, 2007).

The second structure Marysville put in place was literacy coach training. These sessions helped coaches hone the coaching aspect of their job. District administrators encouraged coaches to find partner teachers to practice on and to make their practice public to those in the building (Swanson, 2007, pp. 21–23). These coaching trainings were started in the 2005–06 school year, but coaches were not required to carry out their official coaching duties until the 2006–2007 school year, which gave them time to practice in a safe scenario before rolling the work out to the rest of the school. As part of the coaching days, coaches were also immersed in experiencing coaching cycles through triads that allowed each coach to experience the three roles they would take on as a coach: teacher, scribe, and giver of feedback (Swanson, 2007, p. 22).

In addition to implementing the literacy coach training and the instructional practice sessions, the Marysville district started conducting regular walkthroughs. These walkthroughs included the superintendent’s leadership team and principals, along with consultants from UWCEL. The walkthroughs were used to inform future professional development and to increase staff motivation to implement the balanced reading initiatives they were focusing on as a district. The expectation was always that district staff would see building-level staff implementing what they had seen modeled in their previous professional development sessions (Swanson, 2007, pp. 19–20).

Marysville also implemented Principal Triads designed to provide coaching and support for both principals and district-level staff simultaneously. These triads helped the district-level staff learn how to coach their principals. They also helped the principals take charge of their own learning, discern their own needs, and find the resources necessary to meet those needs. The principal triads would meet at a school site, walk through classrooms, and then debrief what they had seen with the help of UWCEL consultants (Swanson, 2007, pp. 25–28).

Marysville also decided to apply for, and received, waiver days that are “granted by Washington State to waive the requirement of 180 days of student attendance if those student-free days are used for school improvement” (Swanson, 2007, p. 23). On these days, half of the day was used for district level initiatives and the other half was used for teacher needs. The imperative for these days was for leaders to pass on what they had learned in their training and to increase enthusiasm for the balanced reading initiatives (Swanson, 2007, p. 23).

To ensure fidelity of implementation, Marysville also implemented monthly principal instructional leadership sessions. These meetings included only building leaders. They were used to allow UWCEL consultants to help principals develop their ability to write instructional memos, to assist in planning a year-long professional development schedule, and to help principals develop their leadership voice. Principals were taught to share their core beliefs repeatedly and explicitly, to remind staff of their building-level goals, and to remind them why they were committed to the hard work of continually improving practice. These meetings always concluded with a reflective piece to foster understanding of the purpose and importance of what they were working on that day (Swanson, 2007, pp. 28–29).

The final step that was taken to help develop the district’s balanced literacy initiative was to develop studio classrooms. The concept, which was introduced during summer school,

consisted of teachers who volunteered to have UWCEL consultants in their rooms to model best practice, to co-teach with host teachers, and to observe and help reflect as host teachers put into practice what they had observed. The sessions concluded with feedback sessions between UWCEL consultants and host teachers regarding implementation. The district decided, based on limited resources, to do studio classrooms only in buildings where they had established ownership of the balanced leadership initiative (Swanson, 2007, pp. 32–38).

During the first full year of the initiative, Marysville saw significant increases in trust and excitement among involved staff members. The district also saw growth in coaching and leadership ability among their district-level staff, their principals, their literacy coaches, and their teacher-leaders. Finally, they saw an increase in collaboration and openness among staff members regarding educational practice (Swanson, 2007).

During the second year of the initiative, many of the same structures remained in place. Some notable changes in focus affected the different structures developed in year one. The first change in focus was a shift to “habits of thinking” (Swanson, 2007, p. 30). This included reflecting on where teachers were presently (reflecting on results), defining clearly what they needed to know and be able to do (defining purpose), and then planning how to get there (instructional approach) (Swanson, 2007, p. 31). With reflecting on results, defining purpose, and identifying instructional approach in mind, some adjustments were made to the studio classrooms. These adjustments included allowing for equal distribution, reducing visits to two or three a year, and holding a pre-meeting for the UWCEL consultant to determine where the partner teacher had needs. The habits of thinking also led to the development of communities of practice. These existed at each building, but the structure was determined by building-level leaders because only the building-level leader could effectively define what they needed to move

their staff from where they were to where they needed to be. These communities of practice were then followed up by observations by district-level staff to see the work in action (Swanson, 2007, p. 33).

Connection to the Framework. These two studies, along with multitudes of other research, went into developing the CEL 5D+ framework that was eventually adopted by the state of Washington as one of three approved teacher/principal evaluation frameworks. The information gleaned, and the expertise that was refined as part of these pilot processes, provided the bedrock needed to develop the CEL 5D+ framework that is now used by over 80 districts in Washington, and by many others across the nation, to evaluate teacher effectiveness.

Framework Structure-Domain One: Purpose. The first domain, purpose, is essential for providing students and teachers with direction in their lessons. The domain of purpose contains subdomains for standards and learning targets. The subdomain for standards focuses on teachers connecting lessons to grade-level standards, creating meaning that is connected to other aspects of learning, learning transferable knowledge and skills, and intentionally linking current lessons to previous lessons (UWCEL, 2012). The subdomain of learning targets and teaching points looks for learning targets that are “clearly articulated, linked to standards, understood by students ... and measurable” (UWCEL, 2012, p. 1). This subdomain also includes criteria for success that are “clear to students” and performance tasks that “provide evidence that students are able to understand and apply learning in context” (UWCEL, 2012, p. 1). Finally, the subdomain learning targets and teaching points also envision “teaching points that are based on knowledge of students’ learning needs” (UWCEL, 2012, p. 1).

Framework Structure-Domain Two: Student Engagement. A wealth of research indicates that a lack of student engagement has dire implications. Failure to adequately change

how we teach, what we teach, and how we assess does an injustice to our students and places our future at risk (Gilbert, 2007; Prensky, 2005; Robinson & Aronica, 2009; Tapscott, 1999; Willms, 2003). The second domain in the CEL 5D evaluation framework is about assessing a teacher's effectiveness in engaging students. The student engagement domain has three subdivisions: intellectual work, engagement strategies, and talk. The intellectual work subdomain is evidenced by a classroom where student work regularly “embodies substantive intellectual engagement (reading, thinking, writing, problem-solving and meaning-making)” and where “students take ownership of their learning to develop, test and refine their thinking” (UWCEL, 2012, p. 1). The engagement strategies subdomain is evidenced by classrooms that use strategies that “capitalize on and build upon students’ academic background, life experiences, culture and language to support rigorous and culturally relevant learning” and that use engagement strategies that “encourage equitable and purposeful student participation and ensure that all students have access to, and are expected to participate in, learning” (UWCEL, 2012, p. 1). The final subdomain in the student engagement domain is talk. The talk subdomain is evidenced by student talk that reflects “discipline-specific habits of thinking and ways of communicating” and “embodies substantive and intellectual thinking” (UWCEL, 2012, p.1).

Framework Structure – Domain Three: Curriculum and Pedagogy. Having curriculum and pedagogy that are connected to the desired standards and accessible for all students is another significant piece of the learning puzzle. Lunenburg (2011) explored five different definitions of curriculum, including curriculum as content, as learning experiences, as objectives, as a plan of inquiry, and as a nontechnical approach. Domain three is about investigating the multifaceted nature of curriculum through the use of three subdomains. In the first subdomain, curriculum, traditional instructional materials and tasks are investigated. This

includes evidence of materials and tasks that are “appropriately challenging and supportive for all students, are aligned with the learning target and content area standards, and are culturally and academically relevant” (UWCEL, 2012, p. 2). In addition, materials should be appropriately related to the larger unit’s scope and sequence (UWCEL, 2012, p. 2). The second subdomain includes teaching approaches/strategies. In this subdomain, evaluators are looking for approaches and strategies that intentionally support the lesson’s purpose. In addition, approaches should reflect content knowledge in a way that is “culturally responsive and engaging while being differentiated to meet the needs of all learners” (UWCEL, 2012, p. 2). The final subdomain, scaffolding, challenges teachers to “provide scaffolds for the learning tasks that support the development of the targeted concepts and skills and gradually release responsibility, leading to student independence” (UWCEL, 2012, p. 2).

Framework Structure-Domain Four: Assessment of Student Learning. Assessment is key to any teaching and learning process. Without established measures of growth, discerning what was learned becomes a subjective process. To this end, in the fourth domain, observable assessment practices are identified that will allow teachers to effectively establish what students have learned, what they still need to know, and what learning comes next, allowing for more personalized instruction (Fullan et al., 2006). There are two subdomains within the assessment domain in the CEL 5D+ framework: assessment and adjustments. The first subdomain, assessment, describes students who are aware of the learning target and who are engaged in self-assessment. It also puts forth the expectation that there will be multiple opportunities and multiple modalities for students to demonstrate their understanding of the learning targets. In addition, the subdomain specifies that the teacher use “observable systems and routines for recording and using student assessment data” (UWCEL, 2012, p. 2). Finally, it lays out the

purpose, methods, and criteria for assessment as accessible to all and in alignment with the initially established learning target (UWCEL, 2012). The second subdomain focuses on assessment for learning rather than assessment of learning. Within this subdomain, teachers are evaluated on their use of formative data to make “in-the-moment adjustments”, to give specific feedback to students regarding what they do and do not know, and to adjust future lessons to fit current student needs (UWCEL, 2012, p. 2).

Framework Structure-Domain Five: Classroom Environment and Culture.

Classroom environment and culture is broken down into three subdomains: the actual physical environment, the way the classroom runs, and the way the classroom feels. The first subdomain, physical environment, is tremendously important for helping the classroom run smoothly and getting the most out of students. When a classroom is well organized, more positive interactions happen, and that leads to higher levels of positive student behavior (Martella et al., 2003). For that reason, in the physical environment subdomain teachers look for rooms that have physical arrangements that are conducive to student learning. This category includes spaces that are useful for student assessment and support and are easily accessible for students so that they can become increasingly more independent in their learning process as the year progresses (UWCEL, 2012). The second subdivision addresses the routines and rituals in the classroom. Classrooms that are proficient or distinguished in this area will have clear systems and routines to help students grow in responsibility, maximizing time spent on task (UWCEL, 2012). The final subdivision, classroom culture, deals with how the classroom feels. Classrooms that are exemplars in this subdomain are rich with conversations that reflect “high expectations and beliefs about all students’ intellectual capabilities and create a culture of inclusivity, equity and

accountability for learning” (UWCEL, 2012, p. 3). In addition, these classrooms encourage students to take risks, collaborate, and respect one another (UWCEL, 2012).

Marzano Framework for Teacher Effectiveness

The Marzano Framework for Teacher Effectiveness is also a framework for teacher evaluation that is accepted by the state of Washington (TPEP, n.d.). The Marzano framework is like the CEL 5D+ framework in that it is used to evaluate teachers on the required state eight criteria. It differs in the way those criteria are grouped and divided. The following will be a description of the research behind the framework, the organization of the framework, and the continual evolution of the framework to address concerns as they have arisen during action research. See Appendix D.

Marzano Framework Research Base. The research base supporting the Marzano Framework for Teacher Effectiveness derives from research published beginning in 2001. Beginning with *Classroom Instruction That Works* (Marzano et al., 2001), Dr. Robert Marzano and his colleagues began to outline clear educational practices that led to increased student achievement. In *Classroom Instruction That Works*, Marzano outlined nine academic strategies that help students achieve at any level:

1. Identifying similarities and differences
2. Summarizing and note taking
3. Reinforcing effort and providing recognition
4. Homework and practice
5. Non-linguistic representations
6. Cooperative learning
7. Setting objectives and providing feedback

8. Generating and testing hypotheses
9. Cues, questions, and advance organizers (p. 7)

In 2003, Dr. Marzano published *What Works in Schools* (Marzano, 2003) and *Classroom Management That Works* (Marzano et al., 2003). These two works continued to solidify research-based effective practices by using meta-analysis to outline best practices in all areas of school structure, including practices that are school based, teacher based, and student based (Marzano, 2003; Marzano et al., 2003). Identified school-based strategies include a guaranteed and viable curriculum, challenging goals and effective feedback, parent and community involvement, creating/maintaining a safe and orderly environment, and fostering collegial and professional relationships (Marzano, 2003). Teacher-based strategies include high-leverage instructional strategies, classroom management strategies, and classroom level curriculum and design (Marzano, 2003; Marzano et al., 2003). In addition, this meta-analysis revealed student-based strategies that included home atmosphere, learned intelligence/background knowledge, and motivation to learn (Marzano, 2003).

In 2006, Marzano published *Classroom Assessment and Grading that Works*. In this work, Marzano specifically targeted assessment as a high-yield area for teacher and student improvement, noting that “major reviews of the research on the effects of classroom assessment indicate that it might be one of the most powerful weapons in a teacher’s arsenal” (p. 2). He cited specific research by Paul Black and Dylan Wiliam (1998), stating:

The research reported here shows conclusively that formative assessment does improve learning. The gains in achievement appear to be quite considerable, and as noted earlier, amongst the largest ever reported for educational interventions. As an illustration of just how big these gains are, an effect size of 0.7, if it could be achieved on a nationwide scale, would be equivalent to

raising the mathematics attainment score of a country like England, New Zealand or the United States into the top five after the Pacific Rim countries of Singapore, Korea, Japan and Hong Kong. (p. 61)

During the review of the literature for *Classroom Assessment That Works*, Marzano identified four major themes regarding effective assessment:

1. Feedback from classroom assessments should give students a clear picture of their progress on learning goals and how they might improve.
2. Feedback on classroom assessments should encourage students to improve.
3. Classroom assessment should be formative in nature.
4. Formative classroom assessments should be frequent. (p. 3)

To provide feedback for students that offers a clear picture of their learning, Marzano (2006) proposed the concept of unidimensionality. This concept requires teachers, or curriculum coordinators, to take standards, break them down into single topics that can be assessed, and use those scores to help students self-monitor learning (Marzano, 2006). To help students develop an improvement mindset, teachers can provide students with measurement devices that are more centered on item response theory (Birnbaum et al., 1968) and less centered on point systems (Marzano, 2006). In addition, students should be given these assessments more frequently—ideally, biweekly—to achieve the greatest percentile gains (Bangert-Drowns et al., 1991).

In 2007, Marzano published *The Art and Science of Teaching*. This more comprehensive work included many of the elements in his previous research and began to coalesce around the domains that would first be articulated in the Marzano Framework for Effective Teaching, which was originally published in 2010. In *The Art and Science of Teaching*, Marzano specifically addressed the nine domain questions that appear in the framework:

1. What will I do to establish and communicate learning goals, track student progress, and celebrate success? (p. 9)
2. What will I do to help students effectively interact with new knowledge? (p. 29)
3. What will I do to help students practice and deepen their understanding of new knowledge? (p. 58)
4. What will I do to help students generate and test hypotheses about new knowledge? (p. 86)
5. What will I do to engage students? (p. 98)
6. What will I do to establish or maintain classroom rules and procedures? (p. 117)
7. What will I do to recognize and acknowledge adherence and lack of adherence to classroom rules and procedures? (p. 131)
8. What will I do to establish and maintain effective relationships with students? (p. 149)
9. What will I do to communicate high expectations for all students? (p. 162)
10. What will I do to develop effective lessons organized into a cohesive unit? (p. 174)

These works culminated in the formation of the current Marzano Framework for Effective Teaching that is one of the three accepted frameworks for teacher evaluation in Washington State.

Marzano Framework Structure. The original framework, as developed by Dr. Marzano and his team, bears many similarities to the CEL 5D+ framework. Like the CEL 5D+ framework, the Marzano framework begins broadly, with four domains, and narrows down to 19 subdomains. Each of the subdomains are further divided into a total of 60 elements that teachers are then rated on (Marzano, 2013). The newest version of the framework, called the Marzano

Focused Teacher Evaluation Model, was released in 2018. In order to make the framework more user-friendly and to reflect wide-scale movement toward including the common core state standards in statewide teacher evaluation, the four domains were adjusted from classroom strategies and behaviors, planning and preparing, reflecting on teaching, and collegiality and professionalism to standards-based planning, standards-based instruction, conditions for learning, and professional responsibilities (Marzano Focused Teacher Evaluation Model, 2018). In addition to renaming the four domains, Marzano dropped the number of elements to be scored from 60 to 23. In addition, a requirement to score all 23 elements was added to the revised framework (Marzano Focused Teacher Evaluation Model, 2018). The scoring system contains a four-tiered scoring system (beginning, developing, proficient, and distinguished) as mandated by ESSB 6696, which remained unchanged from a 2017 revision. The Marzano framework is currently being used in 88 school districts across the state of Washington in either the focused or the original form (District Framework Selections, n.d.).

Charlotte Danielson Framework

The Charlotte Danielson Framework for teacher evaluation is the final framework for evaluation approved by the state of Washington in response to ESSB 6696 (TPEP, n.d.). Like the other two frameworks, it is based on a hierarchical organization. All evaluation elements are contained within four domains, which are broken down into 22 components, which are further broken down into the 71 elements for specific feedback (Danielson, 2013). The Danielson framework is also like the other two frameworks in that it is based on the State 8 evaluation criteria. Currently, the Danielson Framework is in use in over 100 districts across the state of Washington, making it the most widely used of the three frameworks (District Framework Selections, n.d.). An explanation of the research behind the framework, the structure of the

framework, and the modifications that have taken place to make the framework relevant for current practitioners follows. See Appendix E.

Research Base for Danielson Framework. The Danielson framework was developed as part of research conducted by the Educational Testing Service (ETS) with the intention of developing an observation-based assessment for first-year teachers, which became The Praxis III, Classroom Performance Assessment (Danielson, 2013). The work by ETS was done in response to the body of work done in the 1990s suggesting that value-added evaluation instruments such as the Danielson Framework had positive effects on student outcomes (Saunders & Rivers, 1996; Wright et al., 1997). It was developed as a basis for pedagogical conversations between teachers and administrators that contained common language (Milanowski, 2011). Studies in multiple districts across the country, including in Coventry, RI, Cincinnati, OH, and Washoe County, NV, indicate small to medium positively correlated effects between teacher ratings on the Danielson Evaluation Framework and student achievement (Milanowski, 2011).

Structure of the Danielson Framework. The first version of the Danielson Framework, released in 1996, was created in partnership with ASCD and in response to research by ETS on evaluation of first-year teachers (Danielson, 2013). It was intended to encapsulate essential teaching skills for both beginning and experienced teachers by breaking quality instruction down into four domains and 22 components. Each of the 22 components was then evaluated on a value-added rubric containing the following ratings: 1 unsatisfactory, 2 basic, 3 proficient, and 4 distinguished (Danielson, 2013). Each of the rubrics is specific to the element and offers concrete actions to look for as evidence supporting the score.

In 2007, the framework underwent several key changes. These changes reflected recent research that was relevant to the framework (Danielson, 2013). The main visible change was the original 22 components being broken down further into elements. This change helped “define the component” (Danielson, 2013, p. 1). In addition, the 1996 framework was revised to include frameworks for “non-classroom specialist positions such as school librarians, nurses, and counselors” (Danielson, 2013, p. 1). Finally, some of the components were renamed to clarify language.

The framework was further revised in 2011. These revisions included tightening the rubric language, adding “critical attributes for each level of performance for each component” and examples of things for evaluators to look for at each performance level (Danielson, 2013, p. 2). These enhancements to the framework helped to increase the validity and reliability of the framework for all practitioners (Danielson, 2013).

Impact of TPEP Initiative

Any time a complex policy is implemented, there are predictable and unpredictable consequences of the implementation (Elfers & Plecki, 2017). TPEP policy is no different. Research by the Washington State Office of Superintendent of Public Instruction (OSPI), conducted by Marge Plecki and Ana Elfers (2017), is used to examine the implications of the policy implementation. The research is now in its 5th full year and focuses on the impact of TPEP policy implementation on teacher retention, administrator retention, increases in teacher workload, adequacy of supports provided, and the usefulness of TPEP for professional development (Elfers & Plecki, 2017). Each of these facets will be addressed in the following sections.

Impact on Teacher Shortage

Current teacher shortages in key content areas such as science, math, and special education are a significant concern (Cowan et al., 2016). Elfers and Plecki (2017) found that among teachers surveyed, 31% of teachers agree, or strongly agree, that TPEP has made them consider leaving teaching. When asked about education being a good lifelong career choice, only 67% of teachers expressed the belief that teaching was a good lifelong career choice. These sentiments may be adding to the already identified teacher shortages in key demographics such as high-poverty and urban areas (Engel et al., 2014) and are concerning, as they have the potential to feed the current cycle of disproportionate opportunities for students who are traditionally marginalized (Sutcher et al., 2016). Data also suggest that novice teachers (teachers with 0–4 years of experience) exit the profession at higher rates—20–25%—than teachers with upper or lower levels of mid-range experience—13% and 19%, respectively—and that veteran teachers are leaving the profession at the highest rate of all—44% (Elfers, Plecki, & Van Windekens, 2017). It is my opinion, from practitioner experience, that the fresh perspective of the novice teacher and the seasoned experience of the veteran teacher are very important to having a well-rounded teaching staff. From 2010–2011 to 2015–2016, the number of first- and second-year teachers was more than doubled in an effort to reduce class sizes (Elfers, Plecki, & Van Windekens, 2017). Losing one in five novice teachers in their first 4 years could result in a future teacher shortage. Given a potential teacher shortage that disproportionately affects students of poverty and minority status, examining the impact of TPEP on the teacher workforce seems to be an important component of this research. Currently there is no longitudinal research on year-to-year desire to leave the profession and perception of career choice. This research is intended to add a second data point to help stakeholders better interpret the impact of TPEP on

potential teacher shortages in high-need content areas and traditionally underserved demographics.

Impact on Administrator Shortage

Administrator retention in key demographics is also a concern (Roza et al., 2003; NAASP, 2021a). Work by Plecki, Elfers, and Wills (2017) suggests that administrative retention is higher in Washington state (81%) than the national average (77.4%). Although it is true that administrative retention is higher in Washington state, 39% of principals and 16% of assistant principals did indicate in one study that TPEP has either somewhat or strongly made them consider leaving the profession (Elfers & Plecki, 2017). Finally, in that same study, Elfers and Plecki (2017) found that 11% of principals and 4% of assistant principals were planning to retire in the next 3–5 years, and an additional 13% of principals and 6% of assistant principals were actively pursuing leaving their building leadership positions for other opportunities.

With these numbers looming, it is critical that researchers gather more data regarding teacher and administrator perception of TPEP's impact on their career choice and their career longevity. This study was conducted with the intent of adding to the existing data to help stakeholders better understand any trends that may be emerging.

Impact on Workload and Supports Provided

Increases in Workload. In an article written for The Center for Strengthening the Teaching Profession, teacher Jan Kragen (2015) stated that TPEP was killing her principal. Kragen argued that the increased workload resulting from comprehensive TPEP evaluations was having a negative impact on her principal and cited changes in behavior such as less time in classrooms engaging informally, less time for instructional leadership, and lower levels of energy as evidence. Although Kragen is not a researcher or an expert in the field, her observations are

not isolated. The Association of Washington State Principals (2019) surveyed over 350 of its members during the 2018–19 school year and found that 67% of all principals and 57% of all assistant principals surveyed worked over 6 days a week. The organization also found that 75% of principals and 67% of assistant principals worked over 56 hours per week. In addition, the survey indicated that 91% of principals indicated they didn’t have time every week to complete their workload and 69% said that their workload had a negative effect on their personal relationships (AWSP, 2019). These data are corroborated by Elfers and Plecki’s (2017) finding that 85% of principals said that TPEP has increased their time working on weekends.

TPEP is not the only source of increased workload administrators have experienced over the last 3 years, but it is a significant one. According to Elfers and Plecki (2017), 66% of district administrators saw time to conduct TPEP evaluations as a “great challenge” for their principals, and 74% identified getting more support for principals to lessen their workload as a major focus for them (p. 39).

Supports Available. One of the most common supports provided to building-level administrators is the hiring of additional administrative staff. Since the TPEP Pilot began in 2011, the number of school administrators has gone up, at least in proportion to FTE allocations:

From 2010-11 to 2017-18, the number of full-time equivalent (FTE) principals increased 7.63 percent, and student enrollment increased by 7.59 percent, nearly identical to the growth rate in principal FTE. However, the number of FTE assistant principals increased by 38.6 percent during this same time period. The number of FTE assistant principals at the elementary level nearly quadrupled over this eight-year time period. (Elfers & Plecki, 2019, pp. 14–15)

Hiring more administrators is only one among myriad strategies that districts have used to combat the increased work demands connected to TPEP.

The level of support offered to building-level leaders has a significant impact on the perceived success of TPEP at the building level, as evidenced by the data in Table 1 (Elfers & Plecki, 2019). Over one third (34%) of building-level administrators indicated that they did not receive adequate support to work with the teachers in their building and just under one third (32%) of administrators indicated that their teachers did not have access to sufficient professional development opportunities related to TPEP (Elfers & Plecki, 2019). Table 2.1 provides data indicating that administrators who reported lack of support also reported negative responses to their ability to focus on instructional leadership, address student discipline, interact with families, and spend informal time with students in comparison with their colleagues who reported feeling supported (Elfers & Plecki, 2019).

Table 2.1.

Views of School Leaders Who Report Receiving Adequate Supports to Work With Teachers on TPEP Compared With Those Who Did Not (Supported, n = 290; Not Supported, n = 152)

<i>How has TPEP impacted your other responsibilities as a school leader?</i>	Somewhat or Very Negative Impact		No Impact		Somewhat or Very Positive Impact	
	Principals and APs Supported	Principals and APs Not Supported	Principals and APs Supported	Principals and APs Not Supported	Principals and APs Supported	Principals and APs Not Supported
Time to focus on instructional leadership	22%	41%	11%	9%	67%	49%

Time to address issues of student discipline	57%	74%	27%	22%	16%	3%
General visibility in the building	37%	49%	18%	20%	45%	30%
Time to interact with families	47%	59%	41%	39%	11%	1%
Time to spend informally with students	53%	65%	23%	23%	23%	10%
Time to respond to paperwork, email, and other managerial tasks	77%	82%	15%	16%	8%	0%

Note: Table adapted from *School Leaders and Teacher Evaluation: Learning, Leading, and Balancing Responsibilities* by A. M. Elfers and M. L. Plecki (2019). Center for the Study of Teaching and Policy, University of Washington, p. 42.

District-Level Supports Offered to Building Leaders. As mentioned previously, 74% of district administrators identified getting more support for principals to lessen their workload as a major area of focus (Elfers & Plecki, 2017, p. 39). The validity of this priority is supported by the data in Table 1. When surveyed, administrators who felt supported responded in a more positive manner than those who did not feel supported. Of the 34% of administrators who participated in the survey and indicated that they did not receive adequate support, 65% disagreed (either somewhat or strongly) that teachers in their school have access to sufficient professional development opportunities related to TPEP, compared with 14% of school leaders who felt supported. Sixty-one percent of the group who perceived a lack of support for TPEP also did not find the majority of time spent on TPEP to be useful, compared with 27% of school leaders who felt supported (Elfers & Plecki, 2019). These differences in opinion about the

usefulness of TPEP cannot be attributed to the number of staff evaluated, as there was no significant difference between the supported and unsupported administrator groups in mean number of staff evaluated (Elfers & Plecki, 2019). The importance of perceived central office support was further underscored by one assistant principal who was interviewed as part of the 2019 survey, who said:

Some districts have taken TPEP seriously and some have not. The have nots designed ineffective systems and agreements that have not contributed to improved instruction and positive student learning outcomes. The serious districts are making good progress and teachers have vastly improved instructional practices. (p. 43)

Common professional development opportunities provided in the district that were identified by building-level administrators as beneficial included rater reliability and calibration, collaboration with other school leaders, strategies for making the TPEP process more manageable, feedback on their own performance, framework training, and help with goal setting for teachers (Elfers & Plecki, 2019). Nearly 90% of building-level administrators found the professional development offered by their district and training to support rater reliability to be useful in some capacity. An even higher percentage, 95%, found collaborating with other administrators to be useful in some capacity (Elfers & Plecki, 2019). Between 11% and 15% of principals indicated they did not have these supports available to them, while 5%–18% of assistant principals indicated they did not have these supports available. The need for district leaders to provide more professional development to their building-level administrators is clear. The majority of principals, 62%, and an even larger majority of assistant principals, 78%, indicated that they would benefit from more professional learning opportunities (Elfers & Plecki, 2019).

Some districts also provided support through outside organizations. These organizations included ESDs, consulting contracts with framework authors, the state principals association, the teachers association, OSPI events, and university-based organizations (Elfers & Plecki, 2019). These events offered opportunities for leaders to collaborate with those of like mind about common areas of focus so that they can refine their practice and expand their understanding of the TPEP evaluation process.

Building-Level Supports Offered to Building Leaders. Districts have taken a diversified approach to the building-level supports offered to their principals. Supports offered include full- or part-time assistant principal, full- or part-time counselor, full- or part-time dean of students, full- or part-time social worker, instructional coaches, department heads, teacher leaders, clerical staff time, and electronic tools (Elfers & Plecki, 2019). The top three strategies to alleviate workload were hiring an additional full- or part-time assistant principal, which 90% of respondents found to be somewhat or very useful; hiring a full- or part-time dean of students which 76% of respondents found to be somewhat or very useful; and having the use of an electronic tool to carry out the observations, which 76% of respondents found to be somewhat or very useful (Elfers & Plecki, 2019).

TPEP Usefulness for Professional Development

Elfers and Plecki (2019) identified administrator workload as a primary concern when it comes to TPEP implementation (Elfers & Plecki, 2019). Because of the increase in workload required to evaluate using the TPEP framework, it is essential that the work translate into stronger professional practice. In that vein, I sought in this study to investigate changes in perspective regarding TPEP's usefulness for professional development as teachers transition to administrators.

Professional Development for Administrators. In their 2019 study, Elfers and Plecki found that three-quarters of principals and assistant principals agreed that TPEP helped them plan their professional development and identify teacher leaders. Both of these activities are useful in solving the workload issue, as they help target professional development and help leverage instructional leadership. Another way that administrators have found TPEP to be useful is for assessing progress on their school improvement plan. Sixty-three percent of school administrators found that TPEP helped somewhat, or a great deal, in monitoring their progress on their school improvement plan (Elfers & Plecki, 2019). Finally, 96% of administrators surveyed said that TPEP had helped improve their ability as an instructional leader (Elfers & Plecki, 2019). Although 17% of administrators said TPEP helped them as an instructional leader only a little bit, 48% said it helped them somewhat, and 29% said it helped them a great deal (Elfers & Plecki, 2019). The data in Table 2 indicate that TPEP is an effective professional development tool for administrators.

Professional Development for Teachers. Professional development is the primary focus of TPEP and can happen in many different ways (Elfers & Plecki, 2019). One way professional development can take place is through teacher collaboration. A majority of administrators, 64%, indicate that TPEP has had a positive impact on professional collaboration and conversations (Elfers & Plecki, 2019). As teachers continue to develop professionally, an increase in quality instruction can be expected. A strong majority of administrators, 80%, indicate that TPEP has had a positive impact on quality of instruction and learning outcomes in their buildings (Elfers & Plecki, 2019).

Table 2.2.

School Leaders' Perceptions of the Extent to Which Feedback From TPEP Informed Other Activities (n = 444)

	Not At All	A Little Bit	Somewhat	A Great Deal
Plan professional development activities for teachers	6%	18%	46%	29%
Assist in identifying teacher leaders	10%	14%	41%	34%
Assess progress on my school's improvement plan	12%	24%	44%	19%
Improve my ability as an instructional leader	4%	17%	48%	29%

Note: Table adapted from *School Leaders and Teacher Evaluation: Learning, Leading, and Balancing Responsibilities* by A. M. Elfers and M. L. Plecki. (2019). Center for the Study of Teaching and Policy, University of Washington, p. 39.

Properly implemented, TPEP can help aid teacher professional development by focusing on teaching practice rather than on the teacher themselves, and it gives teachers and administrators a common framework to work from (Elfers & Plecki, 2019). One administrator interviewed as part of the survey summed it up by saying, “The power in TPEP is being able to make the leadership moves in your building to use TPEP to unite your staff around a shared vision of instructional best practices and continuous learning” (Elfers & Plecki, 2019, p. 25). A vast majority of administrators (86%) say that they use TPEP elements in their staff professional development exercises (Elfers & Plecki, 2019). Common uses include involving teachers as active members in the evaluation process, fostering a deeper dialogue about teaching pedagogy, forming shared goals together, and engaging in peer observations (Elfers & Plecki, 2019).

Professional Collaboration. Quality collaboration among staff has been found to increase student achievement in mathematics and reading (Goddard et al., 2007). Because of this

connection, increased quality of professional collaboration is a very important outcome of the TPEP evaluation process. When surveyed, 65% of principals and assistant principals indicated that TPEP had a somewhat or very positive impact on professional collaboration among teachers, and only 4% said it had a negative impact (Elfers & Plecki, 2019).

Summary

The literature review for this study began by addressing the definition of, and history behind, various theoretical frameworks concerning teacher evaluation as teachers transitioned from local community governance to the teacher growth model adopted by many states today. It also investigated the formative and summative purposes of teacher and principal evaluation. The history of the Washington State TPEP initiative was also outlined to develop a thorough knowledge of the research base and structure of each of the three teacher evaluation frameworks accepted by the state of Washington. Finally, a background for each of the four research questions was provided.

Chapter 3 - Research Methodology

The purpose of this study is to gain insight into whether educator perspectives regarding the TPEP evaluation process change as educators move from the classroom to administration—from evaluator to evaluatee. A review of the research revealed insufficient data to answer the questions posed. The current research base has focused either on teachers or on administrators. It has not addressed individuals who have participated in both realms. Since the TPEP initiative began in the 2012–13 school year, opportunities to gain insights from educators who have both been evaluated and been an evaluator has not been conducted. With this study, I hope to contribute to the body of literature by examining these potential changes in perspective, with specific attention given to the support provided for TPEP, impact of TPEP on perception of career choice, impact of TPEP on workload, and usefulness of TPEP for professional growth. Because the sample size is manageable and easily defined, I used the total population sampling method. Chapter 3 begins with an overview of the research questions, followed by the rationale behind the methodology and research design. It goes on to describe the population and the procedure that was followed. The chapter concludes by describing the survey instruments and providing a rationale for the questions included.

Research Questions

The primary research question for this investigation is: Are there any significant statistical differences between educator perceptions of TPEP as teacher and as administrator?

This question is further broken down into four sub questions:

- Is there a significant statistical difference between educator perspectives on TPEP regarding the support provided as a teacher and as an administrator?

- Is there a significant statistical difference between educators' perspectives on TPEP regarding their opinion of their career choice and career mobility as a teacher and as an administrator?
- Is there a significant statistical difference between educators' perspectives on TPEP regarding their perceived workload as a teacher and as an administrator?
- Is there a significant statistical difference between educators' perspective of TPEP regarding TPEP's usefulness for professional growth as a teacher and as an administrator?

The four null hypotheses are:

H1₀ – There is no significant statistical difference between educators' perspectives on TPEP regarding the support provided as a teacher and as an administrator.

H2₀ – There is no significant statistical difference between educators' perspectives on TPEP regarding their opinion of their career choice as a teacher and as an administrator.

H3₀ – There is no significant statistical difference between educators' perspectives on TPEP regarding their perceived workload as a teacher and as an administrator.

H4₀ – There is no significant statistical difference between educators' perspectives on TPEP regarding its usefulness for professional growth as a teacher and as an administrator.

Research Design

This research project used survey research methodology. In this case the survey was designed to look at both overall and varying tendencies of respondents who had experienced two different perspectives: that of evaluator, or administrator, and that of evaluatee, or teacher.

Surveying is an effective way to collect data to delineate trends in attitudes and opinions of participants. The results gathered from a survey sample can, in many cases, be generalized to a larger group (Creswell, 2003, 2007). For this survey, I chose to sample the entire population of evaluating administrators in Washington State using the total population survey method. For this survey, the most appropriate way to sample the population was to use an online survey. This format gives respondents confidentiality, allows for access to increased population sizes, and decreases the impact of geography and distance on sample size (Creswell & Guetterman, 2021; Dillman et al., 2014; Neuman, 2009).

Sample Population

Total population sampling was used in this study. Total population sampling was chosen because the desired population for this study was of manageable size and could be clearly defined. The sample population for this survey was drawn from the list of all 294 public school districts provided on the Washington State OSPI website (OSPI, n.d.f). All administrators serving in public school buildings were included in this population. Deans of students and athletic directors were not included in this population because in most cases people in these roles do not evaluate certificated personnel using the TPEP frameworks. Principal's assistants (PAs), a position unique to the Spokane Public School District, were included in this project, for two primary reasons: because Spokane Public Schools is the third largest school district in Washington State and employs 50 principal's assistants (Spokane Public Schools, 2021), and because I have firsthand experience suggesting that all PAs participated in certificated evaluation using the TPEP frameworks.

The study population included 3,285 administrators. Although all evaluating administrators were surveyed, not all administrators qualified to complete the survey. Only

administrators who had been evaluated as teachers using the TPEP framework and had evaluated teachers using the TPEP framework were prompted to complete the entire survey. To ensure only qualifying administrators responded to the survey, two yes/no exclusion questions were included at the beginning of the survey. If respondents responded no to either of those questions, they were directed to the end of the survey and thanked for their participation. Because the TPEP pilot project began during the 2012–13 school year, administrators with over 8 years of experience or administrators who had transferred in from another state were effectively eliminated. Although the total population of public-school administrators was surveyed, prior research by Elfers and Plecki (2019) indicates that only 48% of those surveyed, including 31% of principals and 72% of assistant principals, will fit the desired demographic (personal communication, March 25, 2020).

Procedures

IRB Approval and Assurance of Human Rights Protection

Prior to survey dissemination, institutional review board (IRB) approval was obtained from Kansas State University. Approval was obtained to secure assurance that no harm would be done to any survey participants. Approval was obtained on August 10, 2021. A copy of the IRB approval can be found in Appendix E. Measures implemented to protect the rights of all survey participants included:

1. An introductory email notifying participants of IRB approval, assuring them of anonymity, and informing them about the survey. The introductory email also assured participants that they had the right to terminate the survey at any time without consequence.

2. An introductory question for the survey that reminded participants that IRB approval was obtained from Kansas State University, that their answers would remain anonymous, and that they could feel free to terminate the survey at any time without consequence. If the participant responded *no* to this introductory question, the survey was ended and they were thanked for their participation.
3. All survey data, including but not limited to survey answers and IRB approval, were maintained by myself as the researcher on an external hard drive for at least 2 years after study completion.

All surveys were sent out using Qualtrics® software provided by Kansas State University. Qualtrics creates a secure link that is emailed to potential respondents. The secure link created makes an anonymous survey and does not capture any of the participants' personal data. In this case, it recorded only the demographic data that the participant provided on the survey.

Pilot Test

Because the survey was modified from an existing survey, it was necessary to conduct a pilot study. The previously administered portion of the survey, that is the administrator portion, was thoroughly vetted by the survey creators, Ana Elfers and Marge Plecki (2019), who explained:

We conducted pilot surveys sent to a small sample of administrators in the field who then provided us with feedback about the content and the design. This helped us clarify questions and make improvements based on their feedback. We also have vetted survey content and items with outside groups, professional associations, and other experts, and used their feedback in our design. We also have conducted quite a few online surveys for

administrators, and we use the data we received from those prior efforts to understand the strengths and weaknesses of prior items, including understanding which items provide variation in responses. Finally, we based some of our items on prior research designs from other published studies. (Personal communication, March 22, 2020)

The teacher perspective portion of the survey was developed by me as the researcher, in conjunction with Elfers and Plecki. The pilot survey was sent out to a team of 11 respondents, including three university researchers, one ESD superintendent, three building principals, two assistant principals, and two PAs. Babbie (1973) stated that a pilot test should be run in the exact manner in which the final test will be conducted. Pilot team members were sent the exact survey that was to be sent to study participants, and responses were collected. Pilot survey members were also asked to give feedback, independently of the survey, that would ensure survey reliability and validity. Feedback was collected and implemented to develop the final version of the survey used for the study. Changes in both questions and the mechanics of the survey were implemented to make the survey more user friendly.

Dissemination of the Survey

Currently, there is no listserv that lists all administrators in Washington State. The OSPI website maintains a list of all school districts in Washington State with links to district home pages. Building administrator contact information, including emails and phone numbers, was gathered by accessing district and school websites via the OSPI-provided links (OSPI, n.d.f). A group email was sent out to all administrators explaining the purpose of the survey and providing a link to the survey and the timeline for completion. The email also offered administrators who completed the survey an entry into a drawing for one of four \$25 Amazon gift cards and a copy of the anonymous survey results. Emails that were returned as invalid were cross-checked, and

corrections were made when necessary. Each week after the introduction email was sent out, an email containing the survey link was sent out that contained a reminder of the purpose, parameters, and deadline for survey completion. The survey window was open from August 1 through January 1 to provide adequate time to obtain as many responses as possible.

Data Collection and Analysis

Administrator contact information was accessed via the Washington State OSPI website directory of websites and contact information. District websites were used to access individual school websites. Contact phone numbers and email addresses were gathered from each school website serving students in K–12th grade. If email addresses were not available on the school or district website, phone calls were made to office personnel who could provide email contact information for all school administrators. This data collection began in February 2020 and concluded in April of 2020. On August 1, 2021, an introduction email was sent out to all email addresses in the database through Qualtrics. This email briefly outlined the purpose and scope of the study, assured all respondents of anonymity and confidentiality, included a link to the survey, and included the author's contact phone number and email for any questions. It also included an offer to send participants the disaggregated data when the survey was complete. Respondents were given 1 month to complete the survey. A reminder email was sent out each week to respondents who had not yet completed the survey. A final reminder was sent out 2 days before the deadline to those who had not yet responded. A personalized follow-up thank you was sent after each survey was completed. Once the data collection was complete, the surveys were closed. The survey data were then exported from Qualtrics to an Excel CSV file and analyzed using basic Excel functions. First, the data were divided into two groups: the teacher data and the administrator data. The data were then analyzed for descriptive statistics, including mean,

median, mode, and frequency distribution. This was done to identify any outliers that might skew the data. Descriptive statistics for both the teacher and the administrator data were presented in table format. Once descriptive statistics were obtained and outliers had been removed for both teacher and administrator groups, each paired question was analyzed using a t-test to discern whether statistically significant differences existed between the groups.

Data security was ensured using multiple steps. The Qualtrics survey tool ensures anonymous responses. In addition, the data were stored on a password protected computer at all times. The data stored on the survey website were deleted upon completion of the project. All data presentation was also done in a manner that preserved the anonymity of all respondents and was disaggregated only by demographic data collected via the survey.

Instrumentation

This quantitative study used an adapted version of the 2018 UW TPEP Principal and Assistant Principal Survey (See Appendix A; Elfers & Plecki, 2019). In collaboration with Elfers and Plecki, I adapted this survey to account for the administrative and the classroom teacher perspectives of those practitioners who had recently entered the field of administration. Validity of the administrator portion of the survey was established using the following strategies:

We conducted pilot surveys sent to a small sample of administrators in the field who then provided us with feedback about the content and the design. This helped us clarify questions and make improvements based on their feedback. We also have vetted survey content and items with outside groups, professional associations, and other experts, and used their feedback in our design. We also have conducted quite a few online surveys for administrators, and we use the data we received from those prior efforts to understand the strengths and weaknesses of prior items, including understanding which items provide variation in

responses. Finally, we based some of our items on prior research designs from other published studies. (A. Elfers & M. Plecki, personal communication, March 22, 2020)

Validity for the teacher portion of the survey was established by submitting the survey to a pilot team that included one ESD superintendent, three university faculty, two building principals, two assistant principals, and two principal's assistants. Feedback was then used to modify the survey to suit the needs of both researchers and practitioners. Appendix A includes a copy of the adapted TPEP Changes in Perspective Survey that was distributed to participants.

The data collection software Qualtrics was used for this study. Qualtrics creates a secure link that can be e-mailed to members of the survey population. Qualtrics creates an anonymous survey that does not capture any of the participants' personal data and records only the demographic information provided by the survey respondent. There were 50 total questions on the survey, which took approximately 15–20 minutes to complete. The survey was divided into four different sections: qualifying questions, demographics, administrator perspectives, and teacher perspectives.

Qualifying Questions

The first section consisted of three survey qualification questions that were yes or no questions. These questions were intended to identify whether participants understood the IRB statement included at the beginning of the study, had been engaged with TPEP as a teacher, and had been engaged with TPEP as an administrator. These questions included skip logic to send participants to the end of the survey if the participants did not agree to the IRB question or had not been involved in TPEP both as an administrator and as a teacher.

Demographic Section

The second section consisted of 14 questions designed to identify the demographics of the respondents. The intent of these questions was to identify whether any trends that emerged were unique to specific demographics.

Administrative Perspective Section

The administrative perspective section contained 17 questions— two multiple choice, two choose-all-that-apply, nine matrix, two short answer, and two yes or no—asking administrators to give feedback regarding TPEP from an administrative perspective.

Teacher Perspective Section

The final section asked administrators to give feedback regarding TPEP from the perspective on their last year as a classroom teacher using 16 questions; three multiple choice, seven matrix, three choose all, two short answer, and one yes or no.

Summary

The purpose of conducting this research was to determine whether there is a statistically significant difference between the way evaluating administrators viewed TPEP as a teacher and how they currently viewed it as an administrator. This chapter focuses on the methodology used in this research. The research used a total population, exploratory, quantitative survey method. The survey gathered information from administrators regarding their perspectives on TPEP both as a teacher and as an administrator. The survey items were created using a survey initially published by Elfers and Plecki (2019) that was targeted at administrators and teachers as individuals. The questions were modified to fit the current research design and purpose.

Chapter 4 - Data, Results, and Analysis

This research was conducted to determine whether educators who had experienced TPEP as both a teacher and as an administrator experienced any statistically significant changes in perspective regarding TPEP's usefulness for professional development, the support received implementing TPEP, their career choice as a teacher or an administrator, and TPEPs impact on their perceived workload. This study consisted of an online survey with 50 questions broken down into four types of questions: qualification questions, demographic information, administrator perspective, and teacher perspective. These surveys were sent during August of the 2020 school year. The survey went out to all evaluating administrators in Washington State. This included principals and assistant principals. Additionally, PAs in the Spokane School District were included because they participate in certificated staff evaluation. Athletic directors and deans of students were not included, as they do not have evaluation duties in Washington State.

TPEP Survey Results

Demographic Information

During the 2020–21 school year, there were 3,285 evaluating administrators practicing in 294 school districts in Washington State. This number was determined by accessing each school district website and entering email contact information for all principals and assistant principals. District website links were accessed using the Washington State OSPI Website (OSPI, n.d.f). Although 3,285 administrators were practicing in Washington State, only a limited number of them qualified for the survey. Elfers and Plecki (2019) estimated that only 48% of those administrators, 31% of the principals, and 72% of the assistant principals would qualify to

complete the survey, leaving the total estimated eligible population at 1,576 administrators (personal communication, March 25, 2020). The total population was sampled to get the greatest possible cross-section of respondents and maximize the ability to generalize survey results to the whole population.

Of the 3,258 administrators surveyed, 232 administrators responded. Respondents who had never been evaluated using TPEP, who had never evaluated teachers using TPEP, or who did not consent to the informed consent notification were not able to complete the survey. This left a total of 95 administrators who completed the survey. Of the 95 administrators who completed the survey, 64.2% were female, 33.6% were male, and two preferred not to say. Table 3 shows the survey respondents by gender and their percentage of total respondents. In addition, 70.5% were principals, 24.2% were assistant principals, and 5.2% had other roles. Table 4 shows the survey respondents by their role and their percentage of total respondents. Most of the respondents, 92.6%, identified as White; 3.2% identified as Asian, 2.1% as Hispanic, and 1% each as American Indian, other race, or no response. Table 5 shows the survey respondents by their ethnicity and their total percentage of respondents. The demographic with the highest number of responses was White female principals (44.2%), followed by White male principals (24.2%), White female assistant principals (12.6%), White male assistant principals (6.3%), Asian female assistant principals (3.2%), and White females with other roles (2.1%). American Indian male principals, Hispanic male assistant principals, White males with other roles, White principals who preferred not to identify their gender, Hispanic female assistant principals, female assistant principals of another race, and people with other roles who did not want to identify their gender or race each comprised 1% of the survey respondents. Table 6 shows the survey results by gender, role, and ethnicity.

Table 4.1. Survey Participants by Gender

Survey Participants by Gender

Gender	Total Number	Percentage of Total Respondents
Female	61	64.2%
Male	32	33.6%
Prefer Not to Respond	2	2.1%
Total	95	99.9%

Table 4.2.

Survey Participants by Role

Role	Total Number	Percent of Total Respondents
Principal	67	70.5%
Assistant Principal	23	24.2%
Other	5	5.2%
Total	95	99.9%

Table 4.3.

Survey Participants by Ethnicity

Ethnicity	Total Number	Percentage of Total Respondents
White	87	91.5%
Asian	3	3.2%
Hispanic	2	2.1%
Native American	1	1.1%
Other	1	1.1%

Ethnicity	Total Number	Percentage of Total Respondents
No Response	1	1.1%
Total	95	99.8%

Table 4.4.

Survey Participants by Gender, Role, and Ethnicity

Ethnicity/Gender/Role	Total Number	Percentage of Total Respondents
White Female Principal	42	44.2%
White Male Principal	23	24.2%
White Female Assistant Principal	12	12.6%
White Male Assistant Principal	6	6.3%
Asian Female Assistant Principal	3	3.2%
White Female Other	2	2.1%
American Indian Male Principal	1	1%
Hispanic Male Assistant Principal	1	1%
White Male Other	1	1%
Prefer Not To Say White Principal	1	1%
Hispanic Female Assistant Principal	1	1%
Other Race Female Assistant Principal	1	1%
Prefer not to say, did not answer, Other	1	1%
Total	95	99.4%

An attempt was made to compare the demographic data of the respondents with the actual state demographics. Contacts were made to researchers at the University of Washington, OSPI,

AWSP, and WASA. The only organization that was able to provide me with some demographic data to compare to my sample was WASA and that comparison will be discussed further in chapter five.

Findings

Statistically Significant Differences

This study was conducted to determine whether there were statistically significant differences between teacher and administrator perspectives regarding TPEP and its impact on these professionals' perspectives as they transition from the role of teacher to administrator. The teacher and the administrator portion of the survey had paired questions to compare the different perspectives. The items on the first half of the survey asked respondents to evaluate their perception of TPEP from an administrator perspective using a variety of different question types. Questions included yes/no, multiple choice, short answer, and matrix-style questions. Several of the questions were measured on Likert scales. Each response on the Likert scale was given a numerical value for analysis. Higher numbers were assigned to positive impact and lower numbers were assigned to negative impact; therefore, the higher the mean score, the more positive the view of TPEP's impact. Scales varied appropriately according to the question type. A list of the questions, their Likert scales, and the numerical value assigned can be found in Table 4.5, Table 4.6, and Table 4.7.

Table 4.5.

Likert Scale for Paired Questions 19/37 and 33/48

Likert Scale	Strongly Agree	Somewhat Agree	Somewhat Disagree	Strongly Disagree
Numerical Value	4	3	2	1

Table 4.6.***Likert Scale for Paired Question 25/43***

Likert Scale	Extremely Useful	Somewhat Useful	A Little Bit Useful	Did Not Participate	Not Available	Not Useful At All
Numerical Value	4	3	2	1	0	-1

Table 4.7.***Likert Scale for Paired Questions 31/46, 32/47, 33/48***

Likert Scale	Extremely Positive	Somewhat Positive	Neither Positive nor Negative	Somewhat Negative	Extremely Negative
Numerical Value	5	4	3	2	1

The teacher and administrator paired questions about perception of TPEP had multiple components. Each component of each question was analyzed using a *t*-test. The *p*-value to determine statistical significance was set at $p < .05$. A total of 24 items were analyzed. Of those 24 items, nine were found to have statistical significance. Because the administrator and teacher survey questions were slightly different, some components needed to be dropped, as there was not adequate comparison to be made between the teacher and administrator portions of the survey. Questions were grouped based on the four null hypotheses and have been labeled H1–H4. These new numbers can be found in Appendix F and will be used throughout this paper when specific items are being discussed.

Group H1 – Support Provided

H1₀ – There is no significant statistical difference between educators’ perspective on TPEP regarding the support provided as a teacher and as an administrator.

This was found to be a valid null hypothesis because none of the items in H1 had a statistically significant difference. Means ranged from 2.67 for item H1-1 (I receive sufficient support for my evaluation) indicating administrators somewhat agreed, to 1.69 for item H1-4 (usefulness of training to support rater reliability) indicating that training for rater reliability was only slightly to somewhat useful. Table 10 provides the mean and standard deviation and p-values for group H1, perceived usefulness of supports. Table 10 also provides the means, standard deviations, and p-values for each of the questions in the group. The items in this group addressed teacher and administrator perception regarding usefulness of supports. Items included receiving sufficient support for my own evaluation, professional development provided by the district, collaborating with other colleagues on TPEP issues, and receiving training to support rater reliability.

Table 4.8.
Group H1 – Perceptions About Usefulness of TPEP Supports

	Administrator		Teacher		
New item number	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>p</i>
H1-1	2.670	0.916	2.500	0.944	0.175
H1-2	1.826	1.485	2.000	1.294	0.389
H1-3	2.278	1.460	2.067	1.510	0.325
H1-4	1.689	1.517	1.667	1.421	0.912
Overall H1	2.116		2.058		

Group H2 – Opinion of Career Choice

H2₀ – Is there a significant statistical difference between educators’ perspectives on TPEP as a teacher and as an administrator regarding their opinion of their career choice?

Evidence suggests this null hypothesis is invalid because two of the four questions in this question group were found to have a statistically significant difference. Items in group H2 sought to explore how teacher and administrator perceptions of their career choice had been impacted by TPEP. Table 4.9 shows the means, standard deviations, and p-values for all H2 items.

Item number H2-1 addressed the level of agreement with the statement “I have better interactions with my teachers/administrator because of TPEP.” The administrator mean was higher than the teacher mean, indicating that administrators believe that TPEP helps them have better interactions with their staff more in their present administrator role than they did when they were teachers. Respondents were more likely to strongly or somewhat disagree with the statement “I have better interactions with my teachers/administrator because of TPEP” from a teacher perspective and more likely to somewhat or strongly agree with the same statement from the administrative perspective. This question combination had a *p*-value of .001, which is lower than the established threshold of .05 that indicates support for statistical significance. The distribution of answers is found in Figure 4.1.

Table 4.9.

Group H2 – Perceptions About Career Choice

	Administrator		Teacher		
New item number	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>p</i>
H2-1	2.600	0.751	2.311	0.745	0.001*
H2-2	2.663	0.854	2.696	0.831	0.767

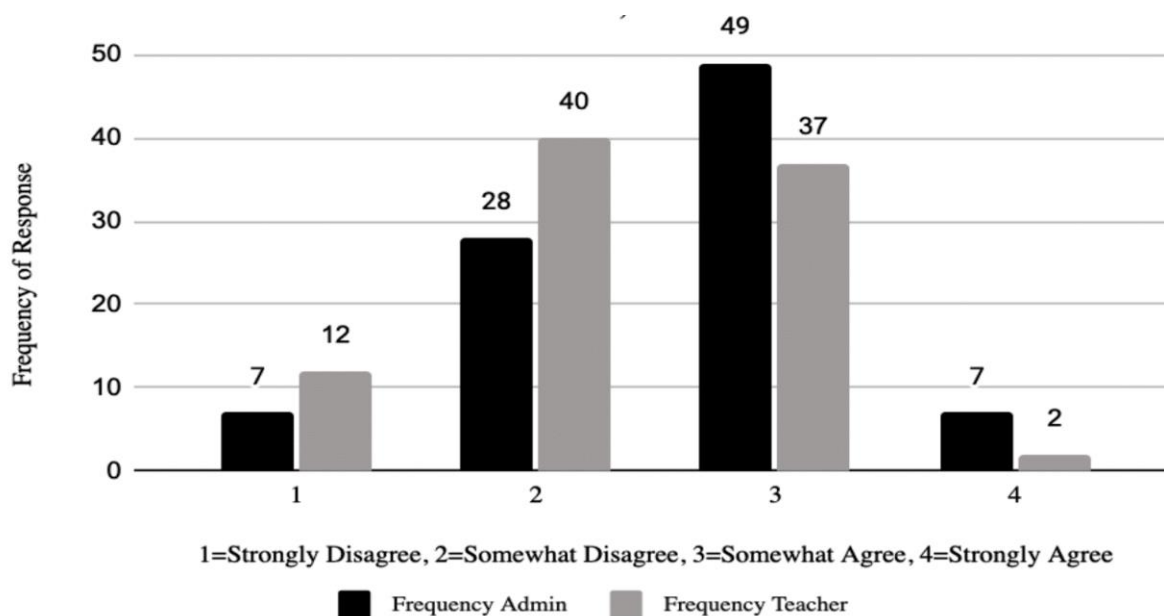
H2-3	2.283	0.770	2.380	0.781	0.191
H2-4	1.747	0.948	1.473	0.762	0.003*
Overall H2	2.323		2.215		

Note. * $p < .05$

Figure 4.1.

Distribution of Item H2-1

(“I have better interactions with my teachers/administrators because of TPEP.”)

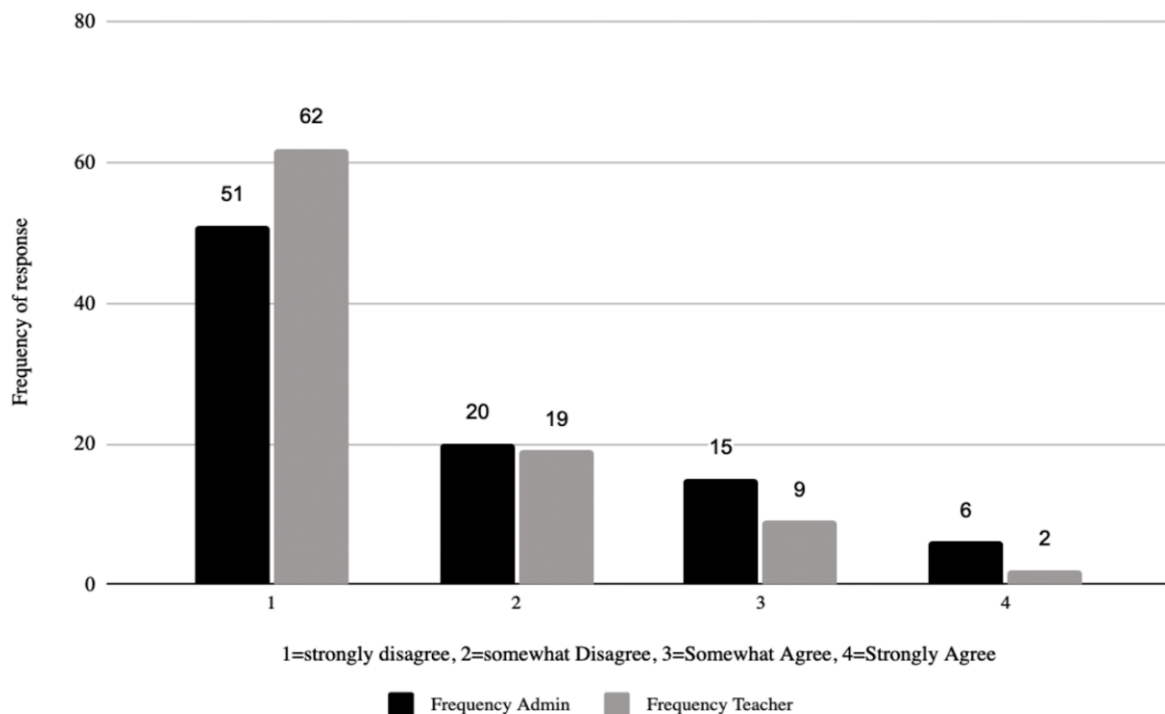


Item H2-4 asked respondents how much they agreed with the statement “TPEP has made me consider leaving my position.” The administrator mean of 1.747 corresponded more closely with the somewhat disagree response while the teacher mean of 1.473 corresponded more closely with strongly disagree. This question combination had a p-value of .003, which is lower than the threshold of .05 and indicates there is statistically significant support for the statement that administrators are more likely to leave their position as a result of TPEP than they were as

teachers. Figure 4.2 shows the distribution of the responses from both the administrator perspective and the teacher perspective.

Figure 4.2. Distribution of Item H2-4

Distribution of Item H2-4 (“TPEP has made me consider leaving my position.”)



It is worth noting that the range for questions H2-1–H2-3 is only .38 while the difference between questions H2-3 (the second lowest mean, 2.28) and H2-4 (the lowest mean, 1.75) is .53, indicating a larger gap between means 3–4 than between means 1-3 combined. This result will be further discussed in Chapter 5.

Group H3 – Impact on Perceived Workload

H30 – There is no significant statistical difference between educators’ perspective on TPEP as a teacher and as an administrator regarding their perceived workload.

Evidence suggests that this null hypothesis is false because six out of seven questions had p-values indicating a statistically significant difference between the teacher and the administrator perspective. Questions in group H3 were about how teacher and administrator perspectives differed in terms of TPEP impact on workload. For items in group H3 (perceived workload), on a 5-point Likert scale, items ranged from 3.25 on H3-1 (time to focus on instructional leadership) to 2.33 on H3-5 (time to respond to paperwork, email, and other tasks). These means indicate that, when responding from an administrative perspective, respondents believed that TPEP had a slightly positive impact on their ability to focus on instructional leadership but a slightly negative impact on all other areas surveyed, including time to focus on student discipline, time to interact with families, time to attend to paperwork and other administrative tasks, and the ability to spend informal time with students. For items in the H3 group on a 4-point Likert scale, means ranged from 2.97 on item H3-6 (TPEP is useful but takes up too much of my time) to 2.81 on H3-7 (I would like more time to work on my own evaluation). These means indicate that, when responding from an administrative perspective, respondents slightly agreed with the statements that TPEP is useful but takes up too much of their time and that they would like more time to work on their own evaluation. Table 4.10 shows the means, standard deviations, and p-values for all questions in the H3 group.

Table 4.10.

Group H3 – Perceptions About TPEP Impact on Perceived Workload

	Administrator		Teacher		
New item number	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>p</i>
H3-1	3.253	.926	3.220	.884	.755
H3-2	2.374	.999	2.956	.910	.000

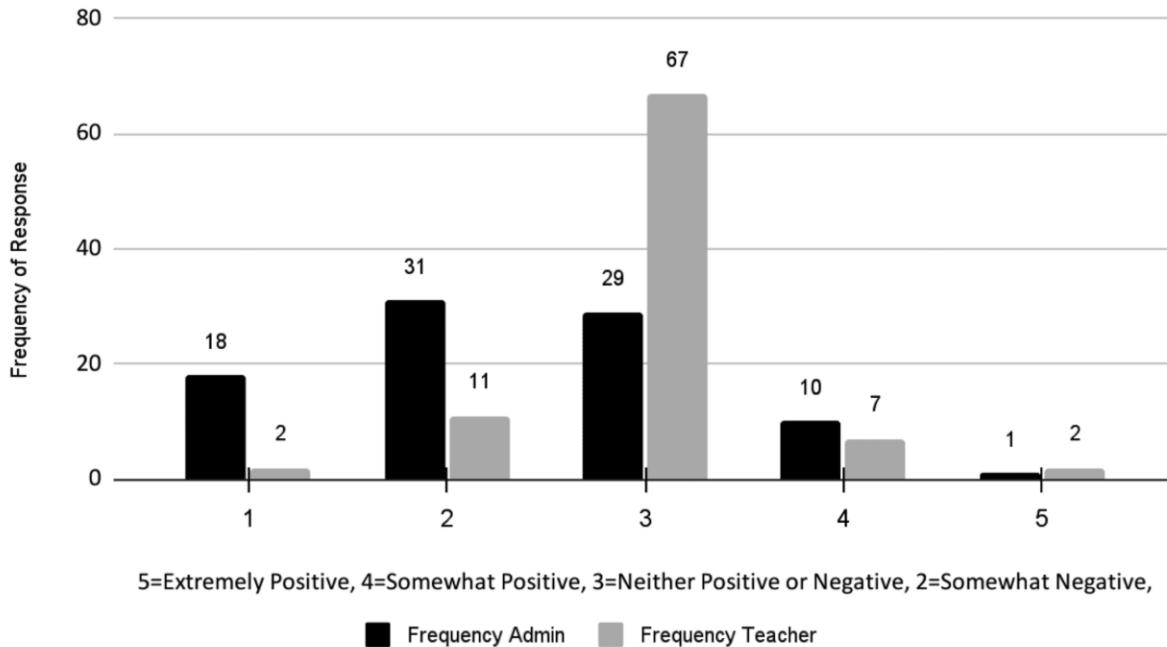
H3-3	2.568	.986	2.977	.624	.000
H3-4	2.637	.803	3.033	.646	.000
H3-5	2.330	1.122	2.750	.699	.000
H3-6	2.967	.848	2.663	.673	.001
H3-7	2.811	.824	2.453	.814	.001
Overall H3	2.706		2.865		

Note. * $p < .05$

For item H3-2, TPEP's impact on my ability to address issues of student discipline, the mean administrator perspective response was 2.374 while the mean for the teacher perspective response was 2.956. This result indicates that respondents were more likely to perceive TPEP as having a somewhat negative impact on their time to address issues of student discipline from the administrator perspective than they were from the teacher perspective. The p-value for H3-2 was .000, indicating a statistically significant difference in the way respondents viewed TPEP's impact on their ability to address issues of student discipline. Figure 4.3 shows the distribution of responses for item H3-2 from both the administrator and the teacher perspective.

Figure 4.3.

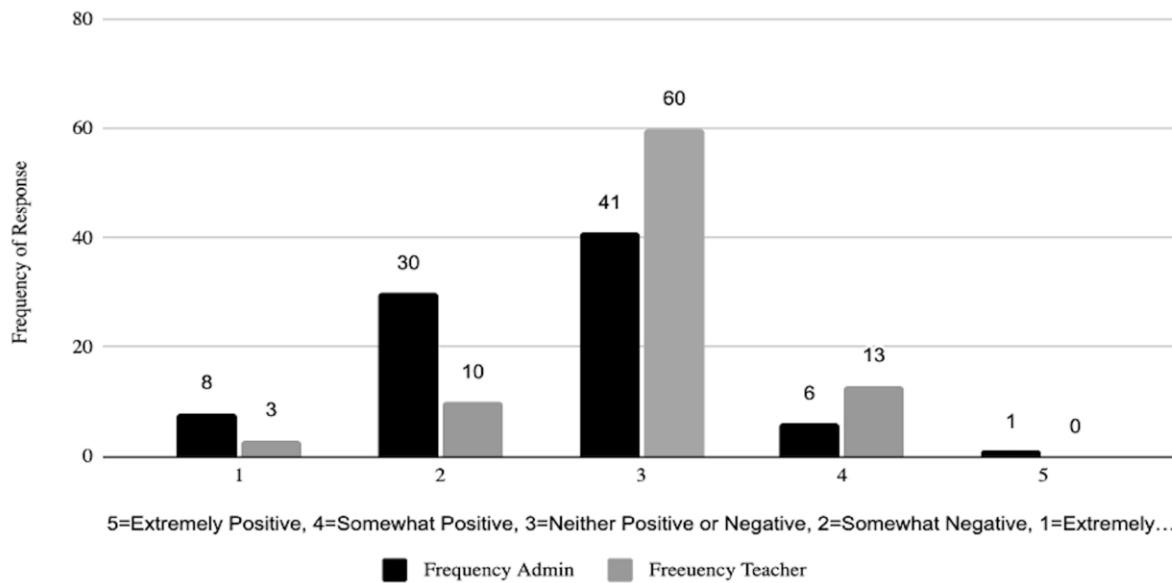
Distribution of Item H3-2 (TPEP's Impact on My Ability to Address Issues of Student Discipline)



For item H3-3, TPEP's impact on my time to interact with families, the mean response for administrator perspective was 2.568, between neither positive nor negative and somewhat negative, while the mean response for the teacher perspective was 2.956, which was very close to neither positive nor negative. This indicates that respondents were more likely to report that TPEP had a negative impact on their time to interact with families from an administrator perspective than from a teacher perspective. The p-value for this question was .000, indicating that there is evidence to suggest the difference between the two perspectives is statistically significant. Figure 4.4 shows the distribution of responses for item H3-3.

Figure 4.4.

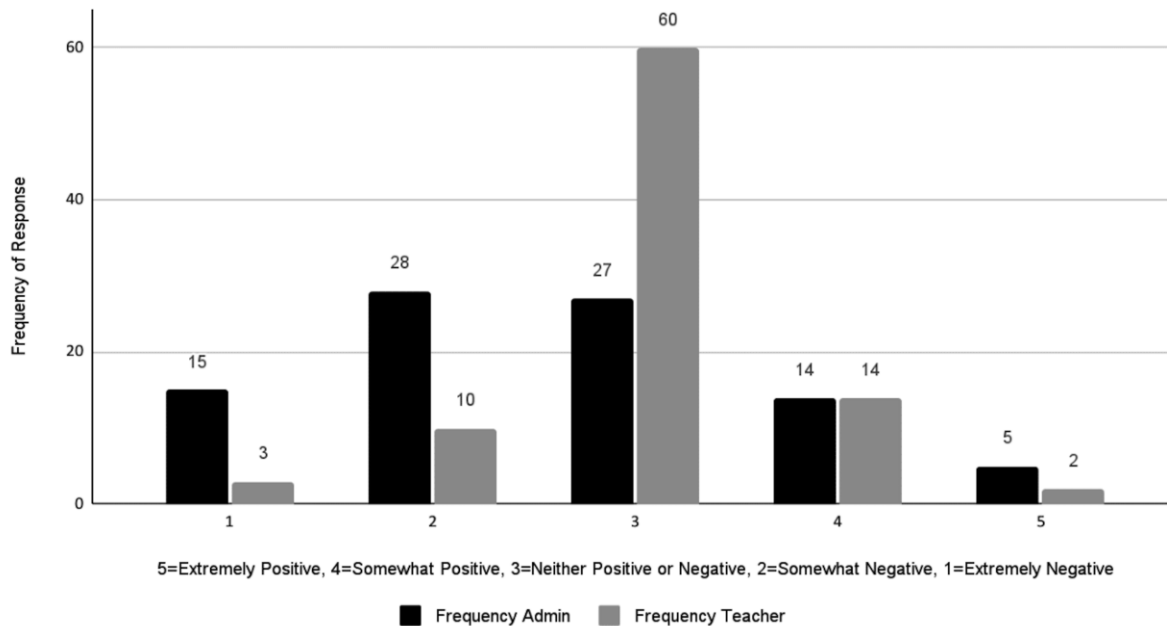
Distribution of Item H3-3 (TPEP's Impact on Time With Families)



For item H3-4, TPEP's impact on my ability to spend informal time with students, the mean for administrator perspective was 2.673 and the mean for teacher perspective was 3.033. These means indicate that respondents were more likely to experience a somewhat negative impact on their ability to spend informal time with students as an administrator than they were as a classroom teacher. The p -value for this question was .000, indicating that there is evidence to suggest that the difference between the two responses is statistically significant. Figure 4.5 shows the distribution of answers for item H3-4.

Figure 4.5.

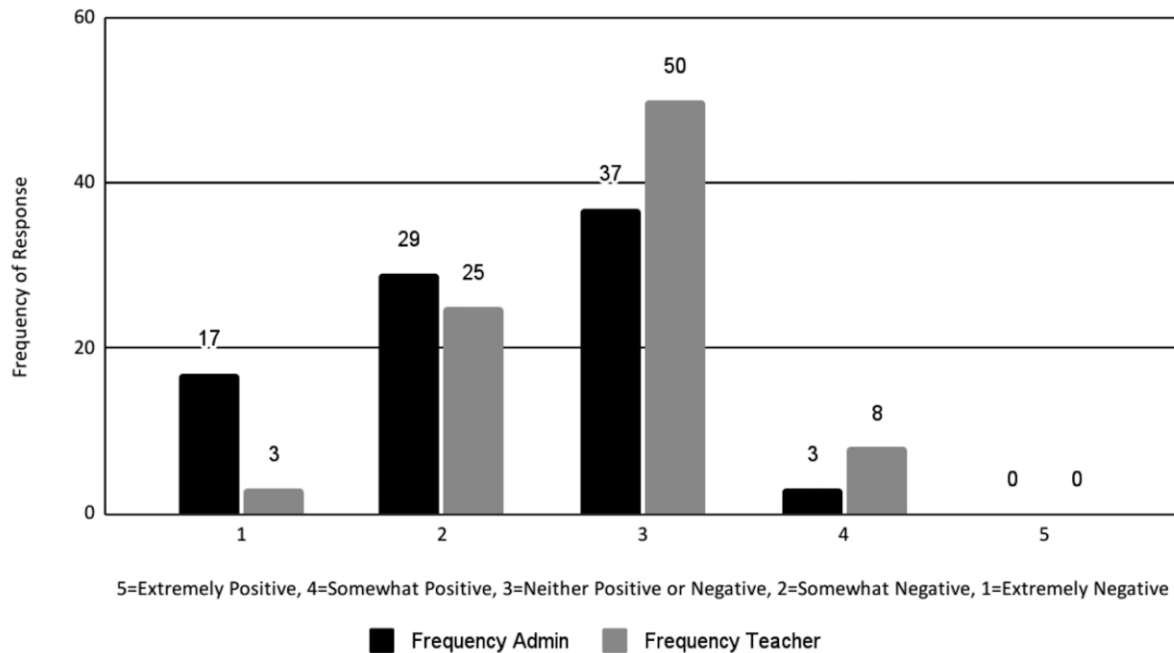
Distribution of Item H3-4 (TPEP's Impact on Informal Time With Students)



For item number H3-5, TPEP's impact on my ability to respond to paperwork, email, and other tasks, the mean for the administrative perspective was 2.330, while the mean for the teacher perspective was 2.750. These means indicate that respondents were more likely to see TPEP as having a somewhat negative impact on their ability to respond to paperwork, email, and other tasks from an administrative perspective than they were from a teacher perspective. The p-value for this question was .000, suggesting that the difference between teacher and administrator perspectives is statistically significant. Figure 4.6 shows the distribution of answers for item H3-5.

Figure 4.6.

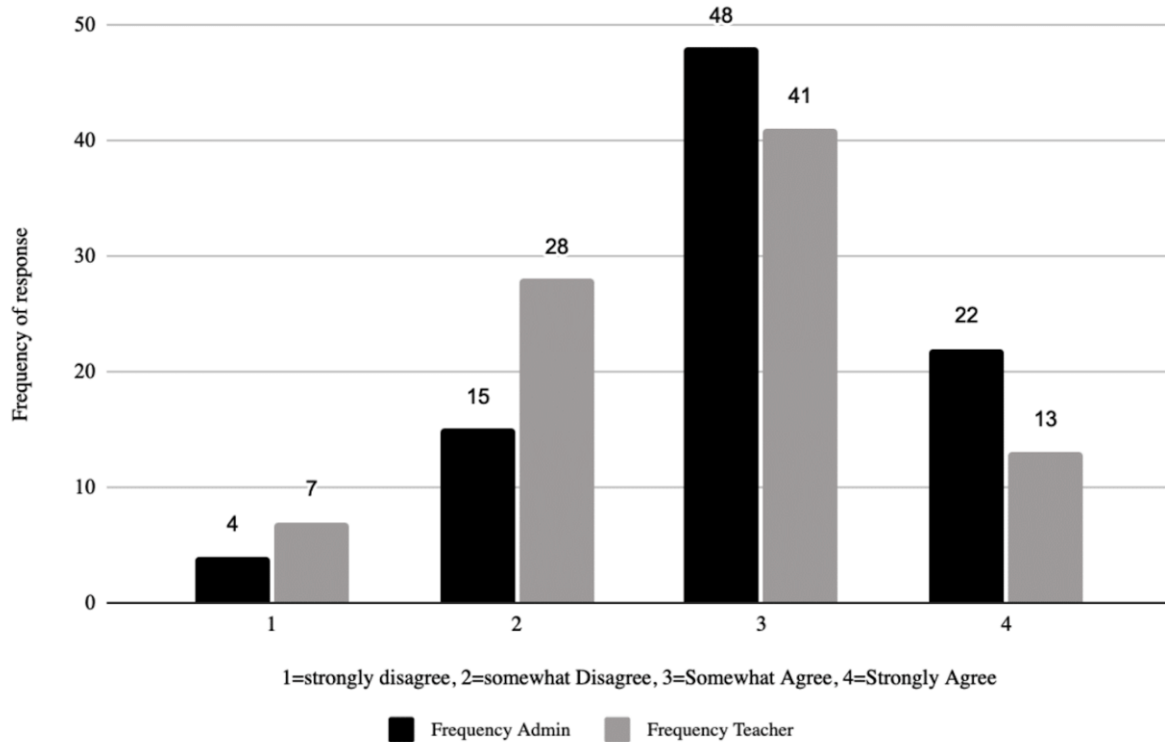
Distribution of Item H3-5 (TPEP's Impact on Time to Respond to Paperwork, Email, and Other Tasks)



For item number H3-6, respondents were asked to rate their level of agreement with the statement “TPEP is useful but takes up too much of my time.” Possible responses included strongly agree (4), somewhat agree (4), somewhat disagree (2), and strongly disagree (1). The mean for item H3-6 from the administrator perspective was 2.967, and the mean from the teacher perspective was 2.663. This difference in means indicates that respondents were more likely to somewhat agree with the statement “TPEP is useful but takes up too much of my time” when responding from the administrative perspective than when responding from the teacher perspective. The p-value for these questions was .001 and indicates that the difference between these two means is statistically significant. Figure 4.7 shows the distribution of responses for item H3-6.

Figure 4.7.

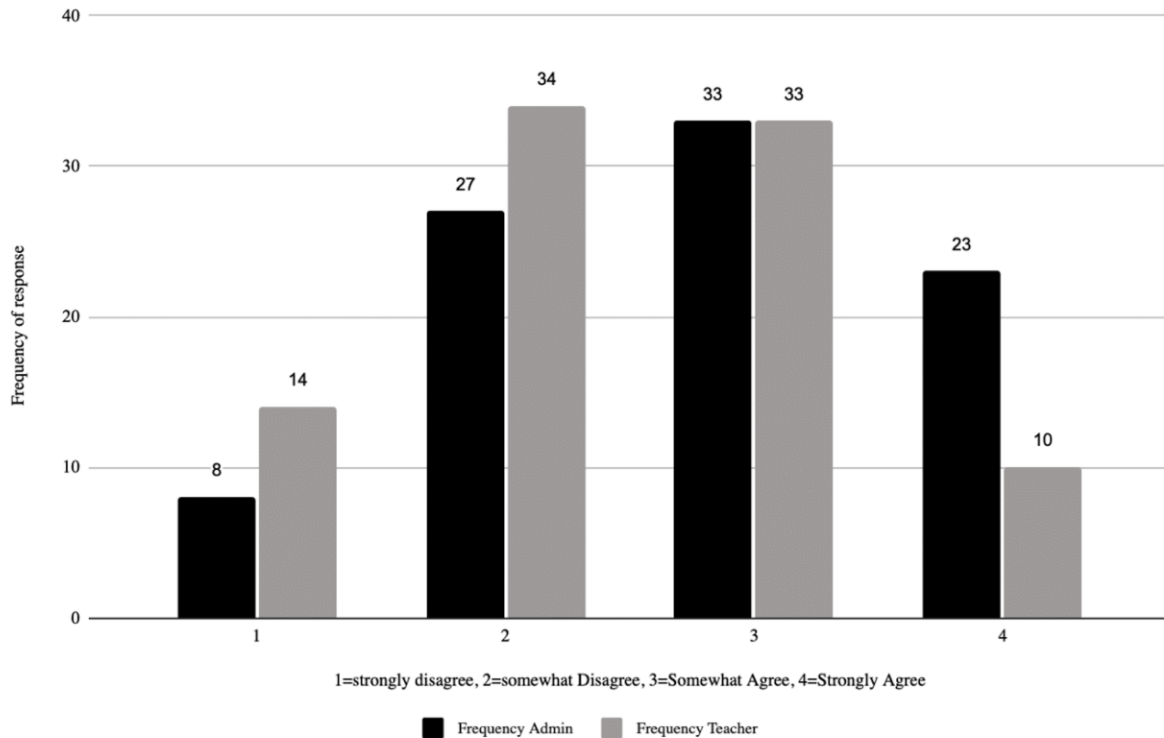
Distribution of Item H3-6 (“TPEP is useful but takes up too much of my time.”)



For item H3-7, respondents were asked to rate their level of agreement with the statement “I would like more time to work on my own evaluation.” Possible responses included strongly agree (4), somewhat agree (3), somewhat disagree (2), and strongly disagree (1). The mean response from the administrative perspective was 2.811 while the mean from the teacher perspective was 2.453. This indicates that administrators are more likely to need more time to work on their own evaluation than teachers are. The p-value for this item was .001, indicating that the difference between responses is statistically significant. Figure 4.8 shows the distribution of responses for item H3-7.

Figure 4.8.

Distribution of Item H3-7 (“I would like more time to work on my own evaluation.”)



Group H4-Usefulness for Professional Development

H40 – There is no statistically significant difference between educators’ perspective on TPEP regarding its usefulness for professional development from educators’ position as a teacher and as an administrator. Evidence collected from H4 items indicated that this null hypothesis is an invalid null hypothesis, as question H4-2 did yield a statistically significant result. Items in the H4 group were intended to investigate how teacher and administrator perspectives differed regarding TPEP’s usefulness for professional development. Six items (items 3–8) in group H4 (usefulness for professional development) were measured on a 5-point Likert scale, which can be found in Table 9. Scores for the administrator perspective items ranged from 3.813 on question H4-4 (professional conversations about what constitutes effective teaching) to 2.956 on question H4-7 (quality of relationships between the district and the teacher’s association). The overall mean for items measured on the 5-point Likert scale was 3.397, meaning that administrators

indicated that they think TPEP has had a slightly positive impact on professional development. Scores for the teacher perspective on the same items ranged from 3.692 on item H4-4 (professional conversations around what constitutes effective teaching) to 3.033 on item H4-7 (quality of relationships between the district and the teacher's association). The overall mean for the teacher perspective on the 5-point Likert scale items was 3.423, indicating that the perspective on TPEP's impact on professional development changed only slightly as respondents considered their transition from the classroom to administrative roles. Means calculated also revealed that respondents saw TPEP as having a slightly positive effect on professional development.

There were also three items in group H4 (H4-1,2, and 9) that were measured on a 4-point Likert scale that can be found in Table 7. Means for the administrator perspective items measured on the 4-point scale ranged from 2.915 on item H4-1 (the TPEP framework is a useful tool for improving my work) to 2.696 on item H4-9 (TPEP helps me align the improvement initiatives in my school). The overall mean for the administrator perspective items measured on a 4-point scale was 2.775, indicating respondents somewhat agreed that TPEP is useful for professional development. Means for the teacher perspective on the 4-point Likert scale ranged from 2.702 on item H4-1 (the TPEP framework is a useful tool for improving my work) to 2.415 on item H4-2 (the feedback I receive on my own evaluation is helpful for me). The overall mean for the teacher perspective items measured on a 4-point scale was 2.564, indicating that when responding from the teacher perspective respondents were slightly less likely to think TPEP was useful for professional development than when responding from the administrator perspective. The significance of Item H4-2 (the feedback I receive on my own evaluation is helpful for me)

being the lowest mean from the teacher perspective will receive further discussion in Chapter 5.

Table 4.11 shows all means, standard deviations, and p-values for each item in the H4 group.

Table 4.11.

Group H4 – Perceptions on TPEP’s Usefulness for Professional Development

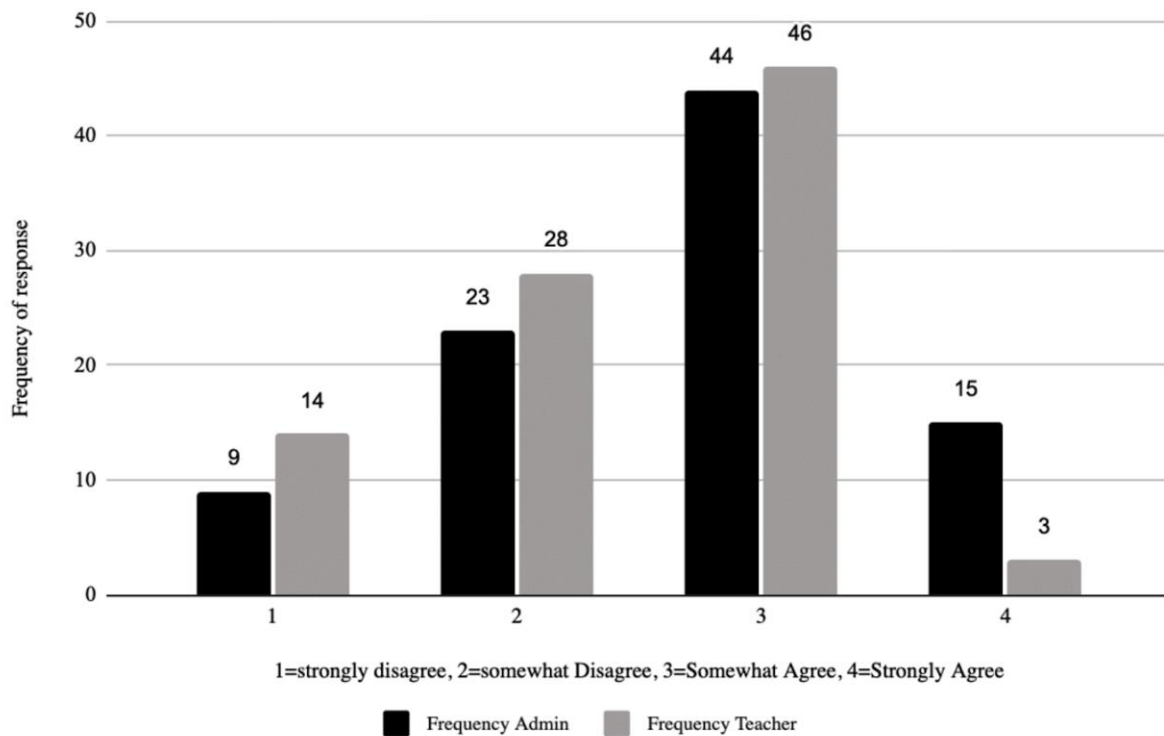
	Administrator		Teacher		
New item number	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>p</i>
H4-1	2.915	0.854	2.702	0.871	.073
H4-2	2.713	0.881	2.415	0.821	.003
H4-3	3.404	0.646	3.393	0.614	.887
H4-4	3.813	0.766	3.692	0.642	.153
H4-5	3.648	.711	3.549	0.685	.243
H4-6	3.626	.730	3.549	.652	.310
H4-7	2.956	0.799	3.033	0.733	.373
H4-8	3.297	0.670	3.319	0.645	.775
H4-9	2.696	0.810	2.576	0.785	.160
Overall H4	3.230		3.127		
5-point mean	3.397		3.423		
4-point mean	2.775		2.564		

For item H4-2, respondents were asked to rate their level of agreement with the statement “The feedback I receive on my own evaluation is helpful for me.” Possible answers included strongly agree (4), somewhat agree (3), somewhat disagree (2), and strongly disagree (1). The mean response for item H4-2 from the administrative perspective was 2.713 while the mean response from the teacher perspective was 2.415. The p-value for this question was .003,

indicating that the difference between the administrator and the teacher perspectives was statistically significant. Figure 4.9 shows the responses and distributions for item H4-2.

Figure 4.9.

Distribution of Item H4-2 (“The feedback I receive on my own evaluation is helpful.”)



Power Analysis for Paired *t*-Tests

A true power analysis could not be done for this sample population for two main reasons. First, there is no way to determine actual effect size because there is no previous population of individuals who have experience with TPEP as both teachers and administrators. In addition, the range of *n* was between 90 and 95 for the paired *t*-test questions, leaving a small sample size. When dealing with a small sample size and a limited body of previous research to establish a population, Cohen (1992) and Bliese and Wang (2019) suggested, it makes sense to use a power

analysis to more accurately and robustly determine the significance of resulting p values (Bliese & Wang, 2019; Cohen, 1992).

The observed power analysis data for questions showing significance can be found in Table 14.

Table 4.12.

Power and Effect Sizes for Items Rejecting the Null Hypothesis

Item #	p value	t value	Power ($1-\beta$)
H2-1	.001	3.639	.953
H2-4	.003	3.105	.874
H3-2	.000	-4.891	.998
H3-3	.000	-4.999	.999
H3-4	.000	-3.821	.969
H3-5	.000	-3.847	.970
H3-6	.001	3.341	.916
H3-7	.001	3.385	.923
H4-2	.003	3.009	.853

Power analysis data suggests that I do not have any false positives, as power values are well above the .5 power threshold.

Other Findings – Administrator Perspective

The administrative perspective portion of the survey contained questions that were similar to the teacher perspective questions but were different enough that a direct comparison

could not be made. Although a direct comparison cannot be made, responses to these questions will be addressed here, as they add both depth and breadth to the data set.

Item 18 – Division of Labor Among Administrators

Question 18 had six sub-questions that asked respondents to identify their level of agreement with the questions listed in Table 4.13 from an administrative perspective. The distribution of the responses is shown in Figure 4.10.

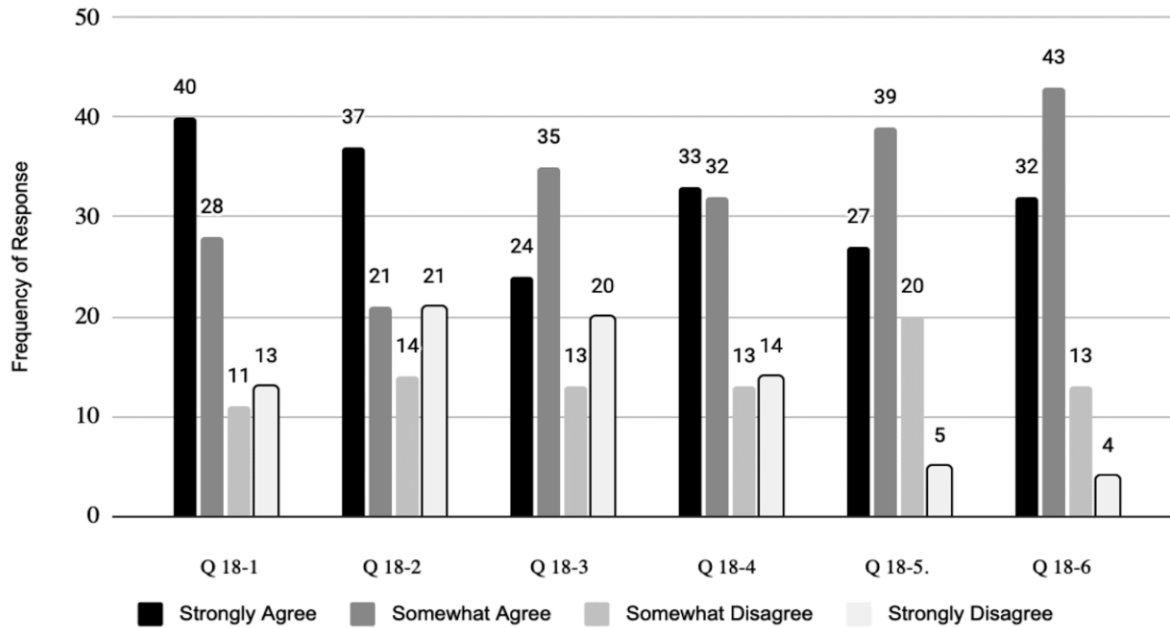
Table 4.13.

Questions 18.1–18.6

Question Number	To what extent do you agree with the following statements:
18-1	The number of teachers to be evaluated is divided equally among evaluators.
18-2	The principal is responsible for a larger portion of teachers to be evaluated.
18-3	The evaluator's subject matter expertise is taken into account when assigning teachers to evaluators.
18-4	Evaluators in the building regularly discuss best practices for teacher evaluation.
18-5	The quality of the feedback teachers receive in my building is similar regardless of who the evaluator is.
18-6	I am confident that teachers would receive similar scores regardless of who evaluates them.

Figure 4.10.

Division of Evaluation Duties



On item 18-1, 73.9% of respondents indicated that they either somewhat or strongly agreed that evaluation duties were distributed equally among administrative team members, while only 26.1% indicated that they somewhat or strongly disagreed with the statement. On item 18-2, 62% of respondents indicated that they either strongly or somewhat agreed that the principal took on a larger portion of the evaluation, while 38% indicated that they either somewhat or strongly disagreed. On item 18-3, 64.1% of respondents indicated that they strongly or somewhat agreed that subject matter expertise was taken into account when assigning evaluations, while 35.8% strongly or somewhat disagreed. On item 18-4, 70.7% of respondents indicated they strongly or somewhat agreed that they discuss best practices in evaluation, while only 29.3% strongly or somewhat disagreed. On item 18-5, 72.8% of respondents indicated they either strongly or somewhat agreed that the quality of the feedback teachers received would be similar regardless of who their evaluator was, while 27.5% strongly or somewhat disagreed. On item 18-6, 81.5% of respondents either strongly or somewhat agreed that the scores teachers

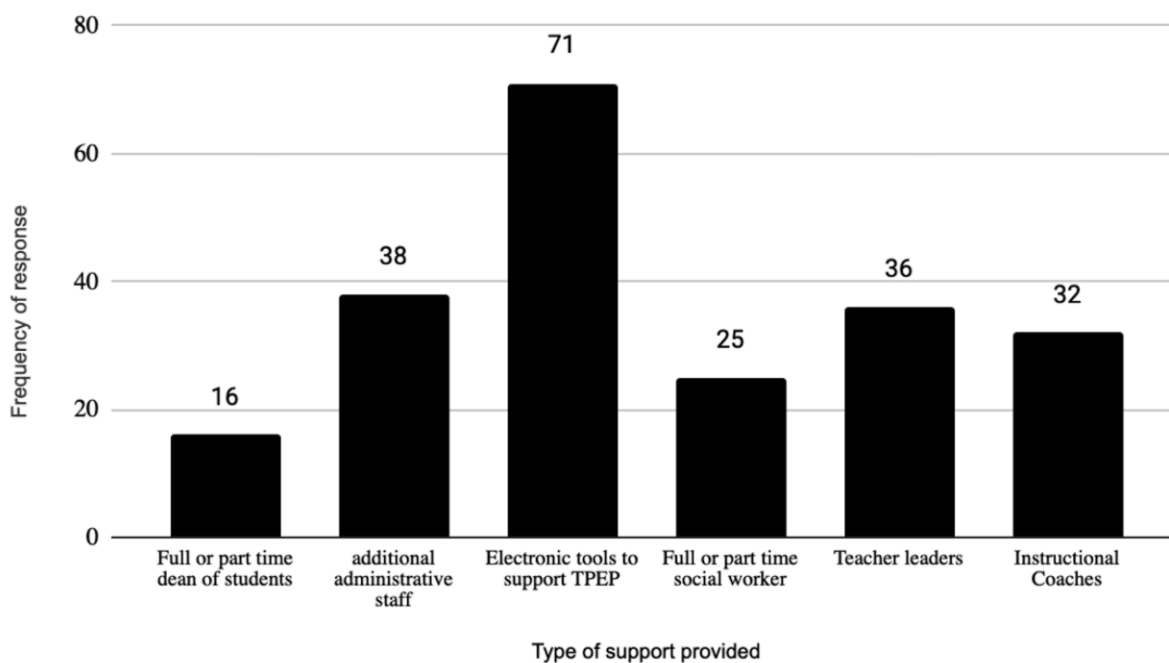
received would be similar regardless of the evaluator, while 18.4% strongly or somewhat disagreed.

Item 20 – Types of Supports Provided for TPEP

For item 20, respondents were asked to identify types of TPEP support that were present in their building from a check-all list including the following options: additional administrative staff, full- or part-time dean of students, full- or part-time social worker, instructional coach, teacher leaders, and electronic support tools. Figure 4.11 shows the types of supports provided and the frequency of response.

Figure 4.11.

Type and Number of TPEP Supports Provided



TPEP support varied widely, with some respondents indicating many different supports and others indicating few or none. Of the respondents surveyed, 26% indicated they had been able to hire a full- or part-time dean of students, 44.2% indicated that they had been given additional administrative staff, 80.6% indicated that they had used electronic tools to help

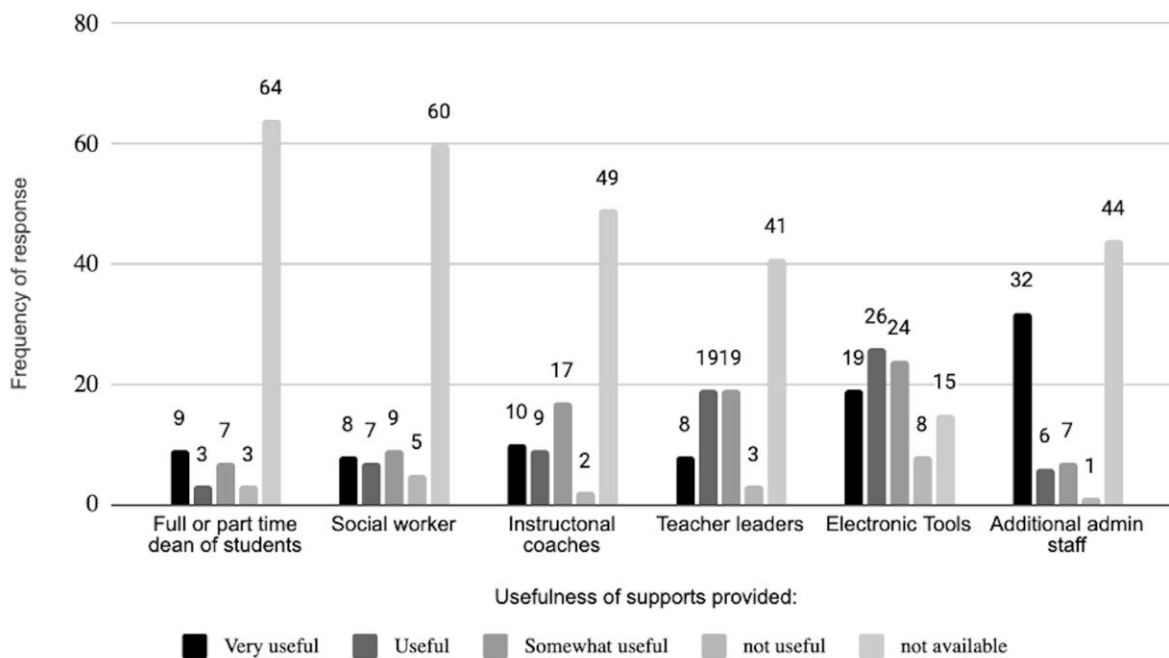
support TPEP, 29.1% were able to hire a full or part time social worker, 41.9% leveraged teacher leaders, and 37.2% were able to hire instructional coaches. Each of these interventions is intended to reduce the workload of administrative staff in a different way; they will all be discussed further in Chapter 5.

Item 21 – Usefulness of TPEP Supports

On survey item 21, respondents were asked to evaluate the effectiveness of each of the six identified TPEP supports on a 5-point Likert scale ranging from very useful to not useful. The question also included an option for not available. Figure 4.12 shows the responses.

Figure 4.12.

Usefulness of TPEP Supports Provided



On item 21-1, access to additional administrative staff, 51.1% of respondents indicated they were given additional administrative staff. Of those receiving additional administrative staff, 97.8% indicated that support was very to somewhat useful. It is worth noting that 69.6% indicated that additional administrative staff was very useful. On item 21-2, access to a full or

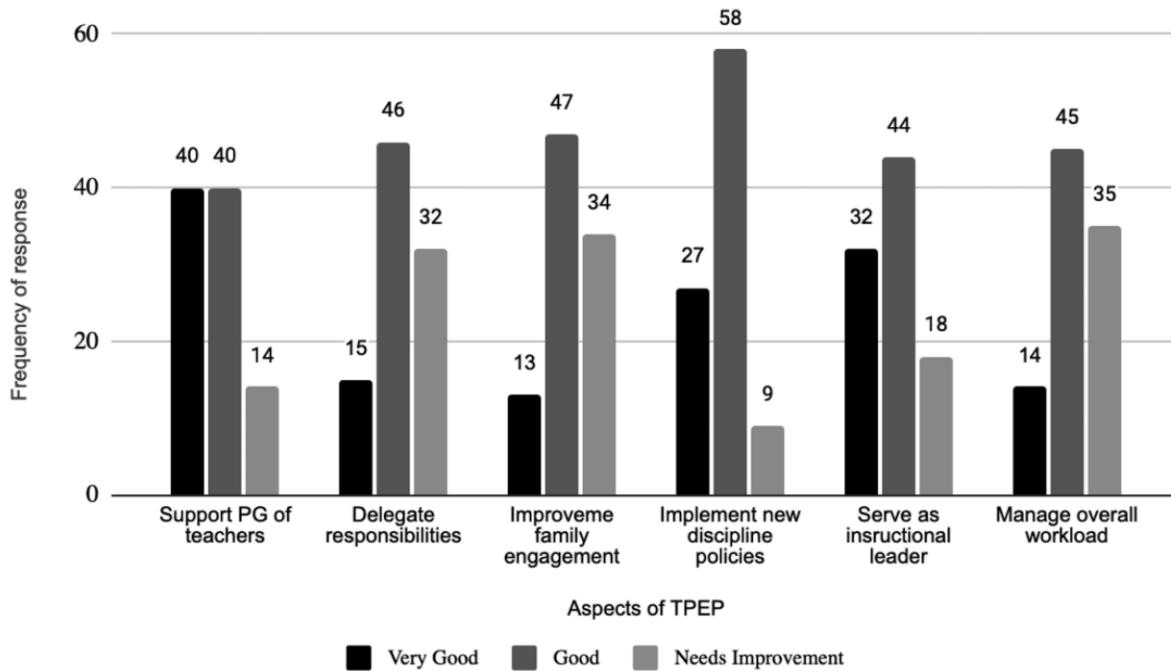
part time dean of students, 26% of respondents indicated that they had access to a full- or part-time dean. Of those having access to a full- or part-time dean of students, 86.4% indicated that it was very to somewhat useful, with 40.9% indicating it was very useful. On item 21-3, access to a full- or part-time social worker, only 32.6% indicated that they had access to a full- or part-time social worker. Of those respondents having access to a social worker, 82.8% indicated that it was very to somewhat useful, with only 27.6% indicating it was very useful. On item 21-4, access to an instructional coach, 56.3% indicating that they had access to instructional coaches. Of those who had access to instructional coaches, 94.7% indicated that their instructional coaches were very to somewhat useful, with 26.3% indicating their instructional coaches were a very useful support. On item 21-5, accessing teacher leaders, 53.9% indicated that they had access to teacher leaders. Of those having access to teacher leaders, 95.8% indicated their teacher leaders were very to somewhat useful, with 20.8% indicating they were very useful. On item 21-6, access to electronic supports, 83.7% of respondents indicated that they had access to electronic supports such as iObservation or eVAL. Of those having access to electronic supports, 89.6% indicated that the electronic supports were very to somewhat useful, with 24.7% indicating the supports were very useful. More attention will be given to the frequency and effectiveness of TPEP supports, as well as to the fiscal commitment involved in providing these supports, in Chapter 5.

Item 23 – Ability to Carry Out Various Aspects of Administrative Duties

Item 23 asked respondents to identify their ability to carry out various aspects of their administrative duties. Respondents were asked to rate themselves as *very good*, *good*, or *needs improvement*. The distribution can be found in Figure 4.13.

Figure 4.13

Ability to Carry Out Aspects of TPEP



Supporting the professional growth of teachers was a relative strength of respondents, with 42.6% indicating the support they provided was very good and 42.6% indicating that the support was good. When asked about their ability to delegate responsibilities, 16.1% indicated they were very good, 49.5% indicated they were good, and 34.4% indicated that they needed improvement. When asked about their ability to improve family engagement only 13.8% indicated they were very good, 50% indicated they were good, and 36.2% indicated they needed improvement. When asked about their ability to implement new discipline policies, 28.7% indicated they were very good, 61.7% indicated they were good, and 9.6% indicated they needed improvement. When asked about their ability to serve as an instructional leader, 34% indicated they were very good, 46.8% indicated they were good, and 19.1% indicated they needed improvement. When asked about their ability to manage workload, 14.9% indicated they were

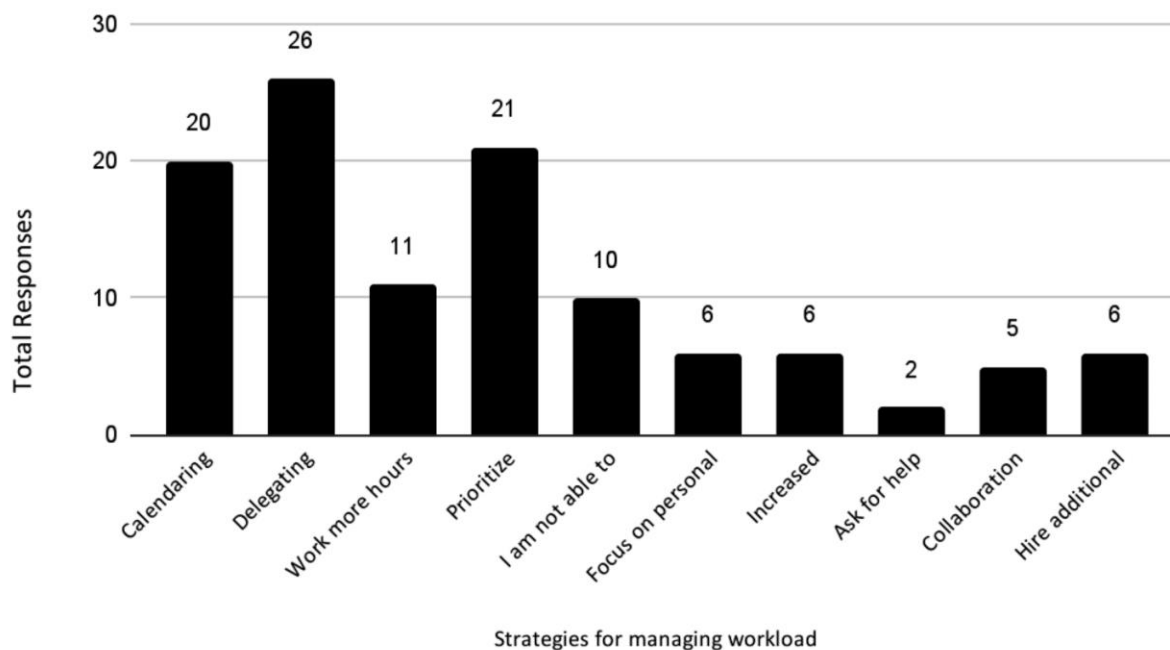
very good, 47.9% indicated they were good, and 32.2% indicated they needed improvement. The implications of these perspectives will be further discussed in Chapter 5.

Item 24 – Strategies for Workload Management

Item 24 asked administrators what strategies they used for workload management. Respondents were allowed to name as many strategies as they wished. Answers were categorized into 10 different categories. Categories and answer frequency are shown in Figure 4.14.

Figure 4.14.

Strategies for Managing Workload



The most frequent responses included delegating (26), prioritizing (21), and calendaring (20). Of the 78 administrators who answered this question, 10 indicated that they were not able to effectively manage their increased workload, and another 11 indicated that the main strategy they used was to spend more hours at work. Both of these data points are concerning and will be discussed further in Chapter 5.

Item 29 – Desired Areas for Training

Item 29 asked respondents to identify types of training they would like to have to help with the implementation of TPEP. Respondents indicated a desire for support in two primary areas. The first area was collaboration time with teachers or administrative colleagues. Of the 44 total responses, 18 indicated that the respondents would like to have time to collaborate with teachers/colleagues. The areas most identified for collaboration were general support (13 responses), inter-rater reliability (7 responses), and time with teachers (1 response). The second area of support identified was area-specific training. Specific areas of support identified were general framework training or refreshers (7 responses), student growth goals (5 responses), collection or artifacts (4 responses), communicating the intent of TPEP (2 responses), and the use of online tools (2 responses). All other responses received only 1 response and therefore were not included.

Item 30 – Helpfulness of Feedback for Performing Specific Duties

Item 30 asked respondents how helpful the feedback they received from TPEP evaluations was for performing specific aspects of their position. The question was broken down into four sub questions about planning professional development, identifying teacher leaders, assessing progress on the school improvement plan, and improvement as an instructional leader. Each sub question will be addressed individually below.

Item 30-1 – Planning Professional Development. Item 30-1 asked respondents to what extent the feedback they received from conducting TPEP evaluations in their building assisted with planning their professional development. Of the 92 individuals who responded to the question, 18.5% indicated the feedback assisted a great deal, 55.47% indicated it assisted somewhat, 15.2% indicated it assisted a little bit, and 10.9% indicated it did not assist at all (see

Figure 4.15). That 89.1% of respondents indicated the feedback from TPEP evaluations assisted in planning professional development is significant, as that response lines up with the mission of the TPEP steering committee (S. Anderson, November 25, 2018). The use of TPEP feedback for planning professional development will be discussed further in Chapter 5.

Item 30 – 2-Identifying Teacher Leaders. Item 30-2 asked respondents to what extent the feedback they received from conducting TPEP evaluations in their building assisted with identifying teacher leaders. Of the 92 individuals who responded to the question, 38.2% indicated the feedback assisted a great deal, 34.8% indicated it assisted somewhat, 15.2% indicated it assisted a little bit, and 9.9% indicated that it did not assist at all (see Figure 4.15). That 90.1% of respondents indicated TPEP assisted in helping identify teacher leaders at least a little bit is significant and will be discussed further in Chapter 5.

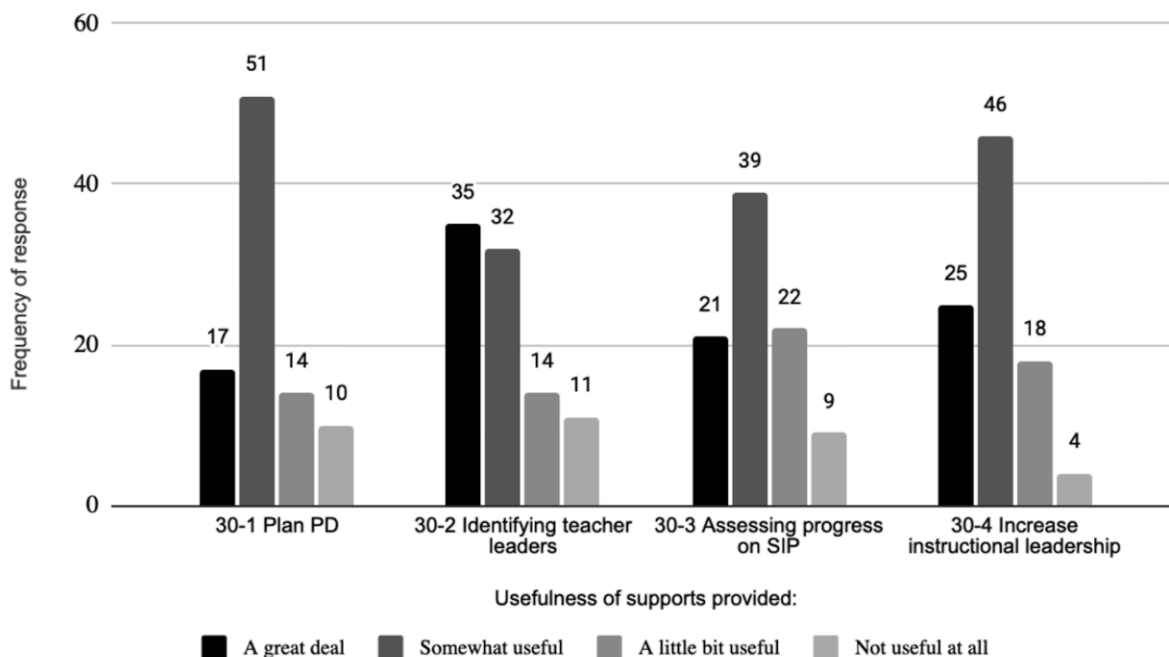
Item 30-3 – Assessing Progress on School Improvement Plan (SIP). Item 30-3 asked respondents to what extent the feedback they received from conducting TPEP evaluations in their building assisted with assessing progress on their school improvement plans. Of the 91 individuals who responded to the question, 23.1% indicated it assisted a great deal, 42.9% indicated it assisted somewhat, 24.2% indicated it assisted a little bit, and 9.9% indicated that it did not assist at all (See Figure 4.15). That 90.2% of respondents indicated TPEP assisted in assessing their progress on school improvement plans is significant and will be addressed further in Chapter 5.

Item 30-4-Improvement as an Instructional Leader. Item 30-4 asked respondents to what extent the feedback they received from conducting TPEP evaluations in their building assisted with their improvement as an instructional leader. Of the 93 individuals who responded to the question, 26.7% indicated the feedback assisted a great deal, 49.4% indicated it assisted

somewhat, 19.4% indicated it assisted a little bit, and 4.3% indicated that it did not assist at all (See Figure 4.15). That 95.7% of respondents indicated TPEP assisted their growth as an instructional leader is significant and will be addressed further in Chapter 5.

Figure 4.15.

Usefulness of Feedback From TPEP



Other Findings – Teacher Perspective

The teacher perspective portion of the survey contained questions that were similar to the administrator perspective questions but were different enough from them that a direct comparison could not be made. Although a direct comparison cannot be made, responses to these questions will be addressed here, as they add both depth and breadth to the data set.

Item 38 – Type of Supports Provided

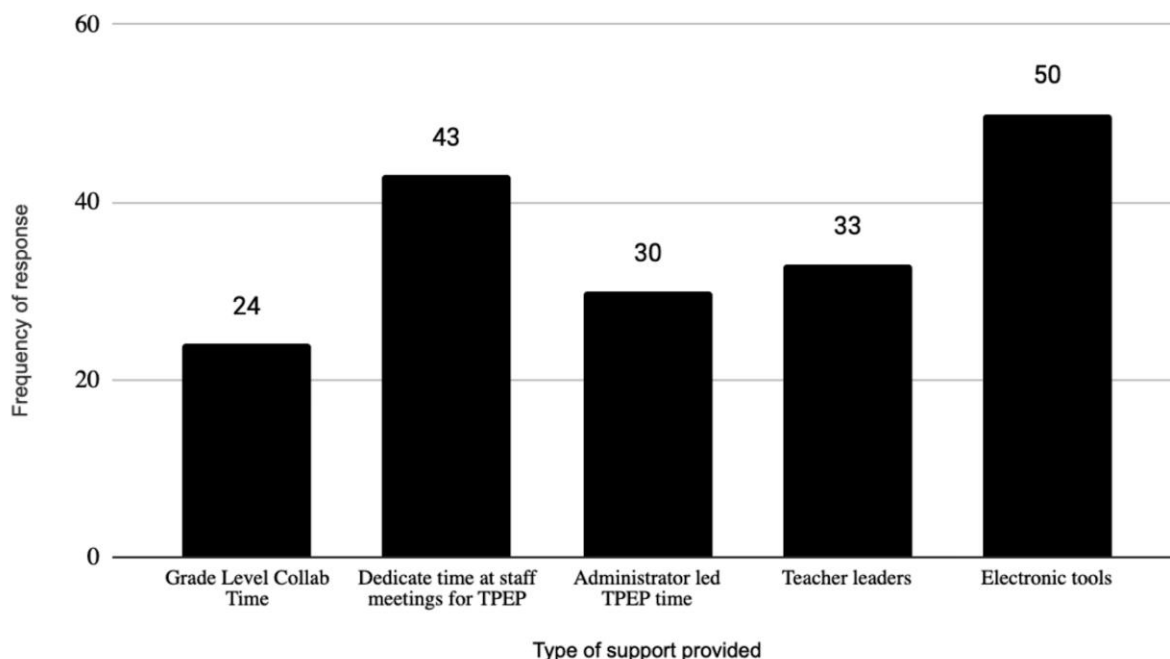
For item 38, respondents were asked to identify types of TPEP supports present in their building from a check-all list including the following options: grade level collaboration time,

dedicated time at staff meetings, administrator-led TPEP time, designated teacher leaders, and electronic tools to support TPEP, including eVAL, Google Docs, iObservation, or other tools.

Figure 4.16 shows the types of supports provided and the frequency of responses.

Figure 4.16.

Type and Number of TPEP Supports Provided – Teacher Perspective



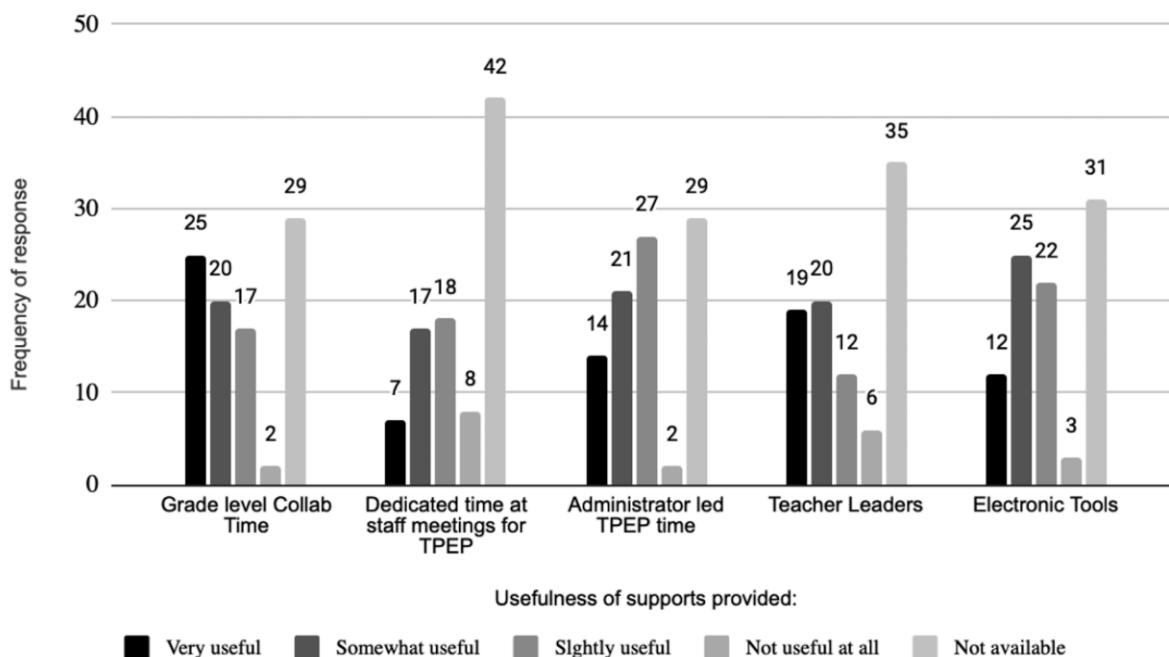
TPEP support varied widely, with some respondents indicating many different supports and others indicating few or none. Of the respondents surveyed, 28.9% indicated they had been given grade-level collaboration time, 51.8% indicated that they had been given dedicated time at staff meetings, 36.1% indicated that they had administrator-led TPEP time, 39.6% had access to teacher leaders, and 60.2% had access to electronic tools. The fact that only 60.2 respondents had access to electronic tools is noteworthy and will be discussed further in Chapter 5.

Item 39 – Usefulness of TPEP Supports

On survey item 39, respondents were asked to evaluate the effectiveness of each of the five identified TPEP supports on a 5-point Likert scale ranging from very useful to not useful. The question also included an option for not available. Figure 4.17 shows the responses.

Figure 4.17.

Usefulness of TPEP Supports Provided



On item 39-1, grade-level collaboration time, 68.1% of respondents indicated they were given time to collaborate with grade-level colleagues. Of those receiving grade-level collaboration time, 39% found it very useful, 31.3% found it somewhat useful, 26.6% found it slightly useful, and 3.1% found it not at all useful. On item 21-2, dedicated time at staff meetings, only 54.3% of respondents reported having access to dedicated time at staff meetings. Of those that had dedicated time, 14% found it very useful, 34% found it somewhat useful, 36% found it slightly useful, and 16% found it not useful at all. On item 31-3, administrator-led TPEP time, 68.1% of respondents indicated that they had this administrator-led time. Of those respondents who had administrator-led time, 21.9% found it very useful, 32.8% found it

somewhat useful, 29% found it slightly useful, and 3.2% found it not useful at all. On item 31-4, access to teacher leaders, 62% of respondents indicated they had access to support from teacher leaders. Of the 62% that did receive support from teacher leaders, 33.3% indicated it was very useful, 35% indicated it was somewhat useful, 21.1% indicated it was slightly useful, and 10.5% indicated it was not useful at all. On item 39-5, access to electronic tools, 66.7% of respondents indicated they had access to electronic supports. Of those respondents, 19.4% indicated the electronic tools were very useful, 40.3 found it somewhat useful, 35.5% found it slightly useful, and 4.8% selected not useful at all. Given the ready and free access to electronic tools, it is troubling to see that one third of all educators surveyed were without access to these tools while they were classroom teachers. This topic will be discussed further in Chapter 5.

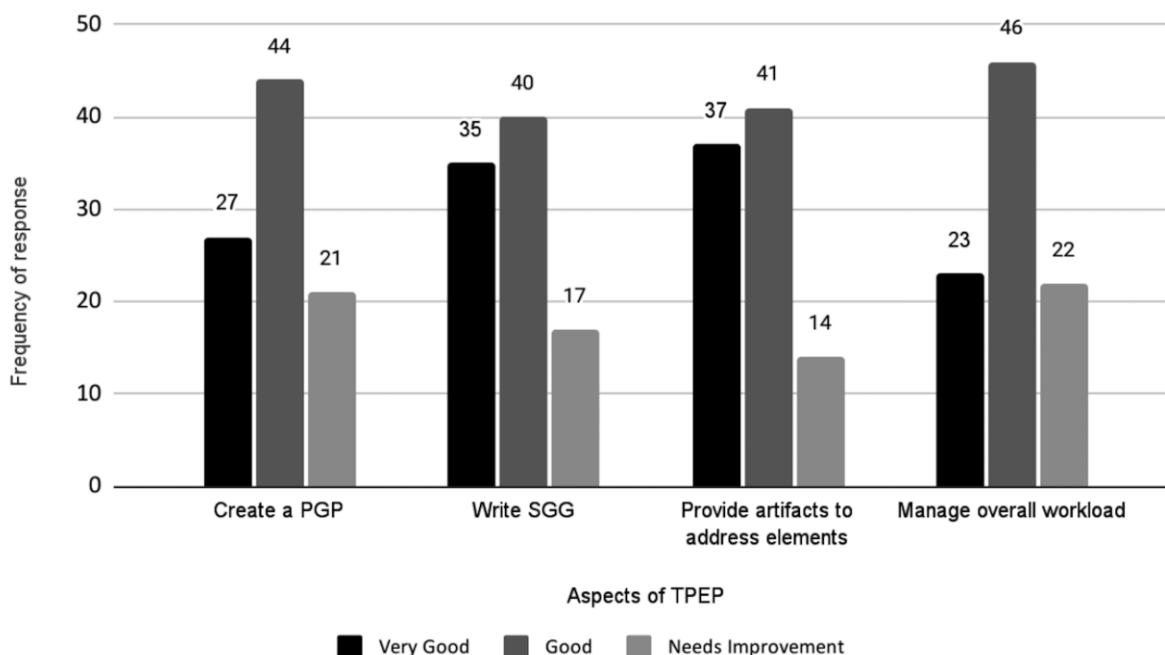
Item 41 – Ability to Carry Out Aspects of TPEP

Item 41 asked respondents to rate their ability to carry out the different aspects of TPEP. Respondents were asked to respond with *very good*, *good*, or *needs improvement* to each segment of the question. Response distributions and frequencies are found in Figure 4.18.

Of the 92 respondents surveyed, 77.1%, indicated they were either very good or good at writing a professional growth plan (PGP). When asked about their ability to create a PGP, 29.3% of respondents indicated they were very good at creating a PGP, 47.8% of respondents indicated they were good at creating a PGP, and 22.8% of respondents indicated they were needing improvement in creating a PGP.

Figure 4.18.

Ability to Carry Out Aspects of TPEP



A vast majority of respondents, 81.5%, indicated that they were either very good or good at writing student growth goals (SGG). When asked about their ability to write SGG, 38% of respondents indicated they were very good at creating SGG, 43.5% indicated they were good at writing SGG, and 18.4% indicated they were needing improvement in writing SGG.

A vast majority of respondents, 84.8%, indicated that they were either very good or good at providing artifacts to support TPEP elements. When asked about their ability to provide artifacts to support TPEP elements, 40.2% indicated they were very good at providing artifacts, 44.6% indicated they were good at providing artifacts, and 15.5% indicated they needed improvement.

The smallest majority of respondents, 75.8%, indicated that they were either very good or good at managing their workload. When asked about their ability to manage their workload, 25.5% of respondents indicated they were very good at managing their workload, 50.5%

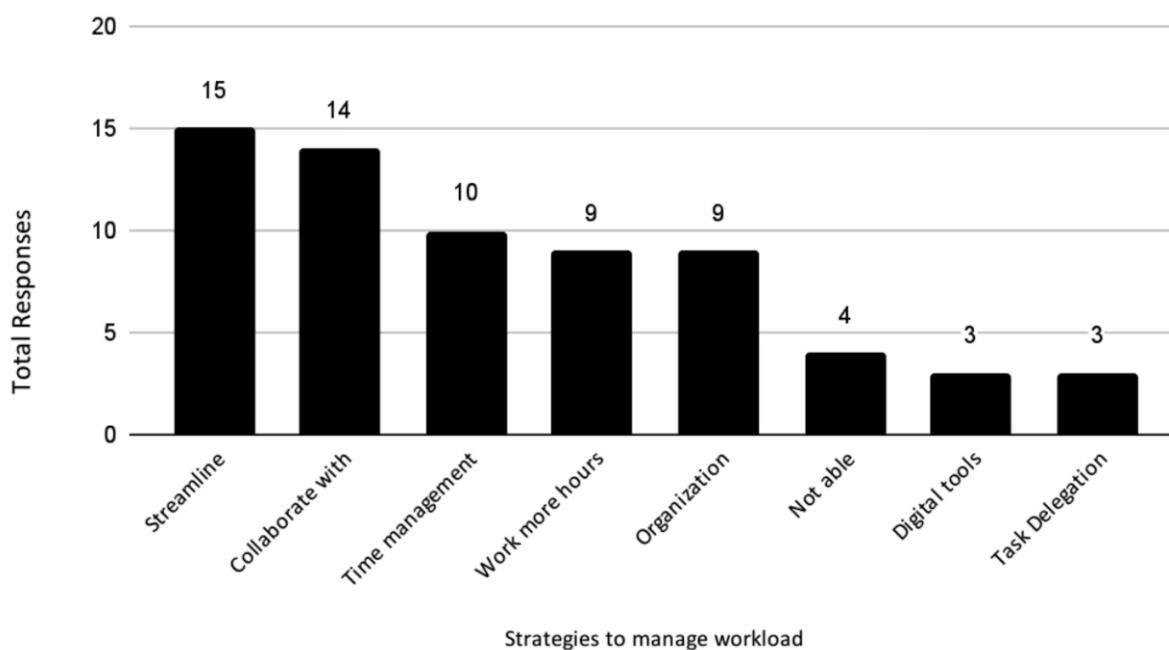
indicated they were good at managing their workload, and 24.2% indicated they needed improvement at managing their workload.

Item 42 – Strategies for Managing Workload, Teacher Perspective

Item 42 asked respondents what strategies they used for workload management as a teacher. Respondents were allowed to name as many strategies as they wished. Answers were broken down into eight categories. Categories and answer frequency are shown in Figure 4.19.

Figure 4.19.

Strategies to Manage Workload, Teacher Perspective



A total of 68 responses were collected for item 42. Among the teacher strategies used by respondents, taking measures to streamline their classroom, a strategy used by 22.1%, was the most common response. Measures taken to streamline classrooms included only grading necessary assignments, focusing on only the essential standards, and simplifying assignments.

The second most common strategy was to collaborate with colleagues. Just 20.6% of respondents indicated that they collaborated with colleagues. Some indicated that they did so

through professional learning communities (PLCs), while others had less formal ways of collaborating. Time management was also a common response (14.7%). Time management strategies used by teachers included maximizing time on task by using entry and exit tasks, calendaring intentionally, and setting times to leave the building at a reasonable hour. Working extra hours and increasing organization were also strategies used by 13.2% of respondents. These strategies included spending time in the evenings and on weekends. Other strategies used included digital tools, such as Google Classroom, and task delegation, both of which were identified by 4.4% of the respondents. A small number of respondents (5.9%) also indicated that they had not effectively implemented any strategies to help alleviate the workload. The percentage of respondents who indicated that they had not been able to successfully manage the workload as a teacher is less than half that of the administrative perspective. This concern will be addressed further in Chapter 5.

Item 45 – Training Desired, Teacher Perspective

Item 45 asked respondents to identify types of training they would like to have to help with the implementation of TPEP. Of the 95 respondents, 50 offered an answer to this question. Their answers fell into three different categories, all of which were similar to the answers to the paired administrative question. The first category of responses centered on elements of TPEP. Respondents indicated they would like training in the general framework of TPEP (9 responses), writing student growth goals (6 responses), collecting and providing artifacts (3 responses), the background of TPEP (2 responses), and the electronic tools used to carry out and track TPEP (1 response). The second category of training desired centered on the idea of collaboration time. Ten respondents indicated that they would like time with colleagues to collaborate on TPEP, and four respondents indicated they would like more training with either an instructional coach or

their evaluator. The third category focused on best teaching practice (7 responses). Given that the purpose of TPEP was to help improve teaching practices, it is encouraging that respondents from the teacher perspective saw it as a training tool to help develop their practice. A small group of respondents (2 responses) indicated that role or content-specific training would have been beneficial to them.

Summary

The purpose of this study was to find out whether current administrators' perspectives on TPEP had changed since they moved from the classroom to an administrative role. With this study, I also sought to determine the impact of TPEP in four different areas: usefulness of TPEP supports, TPEP's impact on perception of career choice, TPEP's impact on perceived workload, and TPEP's usefulness for professional development. A three-part survey was distributed to all evaluating administrators in Washington State that contained 24 paired questions. A t-test was run on the 24 questions to discern whether answers were different in a statistically significant way. Of the 24 paired questions, nine of those showed a statistically significant difference between the teacher and administrator perspective responses. The survey also included questions that were not perfectly paired but added to the depth and breadth of the data. Those data were presented and will be further discussed in chapter five.

\

Chapter 5 - Discussion of Findings

The purpose of conducting this study was to investigate whether educators' perspectives on TPEP had changed in four key areas as they moved from the classroom to an administrative role. The investigation began when I distributed a survey to all evaluating administrators in Washington State asking them to compare their perspectives on TPEP from their last year as a teacher and as an administrator. Once all responses were collected, results to the paired questions were compared using a t-test. Forty percent of the paired questions show a statistically significant difference.

This study reveals that, from the administrator perspective, the impact of TPEP has been at least slightly positive in regard to every research question except for question 3: impact on workload (see Table 5.1). Respondents' answers indicated that they were less affected by TPEP implementation when they were teachers than when they were administrators. Insights from this study can be used to provide training to both teachers and administrators to help increase the usefulness of TPEP for professional development, help manage the increased workload, and increase the application of TPEP training to school level professional improvement. All of these improvements will help increase job satisfaction and the desire to stay in the profession. These results can also be used to help prepare pre-service teachers and administrators in preparatory programs to engage in the collaborative evaluation process as they enter the workforce as certificated practitioners.

Table 5.1.

Overall Mean and Standard Deviation of H1–H4

Group	Administrative Perspective		Teacher Perspective	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
H1	2.116	1.42	2.058	1.34
H2	2.323	0.94	2.215	0.90
H3	2.706	0.98	2.865	0.79
H4	3.230	0.86	3.137	0.85

Implications Based on Findings

Demographics –

At the suggestion of the committee, multiple attempts were made to obtain accurate statewide demographic data from multiple organizations. OSPI, the University of Washington, AWSP, and WASA were all contacted and asked to provide statewide demographic to compare to my response sample. This was done in order to determine how representative my sample was of the population of all administrators in Washington State. After multiple attempts only one organization, WASA, was able to respond with useful data. Unfortunately, out of the 1,314 members WASA has only 62 of them report their jobs as principal, assistant principal, dean of students, or principal's assistant. The remainder of the WASA membership serves in some sort of central office role that would not typically include evaluation duties using the TPEP framework. Data from my sample set indicated that 91.4% while only 61.3% of the WASA members identified as white. This discrepancy is misleading as 35.4% of the WASA membership chose not to identify their race or ethnicity. The obvious implication of this is that

those WASA members that did not respond could change the demographic percentages significantly. If many of the members that chose not to identify their race were white then the WASA data could be in line with my survey respondents. Since this is impossible to know it is a limitation of the study and does not help alleviate the problem of small sample size.

The gender data was slightly more helpful. Respondents for my survey were 33.6% male, 64.2% female, and only 2.1% chose not to respond. The WASA data indicated that 38.7% of respondents were male, 43.5% were female, and 17.8% chose not to respond. While this data is a bit more useful, indicating a higher percentage of male administrators, it is still problematic because of the relatively large portion of WASA members who chose not to reveal their gender. Again, the data from the WASA membership did little to alleviate concerns regarding the small sample size in this study and remains a primary limitation.

H1 – Usefulness of Supports

TPEP is a statewide evaluation framework that was introduced as a pilot program during the 2010–11 school year in eight districts across the state and one consortium of eight smaller districts from ESD 101, which is located in eastern Washington (TPEP, 2013). The establishment of TPEP came with a variety of supports that were available to different districts in different capacities. Elfers and Plecki (2017) explained that these supports were important because of the complexity of the initiative. Supports identified in this survey included training to support rater reliability, district-provided professional development, collaboration with colleagues, and sufficiency of support concerning implementation of teacher and administrator evaluation.

Of the four questions in group H1, all means were below 2.67, meaning that the usefulness of the TPEP support received was only somewhat useful at best. Although none of

these questions exhibited *t*-test results that were statistically significant, the fact that respondents indicated that TPEP supports were only somewhat useful from both the teacher and the administrative perspectives is useful for guiding future practice.

Usefulness and Availability of Administrator Supports

Item 21 provides some useful insight into group H1 questions from the administrator perspective, as it addressed specific supports provided and respondent perception of each support's usefulness. The supports addressed in the survey were additional administrative staff, full- or part-time dean of students, full- or part-time social worker, instructional coaches, teacher leaders, and electronic tools. As might be expected, the most useful of these supports was additional administrative staff. Over the past 10 years, the presence of additional administrative staff has increased significantly (WASA, 2021). Of respondents surveyed, 51.1% indicated having additional administrative staff. Of those who had additional administrative staff, 97.8% indicated that it was very to somewhat useful, with 69.9% indicating it was very useful. Other supports requiring significant fiscal commitments also were seen as useful to different degrees (See Figure 12). With the exception of electronic tools, these supports were reportedly available to less than 60% of respondents. Supports besides additional administrative staff seen as very useful but available to less than 60% of respondents included full- or part-time dean of students (available to 26% of respondents, with 40.9% indicating it was very useful and 86.4% selecting very to somewhat useful); full- or part-time social worker (available to 32.6% of respondents, with 27.6% indicating it was very useful and 82.8% selecting very to somewhat useful); instructional coaches (available to 56.3%, with 26.3% indicating it was very useful and 94.7% indicating it was very to somewhat useful); and teacher leaders (available to 59.3%, with 20.8% indicating it was very useful and 95.8% choosing very to somewhat useful). The most common

support provided was electronic tools (available to 83.7% of respondents). Nearly 25% of respondents indicated that these electronic tools were very useful for supporting their TPEP work, and 89.6% indicated that they were very to somewhat useful.

These data seem to contradict the H1 group mean for the administrator perspective. The mean for group H1 questions was 2.116, slightly higher than the teacher perspective mean of 2.058, indicating that respondents felt that TPEP supports were either slightly useful or somewhat sufficient. These means are lower because respondents who indicated a support was not available were given a score of 1, which would drive the mean down. The implication here is that supports, when present, are useful to administrators in their practice. Given this insight, it is incumbent upon districts to find ways to maintain supports and allocate resources effectively to implement TPEP with fidelity.

Usefulness and Availability of Teacher Supports

Item 39 provides some insight into group H1 questions from the teacher perspective, as it addresses specific supports provided and respondent perception of each support's usefulness. Specific supports investigated included collaboration time, dedicated time at staff meetings, administrator-led TPEP time, access to teacher leaders, and e-tools.

The power of focused educator collaboration is widely documented and is one of the "State 8" evaluation criteria (WAC 392-191A-060, 2013). The power of collaboration is reflected in all five supports provided for TPEP that were investigated in this study. Of the collaboration supports provided, collaboration time with colleagues was reportedly the most helpful. Although only 68.8% of respondents indicated having access to grade-level collaboration time, 39% of respondents indicated that it was very useful and 96.9% of respondents indicated that it was at least slightly useful (See Figure 17). Teacher leaders were

also seen as a useful support for teachers in their implementation of TPEP. Of the respondents surveyed, only 62% had access to trained teacher leaders for TPEP support. Of the respondents who had access to teacher leaders as a support, 33% indicated it was very useful and 89.5% indicated it was at least slightly useful (See Figure 17). Electronic tools have also been provided to support TPEP implementation. Although only 66.7% of respondents indicated that they had access to these electronic tools, 19.4% of those who did have access indicated that they were very useful and 95.2% indicated that the electronic tools were at least slightly useful (See Figure 17). Administrator-led TPEP times were also seen as useful. Although only 68.8% of respondents indicated that they had access to administrator-led TPEP support, 21.9% indicated it was very useful and 96.9% indicated that it was at least slightly useful (See Figure 17). The final type of collaborative support provided was dedicated time at staff meetings. Only 54.3% of respondents indicated they did have dedicated time for TPEP at staff meetings, but 14% found it very useful and 84% found it at least slightly useful (See Figure 17).

Although these data seem to contradict the teacher perspective mean for group H1 of 2.058 (which corresponds to either somewhat disagree with the statement that support for TPEP is sufficient or the support was slightly useful), respondents who indicated the support was not available received a score of 1, which would drive the mean down. Supports, when applied with fidelity, seem to increase the perception that educators feel they have sufficient support for TPEP. Of the five supports investigated in item 39, only one (teacher leaders) requires fiscal support. Districts all across the state have implemented schedules that include grade-level or content-area collaboration time through a 4-day school week, early release, or late start. Staff meetings and administrator-led PD time are built into teacher contracts and offer the opportunity to support staff in TPEP if it is made a priority. Since 2013, Washington has provided eVAL

free of charge. eVAL is a secure, free, continuously updated resource that can be customized to each district's framework of choice (OSPI, n.d.c). There are even training videos and a "sandbox" where educators can practice using the system before being formally evaluated (OSPI n.d.c). The fact that no more than 70% of respondents have access to these supports contributes greatly to the perception that TPEP supports are only slightly useful. Further investigation of why districts are not accessing free resources and prioritizing staff development via TPEP is needed.

H2 – Perception of Career Choice

Research question number 2 focused on perception of career choice. The questions that were chosen for this section were all selected because they connect to elements of educator job satisfaction, which is an indicator of positive career choice perception. Question H2-1 asked respondents how TPEP had impacted their interactions with colleagues. Darling-Hammond (2010) sought to identify working conditions that increase teacher job satisfaction and retention and concluded: "Most important are the conditions that teachers feel enable them to succeed with students—including administrative supports, strong colleagues, and opportunities to participate in decisions" (p. 21). Johnson and colleagues (2012) found that teachers' job satisfaction was much less tied to the demographic of students they served or the community their school was located in than it was to positive school culture, leadership, and collegial relationships. The *p*-value for item H2-1 was .001, which is lower than the .05 that indicates statistical significance for the question. The administrative mean of 2.6 was greater than the teacher mean of 2.3, indicating that the administrative perspective on TPEP's impact on their interactions with colleagues was more positive than the teacher perspective. Increasing the quality of interactions related to evaluation and teacher practice is a significant accomplishment, as that factor has been

connected to teacher satisfaction and retention and will potentially have a positive impact on student achievement (Darling-Hammond, 2010; Johnson et al., 2012).

Another source of job satisfaction is autonomy in the workplace. In a brief from the National Governors Association in 2005, Barnett and Hirsch indicated that one of the things that schools in hard-to-staff districts could do to increase teacher retention is to provide more autonomy. They reported, “Teachers feel that they have no input in the design and organization of schools and that they have minimal prospects for career advancement and professional growth” (p. 5). Item H2-2 asked whether respondents had the flexibility to implement TPEP in the way that works best in their school. Both administrative (2.663) and teacher (2.669) means indicated that respondents somewhat agreed that they had the ability to implement TPEP in a way that worked best in their school. Although the means suggest that there was some level of agreement with this conclusion, there is certainly room for growth. Especially in hard-to-staff districts and buildings, district- and building-level administrators would do well to allow building-level practitioners the freedom to implement TPEP in manners that are most appropriate for their individual buildings. Examples of doing so could include providing more TPEP supports in buildings with new administrators and higher concentrations of younger teachers, giving latitude in student growth goal setting and achievement targets, providing more support for professional growth plan development, and providing weekly TPEP support drop-in meetings directed by administrative staff or teacher leaders who have been trained in the district framework.

Item H2-3 asked respondents whether the majority of time they spent on TPEP was useful. Drew and Sonowski (2019) underscored the connection between purposeful work and job satisfaction. They found that teachers who have a sense of purpose can use it to help them be

resilient in the face of difficult factors and to get the most out of positive ones (Drew & Sonowski, 2019). Comparison of administrative and teacher means for question H2-3 did not show a statistically significant difference between them. The administrative perspective mean (2.283) and the teacher perspective mean (2.380) both indicated that they somewhat disagreed that the majority of the work they did on TPEP was useful. This is concerning, because when educators are forced to engage in work that does not align with their purpose, this can have a negative impact on job satisfaction and can lead to the educators' leaving the profession. Potential solutions to the problem of a disconnect between teaching practice and the work of TPEP include making intentional connections between the framework and everyday teaching practice through regular in-person and digital communication; repeatedly modeling element-specific practices; collecting data that connects TPEP practices to student achievement; establishing TPEP focus criteria or elements each month; and giving appropriate rewards for application of TPEP focus practices.

Item H2-4 asked respondents directly to indicate their level of agreement with the statement "TPEP has made me consider leaving my position." Intent to leave the profession is a clear indicator of lack of job satisfaction/career choice, and this question encompasses all other items in group H2. The *p*-value of .003 is less than the established threshold of .05, indicating the means for administrative and for teacher perspectives were different in a statistically significant way. The administrative mean (1.747) indicates that respondents leaned toward somewhat disagreeing with the statement "TPEP has made me consider leaving my position," while the teacher mean (1.473) indicates a higher tendency toward strong disagreement with the statement. Both of these results are encouraging. The TPEP initiative is an initiative that is not going away in the near future, and current and future educators seem to be adjusting to it

appropriately. Elfers and Plecki (2019) have expressed concern over the impact of TPEP on administrators, especially when they are the sole evaluator in a school or district (Elfers & Plecki, 2019). Concerns surround issues of workload and training and will be discussed further in later sections.

A final point of discussion for item group H2 is the difference between the mean for item H2-4 and that for the rest of the group. Questions H2-1–H2-3 had means ranging from 2.663 to 2.283 (a difference of .38) for administrator perspective and from 2.669 to 2.331 (a difference of .338) for the teacher perspective. Item H2-4 was truly an outlier, as its mean was significantly below that of the rest of the group: Those means were .536 and .858 respectively. The low means for item H2-4 indicate that, although TPEP requires additional training, evaluative rigor, and workload, it is not having a significant impact on the way respondents view their career choice.

H3 – TPEP’s Impact on Perceived Workload

Items in group H3 were connected to TPEP’s impact on perceived workload. H3 items showed the highest number of statistically significant results, with six out of seven items showing statistical significance. Elfers and Plecki (2019) emphasized the need to help administrators manage the increased workload of TPEP and recommended doing so in the following ways: adding additional administrative staff, hiring a full- or part-time dean of students, hiring a full- or part-time social worker, using instructional coaches, using teacher leaders, and accessing electronic tools to organize and track all the different elements of TPEP. Elfers and Plecki also drew particular attention to administrators who were the sole evaluator in their building or district, noting that the workload impact on these people was disproportionate and negative.

Each item in group H3 targeted a different aspect of the daily workload experienced by administrators and teachers. Items H3-1–H3-5 measured TPEP’s impact on a 5-point Likert scale ranging from 5 (extremely positive) to 1 (extremely negative). In each case except for item H3-1, means were significantly higher for the teacher perspective responses than they were for the administrative perspective responses, indicating that TPEP was having a more negative impact on respondents as administrators than it did when they were teachers. Items H3-6–H3-7 were measured on a 4-point Likert scale ranging from strongly agree (4) to strongly disagree (1). On both of these questions, administrator means were significantly higher than teacher means. For all item and group means, standard deviations, and p-values, see Table 12.

Item H3-1 asked respondents to identify TPEP’s impact on their time to focus on instructional leadership and innovation. The p-value was greater than .05, indicating that the difference between the means was not statistically significant. This makes sense because the focus of TPEP is the same regardless of whether the relevant person is taking an administrative perspective or a teacher perspective. Because of TPEP’s increased detail and focus on best instructional practices and its detailed descriptions of what best practice looks like, respondents felt TPEP had a slightly positive impact on their ability to focus on instructional leadership and innovation. This feeling was evidenced by means that were the highest of the group H3 questions (3.252 for administrative and 3.220 for teacher). The fact that respondents indicated an increased focus on instructional leadership is a positive finding, as it is a major focus of the TPEP initiative (S. Anderson, personal communication, November 25, 2018). It is worth noting that this was the only non-statistically significant result among the group H3 items. This also makes sense because this item addresses a matter of focus and not a matter of work volume and therefore does not influence the administrative perspective disproportionately, like all the other

items. As TPEP supports continue to be refined and teacher and administrator preparatory programs further integrate TPEP into their coursework, it is logical to believe that the focus on instructional leadership will continue to increase and the perceived workload will continue to decrease as incoming practitioners will enter the field with a greater capacity to engage in the TPEP process.

Items H3-2–H3-5 all had similar results. Each of the results yielded statistically significant p-values of .000 and had means that ranged from 2.330 to 2.673 for the administrative perspective and 2.750 to 3.033 for the teacher perspective. See Table 12 for the entire list of means, p-values, and standard deviations. On item H3-2, how has TPEP impacted your ability to handle issues of student discipline, respondents reported a somewhat negative impact from the administrative perspective and no impact from the teacher perspective. This result is to be expected, as administrator time is disproportionately impacted by issues of student discipline. Teachers tend to take care of TPEP duties outside of class time, when they would not be dealing with issues of student discipline, so TPEP has no impact on their ability to deal with issues of student discipline. On item H3-3, how has TPEP impacted your time to interact with families, respondents indicated a somewhat negative impact from the administrative perspective and no impact from the teacher perspective. This result from the administrative perspective is not surprising. Large portions of administrator time are dedicated to communicating with families and community members. With the increased demands of TPEP, that time has to come from somewhere. As administrators shift their focus from being building managers to being instructional leaders, there is a need to think creatively about how we can communicate with families effectively. Many systems have been, and will continue to be, developed that will help administrators communicate with constituents consistently and clearly to help lighten the

communication load and allow more time to focus on instructional leadership. The neutral impact on the teacher perspective is somewhat surprising, because state evaluation criterion seven is “Communicating and collaborating with families and school community” (WAC 392-191A-060, 2013). This means that either teachers were already communicating with families and community members in a satisfactory fashion or they have been able to transition and meet the TPEP requirements smoothly and without significant negative impact. Item H3-4, how has TPEP impacted your informal time with students, follows the same pattern. Respondents indicated a somewhat negative impact on informal time with students from an administrative perspective, while teachers indicated no impact. When asked what they are doing to help manage the increased workload arising from TPEP, respondents from the administrative perspective indicated that they were using the strategies of calendaring, delegating, and prioritizing (See Figure 14). Respondents from the teacher perspective indicated they were using the strategies of streamlining their classrooms and using group collaboration (See Figure 19). What this means is that, from the administrative perspective, respondents are having to sacrifice informal time with students because of prioritization whereas, from the teacher perspective, they are not. Item H3-5 asked respondents about TPEP’s impact on their ability to respond to paperwork, email, and other tasks. Administrative and teacher responses were significantly different, as the p-value was .000, which is under the threshold of .05. Item H3-5 is slightly different because it is the only question of this group to which answers ranged from somewhat to slightly negative. Paperwork is something that both administrators and teachers tend to do outside of the hours of the regular school day or during their prep time. From my experience, many of the increased demands of TPEP (engaging in pre and post conferences, answering pre and post conference questions, establishing and tracking student growth goals, and developing a

professional growth plan) also need to be taken care of outside of regular school hours. It follows, then, that this item would be perceived as having the greatest negative impact from the teacher perspective. In my experience, some of the workload issues surrounding paperwork, email, and other administrative tasks can be alleviated by proper delegation and collaboration with office staff trained to help in this area in efforts to minimize the negative impact.

Items H3-6 and H3-7 were both measured on a 4-point Likert scale asking for level of agreement. Both items showed statistical significance having p-values of .001, which is well below the threshold value of .05. Item H3-6 asked respondents about their level of agreement with the statement that TPEP is useful but takes up too much of my time. Respondents indicated that they somewhat agreed with the statement (mean of 2.967) from an administrative perspective and slightly agreed with the statement from a teacher perspective (mean of 2.663). These responses follow the pattern of the previous five workload questions indicating that administrators are disproportionately impacted by the demands of TPEP in a negative way. Effective ways to combat this disproportionate impact on administrators, as indicated by respondents, include; hiring additional administrative staff, hiring teachers on special assignment to take on instructional coach or dean of students duties, or hiring a school social worker to help alleviate social and emotional issues before they turn in to discipline issues that need to be taken care of by administrative staff. The unfortunate truth is that less than 60% of respondents have seen these supports in action (See Figure 12). The only support available to more than 60% of the respondents was the use of electronic tools. This is logical, as electronic tools such as iObservation and eVAL are provided by the state free of charge, so schools that struggle financially can still access these supports. Item H3-7, the final item in the H3 group, asked respondents to indicate their level of agreement with the statement “I would like more time to

work on my own evaluation.” Respondents answering from the administrative perspective indicated that they somewhat agreed with the statement (mean of 2.811) and slightly disagreed (mean of 2.453) from the teacher perspective. Administrators can be responsible for upwards of 50 evaluations a year, and teachers are responsible only for their own evaluation, so the statistically significant difference in the questions in group H3 means makes sense given the increased evaluative scale administrators deal with.

When group H3 questions are taken as a whole, it is clear that the impact of TPEP disproportionately impacts administrators in a negative way. Work by Elfers and Plecki (2019) indicates that administrators in small schools where they are the only evaluator and administrator feel this impact more than do those who have colleagues to share the evaluation role with. It is clear that the null hypothesis—there is no statistically significant difference in the way educators view TPEP as an administrator and as a teacher—can be confidently rejected, as six of the seven questions yielded a statistically significant p-value.

H4 – TPEP’s Usefulness for Professional Development

One of the primary purposes for TPEP was to gather information that would help direct professional development in a more targeted fashion (Elfers & Plecki, 2019). Items in group H4 were all connected to TPEP’s usefulness for professional development. Respondents were asked a series of nine questions labeled H4-1–H4-9. Items H4-1, H4-2, and H4-9 all asked for respondent’s level of agreement with a given statement and were measured on a 4-point Likert scale with answers ranging from strongly agree to strongly disagree. Items H4-3–H4-8 all asked respondents to rate TPEP’s impact on different aspects of professional development and were measured on a 5-point Likert scale where answers ranged from extremely positive to extremely negative. The mean for the administrative perspective on questions measured by the 4-point

Likert scale was 2.775. The mean for the teacher perspective on questions measured by the 4-point Likert scale was 2.564. The mean for the administrative perspective for the 5-point Likert scale questions was 3.457, while the mean for the teacher perspective was 3.423. Only item two showed a statistically significant difference between the means, having a p-value of .003, which is below the threshold of .05.

For items measured on the 4-point Likert scale, administrative responses consistently rated higher than teacher responses. This indicates that respondents consistently agreed more with the statement regarding TPEP from an administrative perspective. For item H4-1, TPEP is useful for improving my work, respondents indicated that they somewhat agreed from the administrative perspective (mean of 2.915) and slightly agreed from the teacher perspective (mean of 2.702). Respondents believed they were seeing benefits, however slight, to their professional practice, meaning that TPEP was helping to fulfill one of its stated purposes: furthering professional development. If we are to maximize TPEP's positive impact on educators' ability to do their work, administrators and teachers must become more familiar with the framework and understand how to implement it with fidelity. The more tightly criteria and elements are focused and student learning data is tracked, the more improvement can be made at the school and district levels (Levine & Lezotte, 1990; Sammons, 1999; Sammons et al., 1995).

Item H4-2 was the only item in group H4 that showed a statistically significant difference between the administrator perspective responses and the teacher perspective responses. The p-value for item H4-2 was .003, which is lower than the threshold of .05 indicating statistical significance. The administrator perspective mean (2.713) was significantly higher than the teacher perspective mean (2.415), indicating that respondents believed more strongly that the feedback they received from their own TPEP evaluation as an administrator was more helpful

than the feedback they received as a teacher (from the teacher perspective). This situation is problematic because each administrator evaluates many teachers, and if the feedback teachers are receiving is not seen as useful to improve their practice, then a huge goal of TPEP is not being met (S. Anderson, personal communication, November 25, 2018). Although all evaluating administrators are required to go through framework-specific training, these data indicate that more training on the practical application of TPEP is needed. In addition, the fact that teachers are not seeing the feedback they receive as useful indicates that implementation fidelity should potentially be called into question. I have firsthand experience with two different schools that had entire teaching staffs who indicated they had received no helpful feedback from their TPEP observations. I believe that the perceived usefulness of TPEP feedback will increase as older administrators, who have been reluctant to adopt the new framework with fidelity, retire and new administrators who have had TPEP training in their principal preparation programs fill those positions.

Teacher professional development has long been the subject of criticism. It has come a long way since the 1960s, when it was believed that high-quality teaching and learning could not overcome student demographic effects (Coleman et al., 1966). The movement from sit and get professional development delivered in one session to professional development that is ongoing, supported by expert practitioners, aligned with school improvement goals, and embedded in practice has been a part of the TPEP initiative even before it was mandated by the state in 2010. Weglinsky (2002) suggested that all students, even those from disadvantaged backgrounds, could learn at high levels with high-quality content and instruction (Weglinsky, 2002). Adding to that research, Mundry (2005) suggested that professional development focused on rich content and high-leverage teaching practices could lead to greater student learning gains. A great example of

these concepts applied effectively in the Marysville School District was covered in depth in the literature review and gives hope that TPEP's impact on teacher professional development will become increasingly positive as administrators and teachers become more adept at engaging in authentic professional development (Gallucci et al., 2005; Swanson, 2007).

TPEP seeks to use evaluation to align individual and building-wide needs, identified through formal and informal observations and student growth goal tracking, with effective professional development. Item H4-3 asked respondents to indicate TPEP's impact on the quality of professional development they received in their district. Respondents indicated that TPEP had a slightly positive impact on the quality of professional development they had received, as the means were 3.404 and 3.393. Five years into the initiative, we still have not seen the large-scale change in data-driven professional development that will truly allow TPEP to have its intended impact on teaching and learning. This difficulty can be addressed if district leadership works with building leaders to clearly articulate priorities for teaching and learning. This collaborative process will allow building-level leaders to align all TPEP goals, data collection, and professional development to best serve the most pressing student needs.

Item H4-4 asked respondents to indicate TPEP's impact on the professional conversations about what constitutes effective teaching. There was no statistically significant difference between the administrative and teacher perspectives. The administrative mean (3.81) indicates that respondents believe TPEP has had a somewhat positive impact on the conversations they are having about what constitutes effective instruction, while the teacher perspective mean (3.62) indicates that the teacher perspective is slightly less positive about the professional conversations that are taking place among staff regarding effective instruction.

Responses to Items H4-5 and H4-6 were both closely related. Item H4-5 asked respondents to respond to TPEP's impact on the quality of instruction, and item H4-6 asked respondents to respond to TPEP's impact on student outcomes. It makes sense that these two questions would be closely related, because instruction and outcomes are directly connected. Respondents indicated that TPEP has had a slightly positive impact on instruction (administrative mean = 3.648, teacher mean = 3.549) and on student outcomes (administrative mean = 3.626, teacher mean = 3.549). Although this is a small impact over a 5-year period, TPEP implementation has been a large-scale, complex change, and sustained effort will be required to continue to implement it with fidelity and to compound the small, but positive, impact.

Item H4-7 asked respondents to identify TPEP's impact on the quality of relationships between the district and the teachers association. Responses indicated that relationships between the district and the teacher's union have not been affected from the teacher perspective and were only slightly negatively affected from the administrative perspective. Washington State is one of 44 states where collective bargaining is either required or allowed (Saenz-Armstrong, 2021). Among those 44 states that allow collective bargaining, Washington is one of 37 states that either requires, or allows, states to collectively bargain evaluation (Saenz-Armstrong, 2021). When an issue such as teacher evaluation is collectively bargained, it follows that these negotiations can lead to tension and negative feelings between the district and the teacher union. The fact that there is only a slightly negative impact is testament to the work that the state, the districts, and the unions have done to convey the importance of teacher evaluation and to build a common vision concerning how it can positively impact the profession moving forward. One would expect that as long as the TPEP initiative continues to be funded and supported properly,

relationships between districts and the teachers unions will continue to improve. Should support and funding wane, burnout could ensue, causing increased tension between districts and teacher unions (Elfers & Plecki, 2017).

Item H4-8 asked respondents to identify the impact that TPEP has had on teachers' ability to meet the instructional needs of students of diverse cultural backgrounds. All three evaluation frameworks include elements that address the teacher's need to meet the instructional needs of the increasingly diverse student body (Danielson, 2013; Marzano, 2018; UWCEL, 2012). Strengthening culturally relevant teaching through evaluation is important because culturally relevant teaching fosters deeper conceptual understanding and democratic citizenship while increasing the value of cultural diversity (Fulton, 2009). Respondents indicated that the framework has had a slightly positive impact on teachers' ability to meet the needs of students from diverse cultural backgrounds. As further culturally relevant training and professional development are connected to the TPEP frameworks, we can only expect this positive impact to increase.

Item H4-9 asked respondents about their level of agreement with the statement TPEP helps me align the improvement initiatives within my school. In their study evaluating the implementation of TPEP, Elfers and Plecki (2017) stated two major purposes: accountability and improvement (Elfers & Plecki, 2017). Elfers and Plecki found that "incorporation of the instructional frameworks has been among the most valuable aspects of the evaluation process because of the coherence it can provide in ongoing instructional improvement efforts" (p. viii). Respondents to item H4-9 indicated that TPEP had a neutral impact on the alignment of improvement initiatives. I believe that this is because superintendent and district office personnel were not surveyed in this sample. Because central office personnel generally set the

school improvement plan, which drives the school improvement initiatives, many respondents may not have had thorough enough knowledge of the school improvement plan to recognize the impact of TPEP on the alignment of improvement initiatives.

Other Findings – Administrator Perspective

Questions that were not analyzed using the paired questions *t*-test also provided valuable insight into the changes in perspective respondents had as they transitioned from the classroom to the administrative office.

Item 18 – Division of Labor Among Administrators

On item 18, respondents were asked to answer questions on a 4-point Likert scale indicating the extent of their agreement with different statements regarding division of labor among evaluating staff. The Likert scale ranged from strongly agree to strongly disagree with somewhat agree and somewhat disagree in the middle.

Item 18-1 asked respondents whether they agreed with the statement that evaluative duties were evenly distributed; 73.9% either strongly or somewhat agreed. This is encouraging because, as evidenced by research from Elfers and Plecki (2019) and by data collected in this research in group H3 items, increased workload is one of the primary negative impacts, from an administrator perspective. Even distribution of evaluative duties can help mitigate these negative impacts and increase longevity for administrators. ESDs and pre-service administrative programs would do well to intentionally teach ways to practice distributed leadership effectively in the area of evaluation.

Item 18-2 asked respondents whether the building principal took on a larger role in the evaluation process. Sixty-two percent of respondents indicated that the principal did take on a larger role in the evaluation process. As building leaders, the first thing principals are charged

with is staff and student safety. Close behind staff and student safety in importance is instructional leadership, and evaluation plays a vital role in being an instructional leader. Evaluation can be used as an effective tool to coach teachers into better practice, and it can also be used as an effective means to move forward building-wide initiatives such as those concerning increases in instructional equity and focus on content standards. Evaluation can also be used to direct professional development, as addressed in the H4 questions in this research. Principals must have their finger on the pulse of many aspects of their school buildings. There are mixed messages in the 62% agreement indicated on this item. It is encouraging that a majority of principals carry a larger portion of the evaluative load. Further questions about why 38% of principals do not carry a larger evaluative role need to be asked. Possible reasons could include administrative role specialization and large school size.

Item 18-3 asked respondents whether subject matter expertise was taken into account when evaluation duties were assigned. It is ideal to have evaluators who have a strong grasp of the content area, especially at the secondary level, so it is encouraging that 64.1% of respondents indicated that they strongly or somewhat agreed with the statement. Having subject matter experts evaluating staff in their field can help provide better feedback to teachers in content-specific areas and allows the evaluatee to gain both content and pedagogical feedback through the evaluation process.

Item 18-4 asked respondents whether they agreed that TPEP caused an increase in discussion of best practice. Over 70% either strongly or somewhat agreed that TPEP did cause an increase in pedagogical discussions. Professional collaboration around best practices is essential to creating an environment of learning for all students (Dufour, 2004). When teachers systematically get together and collaborate about student work, what is working, and what needs

to be changed, all educators learn more, and teaching and learning improve (Dufour, 2004).

Questions still remain about the nature of the discussion of best practice. Worthy questions for further investigation connected to item 18-4 may include:

- Is the discussion of best practice systematic?
- Is the discussion of best practice built into the teacher's workday?
- Does the discussion of best practice include looking at student work samples?

Increased discussion that answers all these questions in the affirmative is highly likely to influence student learning in a positive and sustainable way.

Items 18-5 and 18-6 are similar in nature. Item 18-5 asks whether the quality of the evaluation would be similar regardless of who was doing the evaluating, and item 18-6 asks whether the scores would be similar. On item 18-5, 72.8% of respondents either strongly or somewhat agreed with the statement that the quality of the evaluation would be similar regardless of evaluator. On item 18-6, 81.5% of respondents indicated that teacher scores would be similar regardless of who the evaluator was. Inter-rater reliability has been a concern from the onset of TPEP, so it is encouraging that such a high percentage of respondents believed that there was cohesion in the quality and scoring of their evaluations. Anecdotally, I have had firsthand experience with evaluations that were of equal quality and scoring throughout the building, but the quality of the evaluations was poor, and the scores were not indicative of actual teacher performance. Simply aiming for equal quality of evaluations and equal scoring across the board is not enough. Seeking to ensure not only equal quality and scoring, but high-quality and accurate scoring as well, is the goal. High-quality evaluations, from a practitioner perspective, must be actionable enough to drive teachers toward improvement, and they must be accurate in their assessment of teacher practice. If both of these components are in place, equal

quality and equal scoring will be a strong measure of the effectiveness of the TPEP evaluation frameworks.

Items 20 and 21 – Types of Support Provided for TPEP and Their Perceived Effectiveness

On item 20, respondents were given a list of potential supports that schools have provided to help building leaders manage the increased evaluative load required by TPEP. Supports included electronic tools, additional administrative staff, teacher leaders, instructional coaches, social workers, and deans of students. Each of these supports seeks to support one or more aspects of a building-level administrator's role in efforts to reduce the ever-increasing workload.

Electronic Tools. Electronic tools were the most frequently provided support for administrators. Of the administrators surveyed, 80.6% indicated they had been given access to electronic tools. Common tools included iObservation, eVAL, and Google Docs. The purpose of these tools is to provide a platform for evaluators to house all components of the evaluation process. All certificated staff on comprehensive evaluation are required to have at least two formal observations per year. Each of these formal observations must include three main components. The first component is the pre-observation conference. This is a meeting between the evaluator and the evaluatee where lesson plans are discussed and questions can be asked prior to the observation. This activity is essential because it gives the evaluator context going into the observation. The second component is the formal observation. This observation must be at least 30 minutes long and include detailed notes regarding what the teacher and students are saying and doing during their instructional time and what the classroom environment is like. The final component of the observation is the post-conference. During the post-observation

conference, the evaluator and evaluatee can dialogue regarding what took place during the lesson, talking about areas of strength, areas for improvement, and points of emphasis to continually improve the evaluatee's teaching practice.

Teachers who have at least three years of successful evaluations move from comprehensive to focused evaluation. Teachers on focused evaluation must have only one formal evaluation per year. Focused evaluations also include the pre and post conferences and the formal observation. The major difference, besides the quantity of observations, is that teachers on focused evaluation choose only one of the eight evaluation criteria to focus on throughout the year rather than all eight criteria.

In addition to formal observations, there are also informal observations, student growth goals, and professional growth plans that teachers need to document to successfully complete their evaluation. Informal observations are useful for following up on points of emphasis from formal observations. Although informal observations do not require pre or post conferences and do not need to be at least 30 minutes long, they do still require time to carry out and document. Student growth goals are documented evidence of student progress on an agreed-upon goal. Teachers on comprehensive evaluations have a three-tiered goal whereas those on focused evaluations have only a single tier, which depends on which criteria they choose to focus on. Professional growth plans are required of all teachers on comprehensive evaluation and are a collaborative effort between evaluator and evaluatee to address specifically identified areas for professional growth.

Tools such as eVAL (developed through a partnership between OSPI, the Washington Education Association, and the Bill and Melinda Gates Foundation and offered free of charge to school districts in Washington State); iObservation (produced by Marzano Research Labs and

available for purchase); and Google Docs (available through the Google Suite, which is available to schools free of charge) are all examples of digital tools used regularly by districts in Washington State (Marzano, 2019; OSPI, n.d.c). Given that two of the three electronic evaluation tools are free, it is discouraging that only 80% of administrators have access to them. Possible reasons for the lack of electronic tools for 20% of respondents include small district size, older administrators who resist moving to electronic formats, a “this too shall pass” mentality, and lack of resources to train both teachers and administrators on the electronic tools that are free of charge. Further investigation into why 20% of school districts choose not to use electronic tools is needed.

On item 21, respondents were asked to evaluate the effectiveness of the electronic tools provided. Over 89% of respondents indicated that the electronic tools provided were somewhat to very useful, with 24.7% indicating that the electronic tools were very useful. Of the six supports offered, electronic tools had the fourth-highest overall percentage for usefulness and the fifth highest for very useful (See Figure 12). Given the perceived usefulness of electronic tools indicated by survey respondents, districts that struggle to keep administrators because of high job stress and work demands, in addition to districts that have administrators who are the only evaluator in their building, would do well to access these electronic tools. I have firsthand experience in transitioning a single administrator building from paper evaluation documentation to electronic tools, and I agree with the respondents that electronic tools are useful in helping to manage the workload. Teachers in the building I transitioned from paper to electronic evaluation were appreciative of the transition and said it simplified the process.

Additional Administrators. Another support districts are using to help alleviate the increased evaluation workload was to hire additional administrators. Of survey respondents,

44.2% indicated that they had received additional administrative support, and 97.8% indicated that hiring additional administrators was somewhat or very useful. Most significantly, 69.6% of respondents indicated that hiring additional administrators was very useful. The very useful percentage was 28.7% higher than the next highest choice—hiring a dean of students—indicating it was the most helpful support provided.

The benefits of hiring additional administration are obvious. When additional administrators are hired, duties can be delegated, and evaluation duties are not an exception. As indicated in item 18, nearly 75% of administrators agreed that evaluative duties were evenly distributed, meaning hiring additional administrative staff would lighten the load for all administrators in the building, which would make the workload more sustainable.

Only 44.2% of respondents indicated that they had received additional administrative support. I believe the major barrier to their receiving this support is financial, although further research needs to be done in the area. As of December 2022, the average assistant principal salary in Washington State was \$100,230. This was salary alone and did not include benefits or any other compensation (Salary.com, 2022). In small rural districts, this is a disproportionate size of the budget, so it can become a barrier to providing proper evaluative support. In larger districts that have lower socioeconomic demographics, per pupil funding can be a fraction of what it is in similar-sized districts with higher socioeconomic demographics, causing finances to become a barrier as well. When finances become a barrier to providing the resources necessary to implement a required program successfully, the issue of equity arises. I will elaborate more on this discussion in the future avenues for research portion of Chapter 5.

Teacher Leaders. The third most frequently provided support was teacher leaders (41.9%). Of the 41.9% of respondents who had accessed teacher leaders as a support for TPEP,

95.8% of them indicated that using teacher leaders was somewhat to very helpful, with 20.8% indicating it was very helpful.

Teacher leaders can take both a formal and an informal role. I have been a teacher leader who helped with evaluation as a teacher on special assignment (TOSA). I have also, as a lead administrator, accessed teacher leadership in an informal fashion. I have found both formal and informal use of teacher leadership to be helpful in distributing the evaluation duties. Formal teacher leadership is far more helpful because the TOSA is given allocated time during the contracted day to help with evaluation. This help can take the form of conducting evaluations, providing professional development, or developing evaluation resources to help staff complete their observations. Teacher leaders who assist with evaluation in an informal role can be mentors for teachers, present best practices at staff meetings, and open up their classrooms to be model classrooms for peer observation. Informal teacher leaders can be tremendously helpful in managing the TPEP workload, as they can serve as a go-between between the teacher and the evaluator and can be available for questions when the evaluating administrator is not. I believe the percentage of respondents indicating they had the support of teacher leaders is so low because no distinction was made between the use of formal and informal teacher leaders. I believe that the 41.9% of respondents who indicated they had access to teacher leaders as a support were referring to those who had a formal TOSA of some kind rather than those who effectively used distributed leadership to access teacher leaders. More research can be done regarding drawing this distinction and sharing best practices for leveraging informal teacher leaders, and such research could be especially fruitful for those in districts where finances are a barrier to effective TPEP implementation.

Instructional Coaches. Instructional coaches were the fourth most frequently provided support (37.2%). Of the respondents who indicated they had access to instructional coaches, 94.7% indicated the coaches were somewhat or very useful, with 26.3% indicating they were very useful.

The primary use for instructional coaches in connection with TPEP is to help teachers address areas of improvement and student growth goals during the evaluation cycle. In districts where administrators have larger numbers of evaluatees, it is impossible to follow up with all of the evaluatees in a timely fashion. Instructional coaches can meet with both the teacher and the administrator and help be responsible for monitoring the teacher's implementation of feedback from formal and informal evaluations, thereby helping to lighten the evaluation load for administrators.

Instructional coaches do come at a cost. If a teacher is moved to an instructional coach position, then that teacher position must be covered. That means distributing that teacher's students to other classes, causing increases in class sizes while keeping the budget the same, or hiring additional staff. Instructional coaches tend to be teachers who have been teaching long enough to be recognized as strong teachers. This longevity means they are higher on the salary schedule, so the cost of hiring them as an instructional coach can be prohibitive. It also means that districts with greater financial capacity can hire exemplary teachers away from districts with lower financial capacity for these instructional coach positions. This practice certainly has the potential to increase the gap between those of our students who receive high-quality instruction and those who do not along demographic lines in ways that are not beneficial to our most marginalized students.

Social Worker and Dean of Students. The fifth most frequently provided support was the hiring of a social worker (29.1%). Of respondents who had received a social worker, 82.8% indicated that the social worker support was somewhat or very useful, with 27.6% indicating that the social worker was very useful. The least frequently provided support was hiring a dean of students. Only 18.6% of respondents indicated that they had received a dean of students. Of that 18.6%, 86.4% indicated that hiring a dean was somewhat or very useful, with 40.9% indicating it was very useful.

The main purpose of hiring either a social worker or a dean of students to support TPEP is to help alleviate the discipline load that administrators deal with on a daily basis. School social workers help students develop the skills they need to deal with problems at home, build bridges with at-risk students and families who are most likely to cause discipline problems, and help students in crisis calm down and prepare to return to learning. In schools without a social worker, this work generally falls on school counselors, administrators, or other trusted adults in the building. Deans of students' primary task to help alleviate the TPEP workload is to handle issues of student discipline. They are responsible for supervision and for handling routine student discipline so that the principals and assistant principals can focus on the instructional leadership portion of their duties.

Social workers generally earn between \$40,000 and 50,000 a year in Washington State (Salary.com, 2022). Technically, social workers do not directly affect instruction because they are not involved in the teaching and learning of academic content. Practitioners in schools know that social workers have a significant impact on instruction by helping high-needs students grow the skills they need to work through trauma and learn to regulate their emotions in healthy ways. These skills lead to more time in class, which leads to more learning. The district commitment

to adding social workers is smaller, in most cases, than that to adding extra teaching staff, deans of students, or administrative staff, and social workers were the second least provided support among all the supports offered.

Deans of students are generally TOSA positions, so people in these positions make a teacher's salary. Deans of students tend to be teachers who are either aspiring administrators or well-respected educators who can effectively handle discipline problems, paperwork, and parent relationships. This means they are educators with experience, and in many cases a master's degree of some kind, so they have moved up the pay scale. The average salary for a dean of students is approximately \$115,000 (Salary.com, 2022). Like administrative positions, the dean of students position can be cost prohibitive and therefore out of reach for schools that traditionally serve marginalized students.

Item 23 – Ability to Carry Out Various Administrative Duties

On survey item 23, respondents were asked to rate their ability to carry out various aspects of their administrative duties. A 3-point Likert scale using the categories of very good, good, and needs improvement was used. Data from item 23 can be found in Figure 13.

The area where respondents were most likely to rate themselves as good or very good was in their ability to implement new discipline procedures. A total of 94 respondents answered the question. Of those 94 respondents, 93.4% rated themselves as either very good (28.7%) or good (61.7%). Although this is encouraging, our schools need to be safe places in order for learning to take place, it underscores the difficulty administrators are faced with as they transition from the previous administrative priority, that of building manager, to the new administrative priority of instructional leader.

The areas of ability to support the professional growth of teachers and to serve as an instructional leader were the items on which respondents were next most likely to rate themselves as very good or good. When asked to rate themselves on their ability to support the professional growth of teachers, 85.2% of respondents indicated that they were either very good (42.6%) or good (42.6%), while only 14.9% indicated that they needed improvement. Similarly, when asked to rate their ability to serve as an instructional leader, 81% of respondents indicated they were either very good (34%) or good (46.8%), while only 19% indicated they needed improvement. This breakdown is encouraging because it indicates that, although respondents are more likely to rate themselves as being very good or good in regard to discipline, they are most likely to rate themselves as very good in the areas related to instructional leadership. When administrators are functioning as instructional leaders, they are building the professional capital in their building that Fullan (2014) identified as the “first key” to maximizing impact (p. 70).

Areas where respondents were most likely to see themselves as needing improvement were: ability to delegate (36.2%), ability to improve family engagement (36.2%), and ability to manage overall workload (37.2%). These data confirm earlier findings from Chapter 4 of this study regarding the statistically significant impact of TPEP on workload for administrators. As indicated in the introduction, workload issues related to TPEP have contributed to a sizable number of educators considering leaving the profession (Elfers & Plecki, 2017). Data from this research can help districts and educator preparation programs provide more effective support concerning strategies for effective delegation, family engagement, and workload management in order to help administrators find sustainable careers in education.

Item 24 – Strategies for Workload Management

Item 24 asked respondents to identify strategies they used to help manage the increased workload resulting from TPEP. Responses were grouped into 10 categories. Five of the 10 categories received over 10 responses and will be discussed here. The five categories receiving at least 10 responses were calendaring, delegating, working more, prioritizing, and not able to manage. The full list of categories and response numbers can be found in Figure 14.

Of the five strategies for managing the increased workload of TPEP, three are healthy and two are not. The most frequently implemented healthy strategy for workload management was delegation (33%). Despite being the most frequently used strategy among our respondents, it was also one of the strategies that respondents indicated needed the most improvement in item 23. Prioritizing (26.9%) and calendaring (25.6%) were also identified as healthy strategies being used by respondents to help manage the increased workload resulting from TPEP. The major concern about the data regarding workload management strategies is that only one third of respondents indicated that they delegated responsibilities well, and just over one fourth were using the strategies of prioritizing and calendaring. Although I do not believe this breakdown is indicative of the general administrative population, I do believe that it lends more credence to the need to include more support and training for executive functioning strategies for aspiring administrators. Potential avenues for this training include local school districts, local ESDs, and university administrator preparation programs.

The final two workload management strategies that received over 10 responses were working more hours (14.1%) and “I am not able to” (12.8%). I have deemed both of these responses as unhealthy responses because they lead to burnout and turnover in the administrative ranks. One respondent indicated “I try to keep myself to 12-hour work days.” Although it is encouraging to see that the unhealthy percentages are lower than the percentages for healthy

workload management strategies, there is still work to be done. Nearly 27% of respondents indicated that they used strategies that are not sustainable over the long term, suggesting that there is a need for either further training or hiring of additional support to help manage the workload.

Item 29 – Desired Areas for Training

Item 29 was a free response item that asked respondents to indicate TPEP-related areas where they desired to have further training. Many respondents who answered other questions did not respond to this question. I believe they did not respond because it was a free response question and required more thought and time, so they skipped the question and moved on to the rest of the survey.

Of the 44 respondents who did complete item 29, 40.9% indicated they would like time to collaborate with colleagues regarding TPEP. The most commonly desired areas for collaboration included general support (29.5%) and inter-rater reliability (15.9%). Collaboration time can be difficult to find in the workday, and effective collaboration structures are key to maximizing the impact of any collaboration time.

Needs for collaboration time must be addressed very differently depending on school and district size. An effective strategy in larger districts that I have participated in is called principal labs. In principal labs, principals from across the district gather at a building, observe a classroom together, and then sit down after the observation and discuss how they scored the observation. This sequence of steps builds both individual competence and inter-rater reliability. Principal absences are covered by other building-level administrators and usually take half a school day. In smaller districts, covering these absences is far more difficult. Proximity to colleagues, smaller budgets, and lack of administrative coverage are all barriers to collaboration

for those in small districts. Ways to alleviate these barriers include drawing on organizations like Rural Alliance and New Leaders, which take administrators from rural districts and provide them with training, both in person and virtually, that can be tailored to the unique needs of small-school environments.

Respondents also indicated a desire for specific training. Just under half (43.2%) of respondents indicated that they would like more training in specific aspects of TPEP. Areas for desired training included framework-specific training (15.9%), student growth goal writing (11.4%), the collection and scoring of artifacts (9%), the intent of TPEP (4%), and use of the online tools (2%). This lack of desire for training in specific elements of TPEP might be the result of two situations. The first could be that the administrator received high-quality training provided through the district or ESD that serves that specific administrator. The administrators who saw no need for training in specific areas of TPEP could be highly confident in their abilities to implement TPEP with fidelity and feel that no further training is necessary. Further research on the administrators' perception of their ability to implement all aspects of TPEP with fidelity would help definitively answer this question. The other possibility would be that administrators just do not care to be trained on the implementation of TPEP and therefore do not want any additional training. Elfers and Plecki (2019) have already done research on fidelity of TPEP implementation. A follow-up study could help establish whether fidelity of implementation has increased over time or whether fidelity has gone stagnant.

Item 30 – Helpfulness of TPEP Feedback for Performing Specific Duties

Item 30 asked respondents to rate TPEP's usefulness in four specific areas of their administrative duties: planning professional development, identifying teacher leaders, making progress on the school improvement plan (SIP), and improving as an instructional leader. These

areas were chosen because they connect closely to the mission statement of the TPEP steering committee (personal communication, November 25, 2018). Sub-items were measured on a 4-point Likert scale that asked respondents to rate the helpfulness of TPEP with values of a great deal, somewhat, a little bit, and not at all helpful.

Item 30-1 asked respondents to indicate how helpful TPEP was for planning professional development. This question connects directly to item three of the TPEP steering committee's mission statement (personal communication, November 25, 2018). Of the 92 respondents, 89.1% indicated that TPEP helped at least a little bit, with 18.5% of respondents indicating that TPEP feedback helped a great deal. Implied in this data is that at least 10% of respondents work in buildings or districts that are not using TPEP with fidelity. If TPEP is not being used to identify school wide areas for professional development, then a huge part of the intent of the initiative is not being realized, for one of two reasons. Either the extra work needed to provide the specific feedback in an authentic TPEP evaluation is not being done, or the extra work required to provide the specific feedback provided in an authentic TPEP evaluation is not being maximized. Either way, the intent of the initiative is being compromised. Additional consideration should also be given to the actual practicality of implementing TPEP with fidelity as it currently exists.

Item 30-2 asked respondents to indicate the extent to which they found TPEP helpful in identifying teacher leaders. Of the 92 respondents, 90.1% indicated that TPEP was at least a little bit helpful, with 38.2% saying that TPEP helped a great deal. That administrators are using TPEP to identify teacher leaders in a vast majority of the schools surveyed is encouraging. It indicates that the vision of the TPEP steering committee was being realized, at least in part, as teacher leaders were being identified and empowered to help grow the instructional practices in

the building. Item 30-2 connects directly to the issue of administrator workload as well. At the time of the survey, only 40.9% of respondents indicated they had access to teacher leaders to help with TPEP implementation, while 95.8% of respondents indicated that teacher leaders were useful in some capacity. In addition, a relative area of weakness for respondents was the ability to delegate, which was identified as a strategy for managing TPEP workload by only 33% of respondents. TPEP directly assesses teacher leadership in many of its elements and in all frameworks. In order to receive the highest rating for many of the elements, a teacher must be either implementing an innovative strategy that reaches students who are not served well in the typical classroom or be helping others to grow their practice (TPEP, n.d.). These requirements help call teachers to higher levels of practice and encourage them to help lead others to those higher levels as well.

Item 30-3 asked respondents how helpful TPEP was for assessing progress on their SIP. Of the 91 respondents, 90.2% indicated that TPEP was helpful in monitoring progress on their SIP. TPEP can be helpful for implementing a SIP, as doing so requires a schoolwide goal that uses either a teacher-generated collaborative product or schoolwide student data to assess progress. Again, authentic implementation of TPEP forces schools to turn to data and evidence to assess progress rather than simply using an arbitrary satisfactory or unsatisfactory rating. The 9.9% of respondents who said that TPEP was not helping on their SIP at all fall into the same categories as the respondents from 30-1. Either they are not implementing with fidelity or they are not maximizing the data they collect from authentic implementation. Either way, the purpose of the TPEP initiative is not being fully served.

Item 30-4 asked respondents to indicate how much TPEP helped them improve as an instructional leader. Of the 93 respondents, 95.7% indicated that TPEP helped them improve as

an instructional leader, with 76.1% indicating that TPEP helped them a great deal or somewhat. Only 4.3% of respondents indicated that TPEP did not help them grow as an instructional leader at all. It is encouraging to see the growth in practice, indicated by item 30-4 in that it is in strong alignment with the TPEP Steering Committee's mission statement (personal communication, November 25, 2018). Improvement as an instructional leader was the portion of item 30 that respondents agreed with most and gives reason for optimism. If we have instructional leaders who are willing to grow, then that will trickle down to teachers and student learning!

Other Findings – Teacher Perspective

Items 38 and 39 – Types of Support Provided for TPEP and Their Perceived Effectiveness

Electronic Tools. Of the 93 respondents for items 38 and 39, 66.7% of them indicated having access to electronic tools to help with TPEP. As with the administrative findings, it is curious how only two-thirds of respondents had access to the electronic tools, as they are offered free through OSPI and Google or can also be purchased through the Marzano Research Group. One possible explanation for this low number is the timing of the respondents' experience. At the time of the study, survey respondents were all serving in an administrative role in some capacity, as evaluator experience was a prerequisite for survey completion. Because respondents had been out of the classroom since 2020, many of them may have had teacher experience with TPEP in its early stages. In the early stages of TPEP, electronic tools were not as readily available, potentially resulting in lower access.

Of the 66.7% of respondents who had access, 95.2% said that the electronic tools were very (19.4%), somewhat (40.3%), or slightly (35.5%) useful. It is encouraging to see the high percentage of respondents who saw the value in the electronic tools. With this data, hopefully

more districts that are not currently using electronic tools will take advantage of the opportunity to implement this free and useful resource.

Time at Staff Meetings. Of the 93 respondents for items 38 and 39, 54.3% indicated that they had access to dedicated TPEP time at staff meetings. Of those respondents, 84% indicated that the time was useful in some capacity, with 14% indicating it was very useful, 34% somewhat useful, and 36% slightly useful. It is curious that only 54.3% of respondents indicated that they had access to TPEP time at staff meetings. From personal experience, I know staff meeting time is very important. I also know that what gets covered in staff meetings reflects the priorities of the school. To see that only 54.3% of respondents got any kind of support at staff meetings is disappointing, as it indicates a lack of value for TPEP. Some potential reasons for this low number are the significant amount of time that had elapsed since respondents were a teacher in a staff meeting, the potential for the TPEP time in the staff meeting to have been short and inconsequential, and a “this too shall pass” attitude from acting administration at the beginning of TPEP implementation. It is also discouraging to see that TPEP time in staff meetings was seen as the least useful of all of the supports. This conclusion is fairly easy to explain. When respondents were teachers, TPEP was in the initial stages of implementation and many acting administrators were just beginning their journey toward full implementation between 2012 and 2016.

Teacher Leaders. Of the 93 respondents for items 38 and 39, 62% indicated they had access to teacher leaders to help them with TPEP. Of that 62%, 33% indicated the teacher leaders were very useful, 35% indicated they were somewhat useful, and 21.1% indicated they were slightly useful. As in the administrative portion of the survey, teacher leaders were shown to be a useful resource for their colleagues engaged in the TPEP process. In recent years,

districts have taken measures to increase access to teacher leaders for TPEP. Larger districts have created central office positions designed to train principals and teacher leaders in the TPEP process. Smaller districts have taken advantage of zero- or no-cost opportunities offered by ESDs to help grow staff capacity and further distribute the TPEP workload.

Administrator-Led TPEP Time. Of the 93 respondents for items 38 and 39, 68.1% of them indicated having access to administrator-led TPEP time as a teacher. Of those respondents, 96.2% indicated that administrator-led TPEP time was very to slightly useful. Just 21.9% of respondents indicated that the administrator-led time was very useful; 32.8% of respondents indicated that it was somewhat useful, and 29% indicated that it was slightly useful.

Administrator-led professional development is valuable to schools in many ways. First, it is free of charge. Any time you can access high-quality professional development for your teachers free, that is beneficial to the district. In addition, when principals get the opportunity to show competence in a way that helps grow staff professionally, it builds confidence in the school's leadership. Finally, it builds a collaborative atmosphere rather than a top-down model.

Collaboration between staff and administration regarding best practices builds an environment of relational trust and academic press, which are two of the three components of schools that Salina et al. (2016) identified as being of importance for administrators to enhance in efforts to improve student achievement.

Grade-Level Collaborative Time. Of the 93 respondents surveyed, 68.1% indicated that they had access to grade-level collaborative time. Grade-level collaborative time was seen as the most useful from the teacher perspective. A strong majority of 96.9% of respondents indicated that grade-level collaborative time was useful in some capacity. In addition, 39% indicated it was very useful, 31.3% indicated it was somewhat useful, and 29% indicated it was slightly

useful. Grade-level collaborative time can be difficult to find during the contract day. That almost 70% of respondents indicated they had some sort of grade-level collaborative time around TPEP is encouraging. Collaboration is an essential component of high-performing schools, and structured and focused collaboration is even better. Marzano (2003) outlined best practices in *What Works in Schools*, and structured collaboration around identified standards in a seminal part of that. Although schools have varying degrees of success with collaboration, finding the time in the work day is actually one of the more difficult hurdles to clear. To ensure maximum use of this collaborative time, schools and districts can be encouraged to adopt research-based collaborative structures, such as PLCs.

Item 41 – Ability to Carry Out Aspects of TPEP

Item 41 asked respondents to indicate their level of proficiency on four different aspects of TPEP from a teacher perspective. Respondents could choose very good, good, or needs improvement. A total of 92 respondents completed this item. Complete data for all four aspects can be found in Figure 18.

The first aspect of TPEP that respondents were asked about was their ability to create a professional growth plan, or PGP. Of the 92 respondents, 77.1% indicated they were either very good (29.3%) or good (47.8%) at creating a PGP. These data are encouraging because targeting professional growth was one of the core reasons for implementing TPEP, as previous evaluation systems that rated teachers as either satisfactory or unsatisfactory did not allow for targeted professional development because of a lack of specific feedback. Heidmets and Liik (2014), in their research on teacher well-being, indicated that transformational leadership, as opposed to transactional leadership, had a positive impact on teacher well-being. The PGP is transformational by nature and allows for teachers and administrators to engage in positive,

transformational conversations about personal growth. PGPs can also contribute to teacher efficacy by targeting, collaborating about, and highlighting growth in high-leverage teaching strategies. Ross et al. (2012) connected the concept of teacher efficacy to a teacher's ability to effect change in difficult students or in difficult circumstances (Ross et al., 2012). Underscoring the importance of teacher efficacy, Armor et al. (1976) stated that "teachers' sense of self-efficacy has been found to positively relate to teachers' instructional behavior and student outcomes" (as cited in Ross et al., 2012, p. 2). A teacher's perception that they can create and follow through on a quality PGP focused on positive changes in both the teacher and the students benefits all parties involved, by increasing a teacher's sense of efficacy and providing concrete evidence of the teacher's effect on student learning.

The second aspect of TPEP respondents were asked about was their ability to create student growth goals (SGG). Of the 92 respondents, 81.5% indicated that they were either very good (38%), or good (43.5%) at writing SGGs. These data are encouraging for many of the same reasons the PGP data is encouraging. When teachers identify areas for student growth and see concrete change, their sense of efficacy increases, which, again, leads to positive outcomes (Armor et al., 1976, cited in Ross et al., 2012, p. 2). Attention should be paid to the 18.5% of respondents who said that they needed improvement in creating SGGs. The creating, and tracking, of SGGs can be something that is done as a group during a staff meeting, or it can be done in conjunction with already required pre and post conferences that are attached to formal observations. It is incumbent upon building leaders to make sure all staff feel confident in their ability to write authentic SGGs so that growth can be charted and reasons for lack of growth can be identified.

The third aspect of TPEP that respondents were asked about was their ability to provide artifacts. Of the 92 respondents, 84.8% indicated they were either very good (40.2%) or good (44.6%) at providing artifacts. This datapoint follows a trend we have seen throughout this study. The trend involves things that are the least beneficial being the most prevalent. For example, in supports provided, from both the teacher and the administrative perspective, electronic tools were the most commonly provided but seen as the least beneficial. Further cost benefit analysis of all supports provided is needed to allocate resources in the most efficient way possible.

The final aspect of TPEP that respondents were asked about was their ability to manage their workload. Of the 92 respondents, 75% indicated that they were either very good (25%) or good (50%) at managing their workload. Although a majority of respondents chose one of these positive responses, managing workload did have the lowest percentage of very good responses. It also had the lowest percentage of combined very good and good responses. This outcome raises the same concerns that were raised in the administrative perspective questions. If we have 25% of our respondents indicating they need improvement in managing their workload, it is incumbent on Washington State, school districts, and building leaders to get feedback from teachers regularly about how they are doing personally and professionally and to respond with supports when they are needed. According to McCallum (2017), examples of ways to support teachers include

reflection strategies for insight into professional practice, mindfulness training to manage stress, emotional management strategies, coaching psychology to build learning communities, growth mindset approaches to solving problems, self-care practices to restore when needed, and celebrating achievements and success to feel valued. (McCallum et al., 2017, p. 32)

Further lines of research could include surveys of practices used in schools and perceived ability to manage teacher workload.

Item 42 – Strategies for Managing Workload, Teacher Perspective

Item number 42 asked respondents to identify strategies they used to manage the increased workload caused by TPEP from a teacher perspective. There were 68 responses to this question. The number of respondents was probably lower because of survey fatigue and the open-ended nature of the question. The 68 responses were broken down into 10 different categories that can be found in Figure 19. The top five categories will be discussed in this section.

The most common strategy employed by respondents (22.1%) was to streamline their classrooms. Streamlining strategies employed by teachers included only grading necessary standards, focusing on essential standards, and simplifying assignments. This result is in strong alignment with TPEP, as elements of TPEP are focused directly on essential standards and assignments.

The second most common strategy employed by respondents (20.6%) was to collaborate with colleagues. This was done through PLCs as well as in less formal collaboration time. The necessity of collaboration in education has never been more important, and that is reflected in the TPEP framework. Criteria seven and eight in the Marzano framework both are focused on leadership and collaboration (Marzano, 2010). It is encouraging to know that TPEP frameworks are driving teachers toward collaboration both with their colleagues and with their evaluators.

Time management strategies have also been implemented by a lower number of respondents (14.7%). Strategies used include quick formative assessments to assess student understanding of objectives, entry tasks and exit tickets to maximize time on task, better

calendaring and lesson planning, and setting times to leave the building in efforts to maintain sustainable working hours.

Organizational strategies and working extra hours were also options chosen by respondents (13.2%). Obvious uses of this strategy include working late hours, working in the evening after getting home, and working on weekends. Organizational strategies referenced included calendaring, unit and lesson planning, and increased usage of formative assessment to measure student progress.

Of the respondents, 5.9% did indicate that they were currently unable to manage the workload. This result is particularly worthy of discussion, as it was chosen by less than half of the administrative percentage (12.8%). This result further supports the notion that the impact of TPEP on workload has been far more detrimental to administrators than to teachers. Teachers and administrators alike need training and support to manage their workload in a healthy way that allows them to maintain a sustainable work-life balance.

Item 45 – Training Desired, Teacher Perspective

Of the 95 respondents who participated in the survey, only 50 took the time to answer the last question regarding training desired from the teacher perspective. It is not surprising that the areas where respondents indicated needing support from the administrator perspective are also the areas where they indicated needing support from the teacher perspective. The 50 responses were broken down into three categories for classification.

The first category was elements of TPEP. Of the respondents, 21 indicated that they would like training in various elements of TPEP. Nine of the respondents indicated they would like general framework training, six indicated a need for training in writing SGGs, three indicated they would like training in developing and submitting artifacts, two indicated an

interest in training in the background of TPEP, and one indicated the need for training in the electronic tools. The need for training around all of the different, and relatively new, areas of TPEP seems to be present but not overwhelming. The teacher needs can be met by addressing the expressed needs of the administrators responsible for training their teachers. Currently, Washington State mandates that all administrators who will evaluate certificated personnel undergo framework specific training. These trainings are generally led by ESD personnel and have been, for the most part, virtual over the past 3 years. Although virtual training is effective, it is not a replacement for in-person training where administrators can bring in their own artifacts and discuss desired training topics. As we shift back to more in-person training, I predict the understanding of TPEP will grow and the areas of desired training will evolve into more sophisticated elements, like inter-rater reliability or content-based best practices that will yield greater results in both teacher practice and student learning.

The second area for desired training was collaborative time. Respondents indicated a high desire to collaborate with colleagues for TPEP training from both the administrative and the teacher perspectives. Respondents indicated both a desire to collaborate with grade-level colleagues (80%) and with their evaluator (20%). Many districts have been able to work collaboration time into the teacher workday to help alleviate some of the increased workload. Other districts that have not been able to work the collaborative time into the workday have offered teachers extra hours pay or comp time for the time when they collaborate. It is clear that the increased demands of TPEP have necessitated collaboration. Further research disaggregating workload data between schools that have existing collaborative structure and schools that do not would be beneficial.

The final area for training requested concerned best teaching practices (14%). Best practice training is always beneficial. The major barrier to offering it is time. Many districts have very limited professional development time, and that time can, in many cases, be taken up by logistics rather than by instructional leadership. Educational leaders charged with staff development would do well to respond to staff whenever possible when staff want training in best practices. Building staff capacity can increase staff buy-in, teacher practice, and student learning.

Limitations of the Study

This survey had one primary limitation that had multiple causes. The primary limitation of this study was the low number of respondents. When filtered for incomplete surveys, I was left with 95 responses (N = 95). At the time of the survey, the best estimate of the number of evaluating administrators in Washington State was 3,285. This estimate was gained by accessing district websites for all school districts in Washington State through the OSPI website, visiting each website, and creating a database of all listed administrators, excluding deans of students and athletic directors. I used this method because I was not granted access to the most recent state listserv after requesting it from both OSPI and AWSP.

In addition, estimates had to be used to determine what percentage of surveyed respondents would even be eligible to take the survey given the survey parameters. Elfers and Plecki (2019) indicated that approximately 48% of administrators would be eligible, leaving 1,576 eligible participants. Because respondents needed to have experience with being evaluated and being an evaluator with TPEP frameworks, there was no way to determine the exact N without having access to each administrator's work history, which was not practical to obtain.

Another factor that contributed to the low N was the timing of the survey. The survey was sent out during the fall of the 2020–21 school year. In Washington State, administrators were dealing with having to decide whether school would be online or in person, having to develop reopening plans, developing and updating ever-changing health and safety protocols, and resolving staffing issues that arose because of vaccination requirements. All these COVID-19-related elements contributed to the survey's being a low priority for many administrators, which greatly reduced response numbers.

Another factor contributing to low N was the fact that the survey was not sanctioned by OSPI or any Washington State university. In personal communication with Sue Anderson from OSPI in November of 2020, I was told that OSPI gets many requests for access to administrator listservs and for assistance in disseminating research surveys and that unless the research was sponsored by a university or by OSPI, they would be unable to assist.

Because the N was so small, other limitations arose. One limitation was the inability to disaggregate data by key school demographics, such as size, type, and poverty level. Another limitation that arose because of small N was the inability to disaggregate on the basis of evaluation framework, job role, or years in the profession. Disaggregation of the data points referenced above would have yielded far more useful results. The final limitation that arose due to the small N was the concern about transferability of data and findings to the larger population. In order to try to alleviate this concern an effort was made to gather statewide gender and race/ethnicity data from OSPI, WASA, AWSP, and the UWCEL to see in my research sample showed demographics that were representative of the statewide population. Unfortunately, data from these institutions, discussed earlier in chapter five, did little to alleviate the transferability concerns so the concern regarding small sample size remains.

Knowing that a small response rate would be a limiting factor, I took measures to try to maximize response rate. The first measure I took was to expand the survey window from 2 weeks to 1 month. This shift allowed for more response time and more reminders. The second thing I did to maximize response rate was to send reminder emails to all nonrespondents through Qualtrics weekly during the survey window. Reminders were also sent out 48 and 24 hours before the survey window closed. These reminders proved to be effective—I did see a bump in response numbers after reminders were sent out. The final thing I did to increase the response rate was to offer Amazon gift cards and survey data to all respondents upon completion of the research.

Another measure I took to try to account for small N, and because no previous population existed, was to conduct a post hoc power analysis using the bootstrapping method outlined by Bliese and Wang (2019). For the power analysis, a .104 effect size (well within the small range) was assumed to ensure that statistically significant results were not capitalizing on chance. The power analysis indicated that statistically significant results were valid and not just the result of a small N.

A final limitation, that was independent of sample size, was limited translatability. Many states have laws concerning type, frequency, and method of teacher and principal evaluation. If states do not have similar evaluation frameworks and regulations, this research serves very little purpose for educators in those states.

Recommendations for Further Studies

If I were to repeat this study, I would take increased measures to gain the support of OSPI, AWSP, or a research university in Washington State in an effort to give the survey greater credibility and a greater response rate. Greater response rates would allow future researchers to

disaggregate the data more effectively in the areas identified in the above section on limitations.

In addition, another change that I would make is to change the timing of the survey.

Unfortunately, completion timelines required that I disseminate the survey during a very busy and volatile time for administrators. Collaborating with researchers at state universities, OSPI, AWSP, and WASA regarding the best time to distribute the survey would be useful and could lead to a greater sample size. If support were gained from state institutions, surveys could be disseminated around state conventions or meetings. This approach could allow administrators time to give thoughtful and complete responses to all questions, which might lead to higher response rates.

Another line of inquiry stemming from this research would be to survey central office personnel, specifically in regard to their perceptions of supports provided. Many times those providing supports have a vastly different perspective on the effectiveness of the supports provided than those receiving the supports. Data obtained from this line of inquiry could be very useful to superintendents and building leaders alike. Knowing the perspectives of all parties involved in the provision of, and access to, supports would allow for those responsible for allocating funds to do so in a manner that is most useful to those who were doing the frontline work.

The greatest line for further inquiry that arose from this research concerns teacher and administrator workload. Of the nine questions showing statistical significance in both the t-test and the power analysis, six were in the area of workload. The difference observed indicated that the impact of TPEP on workload was significantly more negative from the administrator perspective than it was from the teacher perspective. A seventh question, “TPEP has made me consider leaving my position,” also indicated a significantly higher negative impact from the

administrator perspective. Further research needs to be done to isolate schools that are effectively helping administrators manage workload in order to translate those practices to other districts who are not. These districts can be identified via high administrator retention and reported job satisfaction, among other elements. Special attention should be paid to districts that have high administrator retention and job satisfaction in typically underserved demographics, such as urban, rural, and low SES schools. Intentionally identifying strategies that work for our most marginalized populations can help increase equity for our staff and students who are traditionally underserved.

All further lines of inquiry should be supported by OSPI, AWSP, WASA, or a research institution in Washington State. Without greater sample sizes, meaningful disaggregation of data cannot take place, so any results cannot be effectively translated into action.

Summary

The purpose of this study was to examine changes in perspective for educators who had gone from being evaluated on TPEP to being an evaluator using a TPEP framework. This study found that individuals experienced TPEP differently in statistically significant ways. Responses from the administrative perspective were significantly more negative than those from the teacher perspective in the areas of workload and career choice. Responses from the administrative perspective were significantly more positive in the areas of usefulness of feedback and staff interactions. In efforts to address current and future staffing issues, representatives at the state level would do well to advocate for resources and training to help alleviate the increased workload caused by TPEP for administrative positions so that these positions can become increasingly more sustainable in the future.

References

- American Statistical Association (2014), “ASA Statement on Using Value-Added Models for Educational Assessment,” available at http://www.amstat.org/policy/pdfs/ASA_VAM_Statement.pdf.
- ASCD. (2020, July 1). ASCD Community in Action. <https://www.ascd.org/el/articles/ascd-community-in-action-june-20>
- ASCD. (2021). About ACSD. <http://www.ascd.org/about-ascd.aspx>
- AWSP. (2014). AWSP leadership framework. https://awsp.org/member-support/leadership-framework_1
- AWSP. (2019). Pay and working conditions survey results. <https://awsp.org/member-support/resources/hot-topics/pay-gap-resources/pay-working-conditions-survey>
- AWSP. (2021). Inside AWSP. (2021). <https://awsp.org/inside-awsp>
- Ayer, F. C. (1954). *Fundamentals of instructional supervision*. Harper & Brothers.
- Babbie, E. (1973). *Survey research methods*. Wadsworth.
- Ballard, K., & Bates, A. (2008). Making a connection between student achievement, teacher accountability, and quality classroom instruction. *The Qualitative Report*, 13(4), 560–580. <https://doi.org/10.46743/2160-3715/2008.1574>
- Bangert-Drowns, R. L., Kulik, J. A., & Kulik, C. L. C. (1991). Effects of frequent classroom testing. *Journal of Educational Research*, 85(2), 89–99. <https://doi.org/10.1080/00220671.1991.10702818>
- Barnett, B., & Hirsch, E. (2005). *Recruiting and retaining teachers for hard-to-staff schools*[Issue Brief]. NGA Center for Best Practices.

[https://www.academia.edu/63359744/Recruiting and Retaining Teachers for Hard to Staff Schools Issue Brief](https://www.academia.edu/63359744/Recruiting_and_Retaining_Teachers_for_Hard_to_Staff_Schools_Issue_Brief)

Barker, R. G., & Gump, P. V. (1964). *Big school, small school: High school size and student behavior*. Stanford University Press.

Beerens, D. R. (2000). *Evaluating teachers for professional growth: Creating a culture of motivation and learning*. Corwin Press.

Best countries for education. (n.d.). *U.S. News & World Report*. Retrieved June 1, 2021, from <https://www.usnews.com/news/best-countries/best-countries-for-education>

Birnbaum, A. (1968) Some latent trait models and their use in inferring an examinee's ability. In F. M. Lord & M. R. Novick (Eds.), *Statistical theories of mental test scores* (pp. 397–479). Addison-Wesley.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, policy & practice*, 5(1), 7–74. <https://doi.org/10.1080/0969595980050102>

Bliese, P. D., & Wang, M. (2019). Results provide information about cumulative probabilities of finding significance: Let's report this information. *Journal of Management*, 46(7), 1275–1288. <https://doi.org/10.1177/0149206319886909>

Bloom, H. S., Ham, S., Melton, L., & O'Brien, J. (2001, November). *Evaluating the accelerated schools approach: A look at early implementation and impacts on student achievement in eight elementary schools*. MDRC. <https://www.mdrc.org/publication/evaluating-accelerated-schools-approach>

Bohrnstedt, G. (2013, October 30). Three decades of education reform: Are we still “a nation at risk?” AIR. <https://www.air.org/resource/spotlight/three-decades-education-reform-are-we-still-nation-risk>

- Bullough, R., Clark, D., & Patterson, R. (2003). Getting in step: Accountability, accreditation and the standardization of teacher education in the United States. *Journal of Education for Teaching*, 29, 35–51. <https://doi.org/10.1080/0260747022000057945>
- Burke, P., & Krey, R. D. (2005). *Supervision: A guide to instructional leadership*. Charles C Thomas.
- Castetter, W. B. (1976). *The personnel function in educational administration*. Macmillan.
- Centra, J. A. (1993). *Reflective faculty evaluation: Enhancing teaching and determining faculty effectiveness*. Jossey-Bass.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112, 155–159.
- Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., & York, R. L. (1966). *Equality of educational opportunity*. National Center for Educational Statistics. <https://doi.org/10.3886/ICPSR06389.v3>
- Cowan, J., Goldhaber, D., Hayes, K., & Theobald, R. (2016). Missing elements in the discussion of teacher shortages. *Educational Researcher*, 45(8), 460–462. <https://doi.org/10.3102/0013189X16679145>
- Cotton, K. (1996, May). School size, school climate, and student performance. *School Improvement Research Series: Close-up #20*. <https://tinyurl.com/school-size>
- Cremin, L. A. (1976). *Public education*. Basic Books.
- Creswell, J. W. (2003). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage.
- Creswell, J. W. (2007). *Qualitative inquiry & research design*. Sage Publications.
- Creswell, J. W., & Guetterman, T. C. (2021). *Educational research: Planning, conducting and evaluating quantitative and qualitative research*. Pearson.

- Danielson, C. (2013). *The framework for teaching evaluation instrument: 2013 edition*.
<https://tinyurl.com/2013-danielson>
- Darling-Hammond, L., Wise, A. E., & Pease, S. R. (1983). Teacher evaluation in the organizational context: A review of the literature. *Review of Educational Research*, 53(3), 285–328. <https://doi.org/10.2307/1170367>
- Darling-Hammond, L., & Berry, B. (1988). *The evolution of teacher policy*. RAND Corporation.
https://www.rand.org/pubs/joint_reports-education/JRE01.html
- Darling-Hammond, L. (2010). Recruiting and retaining teachers: Turning around the race to the bottom in high-need schools. *Journal of Curriculum and Instruction*, 4(1), 16–32.
<https://doi.org/10.3776/joci.2010.v4n1p16-32>
- Dillman, D. A., Smyth, J. D., & Christian, L. M. (2014). *Internet, phone, mail, and mixed-mode surveys: The tailored design method*. Wiley.
- District Framework Selections. (n.d.). Retrieved April 08, 2021, from
<https://tinyurl.com/yubkhffr>
- DJUSD [Davis Joint Unified School District]. (n.d.). Common Core State Standards.
https://www.djUSD.net/instruction/curriculum/common_core
- DOE [Department of Education]. (2016, July 19). Race to the top fund. Retrieved March 27, 2021, from <https://obamawhitehouse.archives.gov/issues/education/k-12/race-to-the-top>
- Donaldson, M. L. (2012, December). *Teachers' perspectives on evaluation reform*. Center for American Progress. <https://cdn.americanprogress.org/wp-content/uploads/2012/12/TeacherPerspectives.pdf>
- Drew, D. E. (2011). *Stem the tide: Reforming science, technology, engineering, and math education in America*. Johns Hopkins University Press.

- Drew, S.V., & Sosnowski, C. (2019). Emerging theory of teacher resilience: A situational analysis. *English Teaching: Practice & Critique*, 18(4), pp. 492–507.
<https://doi.org/10.1108/ETPC-12-2018-0118>
- DuFour, R. (2004). What is a “professional learning community”? *Educational Leadership*, 61(8), 6–11.
- Elfers, A., & Achberger, E. (2017). *Analysis of School Employees Evaluation Survey (SEES)*. University of Washington, Center for the Study of Teaching and Policy.
- Elfers, A. M., & Plecki, M. L. (2017). *Washington's teacher and principal evaluation system: Examining the implementation of a complex policy*. (Rep.). Seattle, WA: Center for the Study of Teaching and Policy, University of Washington.
- Elfers, A. M., & Plecki, M. L. (2019). *School leaders and teacher evaluation: Learning, leading, and balancing responsibilities*. Center for the Study of Teaching and Policy, University of Washington.
- Elfers, A. M., Plecki, M. L., & Van Windekens, A. (2017). *Understanding teacher retention and mobility in Washington State*. Center for the Study of Teaching and Policy, University of Washington.
- Ellett, C. D., & Teddlie, C. (2003). Teacher evaluation, teacher effectiveness and school effectiveness: Perspectives from the USA. *Journal of Personnel Evaluation in Education*, 17(1), 101–128. <https://doi.org/10.1023/A:1025083214622>
- Engel, M., Jacob, B. A., & Curran, F. C. (2014). New evidence on teacher labor supply. *American Educational Research Journal*, 51(1), 36–72.
<https://doi.org/10.3102/00028312135030>
- ESSB 5895 (2013). <https://tinyurl.com/ESSB-5895>

- ESSB 6696. (2010). <https://lawfilesexternal.wa.gov/biennium/2009-10/Pdf/Bills/Senate%20Passed%20Legislature/6696-S2.PL.pdf>
- Every Student Succeeds Act, 20 U.S.C. § 6301 (2015). <https://www.congress.gov/bill/114th-congress/senate-bill/1177>
- Fullan, M. (2014). *The principal: Three keys to maximizing impact*. Jossey-Bass.
- Fullan, M., Hill, P., & Crevola, C. (2006). Critical learning instructional paths: Assessment for learning in action. *Orbit*, 36(2), 10–14.
- Fulton, R. (2009). *A case study of culturally responsive teaching in middle school mathematics* [Doctoral dissertation, University of Denver]. *Electronic Theses and Dissertations*, 217. <https://digitalcommons.du.edu/cgi/viewcontent.cgi?article=1216&context=etd>
- Gardner, D. P. (1983, April). *A nation at risk: The imperative for educational reform*. National Commission on Excellence. http://edreform.com/wp-content/uploads/2013/02/A_Nation_At_Risk_1983.pdf
- Gilbert, J. (2007). Catching the knowledge wave?: Redefining knowledge for the post-Industrial Age. *Education Canada*, 47(3), 4–8.
- Goals 2000: Educate America Act, (1994), H.R. 1804 – 103rd Congress. <https://www.congress.gov/bill/103rd-congress/house-bill/1804>
- Goddard, Y., Goddard, R., & Tschannen-Moran, M. (2007). A theoretical and empirical investigation of teacher collaboration for school improvement and student achievement in public elementary schools. *Teachers College Record*, 109(4), 877–896. <https://doi.org/10.1177/016146810710900401>
- Goldhammer, R. (1969). *Clinical supervision: Special methods for the supervision of teachers*. Holt McDougal.

- Gordon, R. J., Kane, T. J., & Staiger, D. (2006). *Identifying effective teachers using performance on the job*. Brookings Institution.
- Gruber, C., & Onwuegbuzie, A. (2001). Effect of block scheduling on academic achievement among high school students. *High School Journal*, 84(4), 32–42.
<https://doi.org/10.1353/hsj.2001.0010>
- Graham, E. (2013, April 25). A Nation at Risk turns 30: Where did it take us? *NEA Today*.
<http://bigeducationape.blogspot.com/2013/04/a-nation-at-risk-turns-30-where-did-it.html>
- Gallucci, C., Boatright, E., Lysne, D., & Swinnerton, J. (2006). *The pedagogy of third-party support for instructional improvement: A partnership between CEL and Highline School District (interim report #1)*. University of Washington Center for Educational Leadership.
- Heidmets, M., & Liik, K. (2014). School principals' leadership style and teachers' subjective wellbeing at school. *Problems of Education in the 21st Century*, 62, 40–50.
http://www.scientiasocialis.lt/pec/files/pdf/vol62/40-50.Heidmets_Vol.62.pdf
- Hurley, J. C. (1997). The 4×4 block scheduling model: What do students have to say about it? *National Association of Secondary School Principals Bulletin*, 81, 53–64.
<https://doi.org/10.1177/019263659708159308>
- Improving America's Schools Act of 1994. H.R. 6 – 103rd Congress.
<https://www.congress.gov/bill/103rd-congress/house-bill/6>
- Jefferson, Thomas. (January 6, 1816). [Letter from former president Thomas Jefferson to Charles Yancey, 1816]. Retrieved from <https://founders.archives.gov/documents/Jefferson/03-09-02-0209>. [Original source: *The Papers of Thomas Jefferson*, Retirement Series, vol. 9,

- September 1815 to April 1816*, ed. J. Jefferson Looney. Princeton: Princeton University Press, 2012, pp. 328–331.]
- Johnson, S. M., Kraft, M. A., & Papay, J. P. (2012). How context matters in high-need schools: The effects of teachers' working conditions on their professional satisfaction and their students' achievement. *Teachers College Record*, 114(10), 1–39.
<https://doi.org/10.1177/016146811211401004>
- Kanter, D. E., & Konstantopoulos, S. (2010). The impact of a project-based science curriculum on minority student achievement, attitudes, and careers: The effects of teacher content and pedagogical content knowledge and inquiry-based practices. *Science Education*, 94, 855–887. <https://doi.org/10.1002/sce.20391>
- Kragen, J. (2015, January 5). TPEP is killing my principal. <https://storiesfromschool.org/tpep-is-killing-my-principal/>
- Lee, V., Croninger, R., & Smith, J. (1997). Course-Taking, equity, and mathematics learning: Testing the constrained curriculum hypothesis in U.S. secondary schools. *Educational Evaluation and Policy Analysis*, 19(2), 99–121. <https://doi.org/10.2307/1164205>
- Leithwood, K., Seashore Lewis, K., Anderson, S., & Whalstrom, K. (2004). *How leadership influences student learning*. Center for Applied Research and Educational Improvement.
- Lillard, D. R., & DeCicca, P. P. (2001). Higher standards, more dropouts? Evidence within and across time. *Economics of Education Review*, 20(5), 459–473.
<https://www.learntechlib.org/p/206871/>
- Linn, R. L. (2003). Accountability: Responsibility and reasonable expectations. *Educational Researcher*, 32(7), 3–13. <https://doi.org/10.3102/0013189X032007003>
- Lucio, W. H., & MacNeil, J. D. (1969). *Supervision: A synthesis of thought and action*. ASCD.

- Lunenburg, F. C. (2011). Theorizing about curriculum: Conceptions and definitions. *International Journal of Scholarly Academic Intellectual Diversity*, 13(1), 1–6.
- Marks, J. R., King-Stoops, J., & Stoops, E. (1985). *Handbook of educational supervision: A guide for the practitioner*. Allyn and Bacon.
- Martella, R. C., Nelson, J. R., & Marchand-Martella, N. E. (2003). *Managing disruptive behaviors in the schools: A schoolwide, classroom, and individualized social learning approach*. Allyn and Bacon.
- Marzano, R. J. (2003). *What works in schools: Translating research into action*. Association for Supervision and Curriculum Development.
- Marzano, R. J. (2006). *Classroom assessment & grading that work*. ASCD.
- Marzano, R. J. (2007). *The art and science of teaching: A comprehensive framework for effective instruction*.
- Marzano, R. J. (2013, May). The Marzano teacher evaluation model. <https://tinyurl.com/23vs3d7j>
- Marzano, R. (2018). Marzano Focused Teacher Evaluation Model. Retrieved June 17, 2021, from <https://tinyurl.com/Marzano-focused>
- Marzano, R. (2019, August 12). *Dr. Marzano's suite for improving teacher effectiveness*. iObservation. Retrieved January 21, 2023, from <https://www.iobservation.com/marzano/>
- Marzano, R. J., Marzano, J., & Pickering, D. J. (2003). *Classroom management that works: Research-based strategies for every teacher*. ASCD.
- Marzano, R. J., Pickering, D., & Pollock, J. E. (2001). *Classroom instruction that works: Research-based strategies for increasing student achievement*. Association for Supervision & Curriculum Development.

- McCallum, F., Price, D., Graham, A., Morrison, A. (2017). *Teacher wellbeing: A review of the literature*. Association of Independent Schools of NSW.
- McGorry, E., & McGorry, S. Y. (1998). Intensive scheduling: A hybrid model for the junior high. *Clearing House: A Journal of Educational Strategies, Issues and Ideas*, 71(3), 149–152. <https://doi.org/10.1080/00098659809599348>
- McNergney, R. F., Imig, S. R., & Pearlman, M. A. (n.d.). Teacher evaluation—Overview, methods. <https://tinyurl.com/cayzyv95>
- McNergney, R. F., Imig, S., & Pearlman, M. A. (1984). Teacher evaluation. In J. M. Cooper (Ed.), *Developing skills for instructional supervision* (pp. 147–178). Longman.
- Melchior, W. T. (1950). *Instructional supervision: A guide to modern practice*. Heath.
- Milanowski, A. (2011). Strategic measures of teacher performance. *Phi Delta Kappan*, 92(7), 19–25.
- Millman, J. (1981). *Handbook of teacher evaluation*. Corwin.
- Millman, J., & Darling-Hammond, L. (Eds.). (1990). *The new handbook of teacher evaluation: Assessing elementary and secondary school teachers*. Corwin.
- Mistretta, G. M., & Polansky, H. B. (1997). Prisoners of time: Implementing a block schedule in the high school. *NASSP Bulletin*, 81(593), 23–31.
<https://doi.org/10.1177/019263659708159304>
- NAEP [National Assessment of Educational Progress]. (2021). *An overview of NAEP*.
https://nces.ed.gov/nationsreportcard/subject/about/pdf/naep_overview_brochure_2021.pdf
- NASSP. (2019, July). Value-added measures in teacher evaluation. Retrieved March 27, 2021, from <https://tinyurl.com/value-added-measures>

NCES [National Center for Education Statistics]. (n.d.). Fast facts: Title 1.

<https://nces.ed.gov/fastfacts/display.asp?id=158>

NEA [National Education Association]. (2021). How to join: Membership types.

<https://www.washingtonea.org/membership/join/>

Neuman, W. L. (2009). *Social research methods: Qualitative and quantitative approaches*.
Pearson.

No Child Left Behind Act of 2001, H.R.1 – 107th Congress (2001–2002).

<https://www.congress.gov/bill/107th-congress/house-bill/1>

Orphanos, S. (2014). What matters to principals when they evaluate teachers? Evidence from
Cyprus. *Educational Management Administration & Leadership*, 42(2), 243–258.

<https://doi.org/10.1177/1741143213499262>

OSPI [Washington Office of Superintendent of Public Instruction]. (n.d.a). About OSPI. (n.d.).

<https://www.k12.wa.us/about-ospi>

OSPI. (n.d.b). Educational service districts (ESD). <https://tinyurl.com/ruejmu55>

OSPI. (n.d.c). eVAL. Retrieved January 21, 2023, from <https://www.k12.wa.us/educator-support/teacherprincipal-evaluation-program/policy-and-resources/eval>

OSPI. (n.d.d). TPEP steering committee members. Retrieved March 27, 2021, from

<https://tinyurl.com/846nfcwk>

OSPI. (n.d.e). School Employee Evaluation Survey (SEES) Help. Retrieved February 20, 2021,

from <https://www.k12.wa.us/sites/default/files/public/tpep/pubdocs/SEESFAQ-UpdatedHelp.pdf>

OSPI. (n.d.f). Websites and contact info. Retrieved May 19, 2021, from

<https://www.k12.wa.us/about-ospi/about-school-districts/websites-and-contact-info>

- Pearson, P. D., & Gallagher, M. C. (1983). The instruction of reading comprehension. *Contemporary Educational Psychology*, 8(3), 317–344. [https://doi.org/10.1016/0361-476X\(83\)90019-X](https://doi.org/10.1016/0361-476X(83)90019-X)
- Peterson, K. D. (2000). *Teacher evaluation: A comprehensive guide to new directions and practices*. Corwin Press.
- Plecki, M. L., Elfers, A. M., & Van Windekens, A. (2017). *Examining Beginning Teacher Retention and Mobility in Washington State*. Research commissioned by the Washington State Office of the Superintendent of Public Instruction. Center for the Study of Teaching and Policy, University of Washington.
- Plecki, M. L., Elfers, A. M. & Wills, K. (2017). *Understanding Principal Retention and Mobility in Washington State*. Research commissioned by the Washington State Office of the Superintendent of Public Instruction. Seattle, WA: Center for the Study of Teaching and Policy, University of Washington.
- Popham, J.W. (1988). The dysfunctional marriage of formative and summative teacher evaluation. *Journal of Personnel Evaluation in Education*, 1, 269–273. <https://doi.org/10.1007/BF00123822>
- Popham, J. W. (2013). On serving two masters: Formative and summative teacher evaluation. *Principal Leadership*, 13(7), 18–22.
- Prensky, M. (2005, September 12). “Engage me or enrage me”: What Today's™ learners demand. *Educause Review*, 40(5), 61–64. <https://er.educause.edu/articles/2005/9/engage-me-or-enrage-me-what-todays-learners-demand>
- Reavis, C. A. (1976). Clinical supervision: A timely approach. *Educational Leadership*, 33(5), 360–363. <https://www.ascd.org/el/articles/clinical-supervision-a-timely-approach>

- Robert, M. S., Stoops, E., & King-Stoops, J. (1985). *Handbook of educational supervision: A guide for the practitioner*. Allyn and Bacon.
- Robinson, K., & Aronica, L. (2009). *The element: How finding your passion changes everything*. Random House.
- Roza, M., Celio, M., Harvey, J., & Wishon, S. (2003). *A matter of definition: Is there truly a shortage of school principals?* (pp. 7–9). Center for Reinventing Public Education.
- Rury, J. L. (1991). *Education and women's work: Female schooling and the division of labor in urban America, 1870-1930*. State University of New York Press.
- Saenz-Armstrong, P. (2021, November 11). *Do state collective bargaining rules influence district teacher policies?* National Council on Teacher Quality (NCTQ). Retrieved October 16, 2022, [from https://tinyurl.com/bde35df4](https://tinyurl.com/bde35df4)
- Salary.com. (2022, December 27). *Assistant school principal salary in Washington*. Retrieved January 28, 2023, [from https://tinyurl.com/3rmf2dyf](https://tinyurl.com/3rmf2dyf)
- Salina, C., Girtz, S., & Eppinga, J. (2016). *Powerless to powerful: Leadership for school change*. Rowman & Littlefield.
- Scriven. (1966). *The methodology of evaluation*. Social Science Education Consortium. (ERIC Document Reproduction Service No. ED 014-001)
- Shavelson, R. J., Linn, R. L., Baker, E. L., Ladd, H. F., Darling-Hammond, L., Shepard, L. A., Barton, P. E., Haertel, E., Ravitch, D., & Rothstein, R. (2010, August 29). *Problems with the use of student test scores to evaluate teachers*. EPI Briefing Paper #278. Economic Policy Institute. <https://files.epi.org/page/-/pdf/bp278.pdf>
- Shinkfield, A. J., & Stufflebeam, D. L. (1995). School professionals' guide to improving teacher evaluation systems. In *Teacher evaluation* (pp. 81–172). Springer Dordrecht.

- Sipple, J., & Killeen, K. (2004). Context, capacity, and concern: A district-level analysis of the implementation of standards-based reform in New York State. *Educational Policy*, 18(3), 456–490. <https://doi.org/10.1177/0895904804265639>
- Soar, R., Medley, D., & Coker, H. (1983). Teacher evaluation: A critique of currently used methods. *Phi Delta Kappan*, 65(4), 239–246. <http://www.jstor.org/stable/20386992>
- Spokane Public Schools. (2021, May 18). Homepage. Retrieved May 19, 2021, from <https://www.spokaneschools.org/>
- Spring, J. H. (2010). *American education*. McGraw-Hill Higher Education.
- Stein, M. K., & Nelson, B. S. (2003). Leadership content knowledge. *Educational Evaluation and Policy Analysis*, 25(4), 423–448. <https://doi.org/10.3102/01623737025004423>
- Strauss, V. (2012, July 2). Why US can't get back to the head of the class (because it was never there). *Washington Post*. https://www.washingtonpost.com/blogs/answer-sheet/post/why-us-cant-get-back-to-head-of-the-class-because-it-was-never-there/2012/07/01/gJQAwpqAHW_blog.html
- Sutcher, L., Darling-Hammond, L., & Carver-Thomas, D. (2016). *A coming crisis in teaching? Teacher supply, demand, and shortages in the US* [Research Brief]. Learning Policy Institute. https://learningpolicyinstitute.org/sites/default/files/product-files/A_Coming_Crisis_in_Teaching_BRIEF.pdf
- Swanson, J. (2007). *Learning at every level: A partnership between the Center for Educational Leadership and Marysville School District*. Center for the Study of Teaching and Policy, University of Washington.
- Tanner, D., & Tanner, L. (1987). *Supervision in education: Problems and practices*. Macmillan.
- Tapscott, D. (1999). *Growing up digital: The rise of the net generation*. McGraw-Hill.

- Tarr, J. E., Mcnaught, M. D., & Sears, R. (2010). *Conceptualizing and measuring fidelity of implementation of secondary mathematics textbooks: Results of a three-year study*. Paper presented at the Annual Meeting of the American Educational Research Association, April 30–May 4, Denver, CO. <https://tinyurl.com/wxkdyfv7>
- Taylor, F. W. (1911). *Scientific management*. Routledge.
- TPEP. (n.d.). *Frameworks and rubrics*. Retrieved February 14, 2021, from <https://www.k12.wa.us/educator-support/teacherprincipal-evaluation-program/frameworks-and-rubrics>
- TPEP. (2013). September 9). About TPEP. Retrieved February 20, 2021, from <http://www.tpep-wa.org/about-tpep/index.html>
- Tucker, P. D., & Stronge, J. H. (2005). *Linking teacher evaluation and student learning*. ASCD.
- UWCEL. (2012). *5 dimensions of teaching and learning*.. Center for Educational Learning. Retrieved March 14, 2021, from <https://info.k-12leadership.org/hs-fs/hub/381270/file-797242241.pdf>
- UWCEL. (2022, January 27). *The University of Washington’s Center for Educational Leadership (CEL) partners with courageous leaders in classrooms, schools and the systems that* [Status update]. Facebook. <https://www.facebook.com/UWCollegeofEducation/photos/a.370736402532/10160117706972533/?type=3>
- WAC 392-191A-060. (2013, March 10). Minimum evaluation criteria—Certified classroom teachers. Retrieved April 6, 2021, from <https://tinyurl.com/WAC-391-191A-060>

WAC 392-191A-150. (2013, March 10). Minimum evaluation criteria—Certificated principals and assistant principals. Retrieved April 6, 2021, from

<https://apps.leg.wa.gov/WAC/default.aspx?cite=392-191A-150>

WASA [Washington Association of School Administrators]. (2021). *Annual report 2020–2021*.

https://wasa-oly.org/WASA/images/WASA/6.0Resources/6.1.1Publications/Download_Files/2021AnnualReport.pdf

Willms, J. D. (2003). *Student engagement at school: A sense of belonging and participation: Results from PISA 2000*. OECD.

Wolf, E. (2013). *Memorandum: Implementation of the Teacher/Principal Evaluation Project*. Senate Early Learning and K-12 Education Committee Staff.

World Top 20 Project. (2020, December 29). 2022 World Best Education Systems—Final Rankings. Retrieved February 14, 2021, from

<https://worldtop20.org/worldbesteducationsystem>

WPTA [Washington State Parent Teacher Association]. (2021, January 29). *PTA history*.

Retrieved February 21, 2021, from <https://www.wastatepta.org/about/history/>

Wright, S. P., Horn, S. P., & Sanders, W. L. (1997). Teacher and classroom context effects on student achievement: Implications for teacher evaluation. *Journal of Personnel Evaluation in Education*, 11(1), 57–67.

Zackoff, M., Jerardi, K., Unaka, N., Sucharew, H., & Klein, M. (2015). An observed structured teaching evaluation demonstrates the impact of a resident-as-teacher curriculum on teaching competency. *Hospital Pediatrics*, 5(6), 342–347.
<https://doi.org/10.1542/hpeds.2014-0134>

Appendix A - Survey Instrument

Start of Block: Default Question Block

Q1 I understand that this survey has been approved by the Institutional Review Board (IRB) of Kansas State University and choose to engage in this survey by my own free will. I also understand that my answers will remain anonymous and that I am free to terminate the survey at any time without consequence.

- ☐ Yes, I understand the above stated parameters for the survey and agree to participate in the survey (1)
- ☐ No, I choose not to participate in the survey (2)

If answer is Yes, I understand, skip to end of survey. If answer is No, I choose not to participate, your participation in the survey will end here.

Q2 Do you have at least one year of experience being evaluated as a certificated teacher using one of the TPEP frameworks (Marzano, Danielson, or CEL 5D+)? If you answer no, you will be directed to the end of the survey.

- ☐ Yes (1)
- ☐ No (2)

If answer to this question about having one year of experience is No, skip to the end of the survey.

Q3 Do you have experience evaluating certificated teachers on any of the TPEP frameworks (Marzano, Danielson, or CEL 5D+)? If you answer No, you will be directed to the end of the survey.

- ☐ Yes (1)
- ☐ No (2)

Skip To: End of Survey If Do you have experience evaluating certificated teachers on any of the TPEP frameworks (Marzano, D... = No)

Q4 What is your Gender?

- ☐ Male (1)
- ☐ Female (2)
- ☐ Non-binary / third gender (3)
- ☐ Prefer not to say (4)

Q5 Which category best describes your ethnicity?

- ☐ Hispanic or Latino or Spanish origin (1)
- ☐ Non Hispanic, Latino, or Spanish origin (2)

Q6 Which category best describes your race (select all that apply)

- ☐ White or Caucasian (1)
- ☐ Hispanic or Latino (2)
- ☐ Asian or Asian American (3)
- ☐ American Indian or Alaska Native (4)
- ☐ Native Hawaiian or Other Pacific Islander (5)
- ☐ Another Race (6)

Q7 What grades does your building serve (check all that apply)

- ☐ Elementary (P–5) (1)
- ☐ Middle School (6–8) (2)
- ☐ High School (9–12) (3)

Q8 Which category best describes the type of school you work in?

- ☐ Urban (1)
- ☐ Suburban (2)
- ☐ Rural (3)

Q9 What is your school's percentage of free and reduced lunch?

- ☐ 0–25% (1)
- ☐ 26–50% (2)
- ☐ 51–75% (3)
- ☐ Over 76% (4)

Q10 Which of the following best describes your current position?

- ☐ Principal (1)
- ☐ Assistant Principal (2)
- ☐ Principal's assistant (3)
- ☐ Dean of Students (4)
- ☐ Other, Please specify (5)

Q11 How many certificated classroom teachers are you evaluating this year on a COMPREHENSIVE evaluation?

Q12 How many certificated teachers are you evaluating on FOCUSED evaluation?

Q13 Besides certificated classroom teachers, how many other staff in your building are you evaluating this year?

Q14 How many years have you been evaluating teachers using TPEP?

- ☐ This is my first year (1)
- ☐ This is my second year (2)
- ☐ I have been evaluating teachers on TPEP for three years or more (3)

Q15 Which instructional framework has your district adopted for certificated teacher evaluation

- ☐ CEL 5D+ (1)
- ☐ Danielson (2)
- ☐ Marzano (3)

Q16 Do you have more than one evaluator of teachers in your building? No – Proceed to question 17; Yes – Skip to question 1.

- ☐ Yes (1)
- ☐ No (2)

Skip To Q18 If answer to Do you have more than one evaluator of teachers in your building? is No. If the answer is Yes, proceed to question 17.

Q17 Please identify the reasons why you are the only evaluator in your building (Check all that apply)

- ☐ Enrollment in my building is too low (1)
- ☐ The district does not have resources to support an assistant principal (2)
- ☐ Instead of an assistant principal I have been given other supports (3)
- ☐ There are no assistant principals in my district (4)
- ☐ My district only provides assistant principals at the secondary level (5)
- ☐ Other, please specify (6)

Q18 To what extent do you agree or disagree with the following statements:

	Strongly Agree (1)	Somewhat Agree (2)	Somewhat Disagree (3)	Strongly Disagree (4)
The number of teachers to be evaluated is divided equally among evaluators (1)	☐	☐	☐	☐
The principal is responsible for a larger portion of teachers to be evaluated (2)	☐	☐	☐	☐
The evaluator's subject matter expertise is taken into	☐	☐	☐	☐

	Strongly Agree (1)	Somewhat Agree (2)	Somewhat Disagree (3)	Strongly Disagree (4)
account when assigning teachers to evaluators (3)				
Evaluators in the building regularly discuss best practices for teacher evaluation (4)	☐	☐	☐	☐
The quality of the feedback teachers receive in my building is similar regardless of who the evaluator is (5)	☐	☐	☐	☐
I am confident that teachers would receive similar scores regardless of who evaluates them (6)	☐	☐	☐	☐

Q19 To what extent do you agree or disagree with the following statement about your own evaluation as a principal or assistant principal:

	Strongly Agree (1)	Somewhat Agree (2)	Somewhat Disagree (3)	Strongly Disagree (4)
The leadership framework is a useful tool for improving my work as an administrator (1)	☐	☐	☐	☐

Feedback I receive from my own evaluation is helpful for me (2)	☐	☐	☐	☐
I receive sufficient support for my own evaluation (3)	☐	☐	☐	☐
I would like more time to work on my own evaluation with my supervisor (4)	☐	☐	☐	☐

Q20 What types of supports are available in your building (check all that apply)

- ☐ Additional administrative staff (1)
- ☐ Full- or part-time dean of students (2)
- ☐ Full- or part-time social worker (3)
- ☐ Instructional coaches (4)
- ☐ Teacher leaders (5)
- ☐ Electronic tools to support TPEP (eVAL, Google docs, iObservation, etc.) (6)

Q21 Please rate the usefulness of these supports for TPEP in your building:

	Very Useful (1)	Useful (2)	Somewhat Useful (3)	Not Useful (4)	Not Available (5)
Additional administrative staff (1)	☐	☐	☐	☐	☐
Full- or part-time dean of students (2)	☐	☐	☐	☐	☐
Full- or part-time social worker (3)	☐	☐	☐	☐	☐

Instructional coaches (4)	☐	☐	☐	☐	☐
Teacher leaders (5)	☐	☐	☐	☐	☐
Electronic tools to support TPEP (eVAL, Google Docs, iObservation, etc.) (6)	☐	☐	☐	☐	☐

Q22 Have teacher leaders, including instructional coaches and mentors, supported TPEP in your building in any of the following ways? (check all that apply)

- ☐ Provide training and support around goal setting or evidence collection (1)
- ☐ Mentoring novice teachers (2)
- ☐ Providing training and support around the instructional framework (3)
- ☐ Help developing classroom-based assessments (4)

Q23 Please rate your ability to do the following:

	Very good (1)	Good (2)	Needs Improvement (3)
Support professional growth of teachers (1)	☐	☐	☐
Delegate some of my responsibilities (2)	☐	☐	☐
Improve family engagement in my school (3)	☐	☐	☐

	Very good (1)	Good (2)	Needs Improvement (3)
Implement new student discipline policies (4)	☐	☐	☐
Serve as an instructional leader in my building (5)	☐	☐	☐
Manage my overall workload as a school leader (6)	☐	☐	☐

Q24 What strategies have you used to manage your workload this year?

Q25 Please rate the usefulness of the following supports for TPEP if available:

	Very useful (1)	Somewhat useful (2)	Slightly useful (3)	Not at all useful (4)	Did not participate (5)	Not available (6)
Professional development provided by the district (1)	☐	☐	☐	☐	☐	☐
Collaborating with other administrators on TPEP issues (2)	☐	☐	☐	☐	☐	☐

Training to
support rater
reliability (3)

☐ ☐ ☐ ☐ ☐ ☐

Q26 Teachers in my school have access to sufficient professional development opportunities related to TPEP.

- ☐ Strongly agree (1)
- ☐ Somewhat agree (2)
- ☐ Somewhat disagree (3)
- ☐ Strongly disagree (4)

Q27 I received adequate supports to work with teachers in my building on TPEP-related activities.

- ☐ Strongly agree (1)
- ☐ Somewhat agree (2)
- ☐ Somewhat disagree (3)
- ☐ Strongly disagree (4)

Q28 Would you benefit from more professional development or support related to TPEP?

- ☐ Yes, Respond to question 29 (1)
- ☐ No, Skip to item 30 (2)

Q29 What kinds of training or support materials related to TPEP would be most beneficial for you?

Q30 To what extent does the feedback you receive from conducting TPEP evaluations in your building assist with the following:

	A great deal (1)	Somewhat (2)	A little bit (3)	Not at all (4)
Plan professional development (1)	☐	☐	☐	☐
Assist in identifying teacher leaders (2)	☐	☐	☐	☐
Assess progress on my school's improvement plan (3)	☐	☐	☐	☐
Improve my ability as an instructional leader (4)	☐	☐	☐	☐

Q31 How has TPEP impacted your other responsibilities as a school leader?

	Extremely positively (1)	Somewhat positively (2)	Neither positively nor negatively (3)	Somewhat negatively (4)	Extremely negatively (5)
Time to focus on instructional leadership (1)	☐	☐	☐	☐	☐
Time to address issues of	☐	☐	☐	☐	☐

student
discipline (2)

General
visibility in
the building
(3)

Time to
interact with
families (4)

Spending
informal time
with students
(5)

Time to
respond to
paperwork,
email, and
other
managerial
tasks (6)

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

Q32 In what ways has TPEP impacted your school in the following areas:

Extremely
positive (1)

Somewhat
positive (2)

Neither
positive
nor

Somewhat
negative (4)

Extremely
negative (5)

	negative (3)				
-Quality of professional development (1)	0	0	0	0	0
Professional conversations about what constitutes effective teaching (2)	0	0	0	0	0
Quality of instruction (3)	0	0	0	0	0
Student learning outcomes (4)	0	0	0	0	0
Quality of the relationships between the district and the teachers association (5)	0	0	0	0	0

Teacher's skills to meet the instructional needs of the students from diverse cultural backgrounds (6)	☐	☐	☐	☐	☐
--	-----------------------	-----------------------	-----------------------	-----------------------	-----------------------

Q33 Please indicate the extent to which you agree or disagree with the following statements:

	Strongly agree (1)	Somewhat agree (2)	Somewhat disagree (3)	Strongly disagree (4)
TPEP is useful but it takes up too much of my time (1)	☐	☐	☐	☐
I have better interactions with my teachers because of TPEP (2)	☐	☐	☐	☐
I incorporate elements of TPEP in the professional development	☐	☐	☐	☐

offerings in my
school (3)

TPEP helps me
align the
improvement
initiatives in my
school (4)

I have flexibility
to implement
TPEP in a way
that works best at
my school (5)

Implementation
of TPEP in my
district is very
consistent across
schools (6)

The majority of
time I spend on
TPEP is useful
(7)

TPEP has made
me consider
leaving my
position as a
school leader (8)

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

Q34 In total, how many years have you worked as a school administrator?

- ☐ 1–2 (1)
- ☐ 3–4 (2)
- ☐ 5–6 (3)
- ☐ 7 or more (4)

Q35 As a teacher, how many years were you evaluated using the TPEP process?

- ☐ 1–2 years (1)
- ☐ 3–4 years (2)
- ☐ 5–6 years (3)
- ☐ 7 or more years (4)

Q36 In your last year as a teacher, which instructional framework did your district use for teacher evaluation?

- ☐ Marzano (1)
- ☐ Dainelson (2)
- ☐ CEL 5D+ (3)

Q37 To what extent do you agree with the following statements about your own evaluation as a teacher?

	Strongly agree (1)	Somewhat agree (2)	Somewhat disagree (3)	Strongly disagree (4)
The TPEP framework was a useful tool for improving my work as a teacher (1)	☐	☐	☐	☐

The feedback I received from my evaluation was helpful for me (2)

☐

☐

☐

☐

I received sufficient support for my evaluation (3)

☐

☐

☐

☐

I would have liked more time to work on my evaluation (4)

☐

☐

☐

☐

Q38 What type of building level supports were available for you as a teacher? (Check all that apply.)

- ☐ Grade-level collaborative time to write and review student growth goals and professional growth and development plans (1)
- ☐ Dedicated time at staff meetings for TPEP information (2)
- ☐ Administrator-led TPEP support time before or after school (3)
- ☐ Teacher leaders (4)
- ☐ Electronic tools to support TPEP (e.g., eVAL, Google Docs, iObservation, etc. (5)

Q39 Recalling your last year as a teacher, please rate the usefulness of the building-level supports for TPEP:

Very
useful
(1)

Somewhat
useful (2)

Slightly
useful (3)

Not
useful at
all (4)

Did not
participate
(5)

Not
available
(6)

-Grade-level collaboration, time to write and review student growth goals, and professional growth plans (1)	☐	☐	☐	☐	☐	☐
-Admin-led TPEP support time before or after school (2)	☐	☐	☐	☐	☐	☐
-Dedicated time at staff meetings for TPEP (3)	☐	☐	☐	☐	☐	☐
-Teacher leaders (4)	☐	☐	☐	☐	☐	☐
-Electronic tools (5)	☐	☐	☐	☐	☐	☐

Q40 In your last year as a teacher, did teacher leaders, including instructional coaches, support your TPEP process in any of the following ways? (check all that apply)

- ☐ Providing training and support around goal setting and evidence collection (1)
- ☐ Mentoring novice teachers (2)
- ☐ Providing training and support around the instructional framework (3)
- ☐ Helping to develop classroom-based assessments (4)
- ☐ Helping teachers learn to use the electronic tools (5)

Q41 Reflect on and rate your ability to do the following during your last year as a teacher being evaluated on TPEP:

	Very good (1)	Good (2)	Needs Improvement (3)
Create a professional growth and development plan (1)	☐	☐	☐
Write student growth goals that are authentic and meaningful (2)	☐	☐	☐
Provide artifacts to address elements not easily seen in the classroom (3)	☐	☐	☐
Manage my overall workload as a teacher (4)	☐	☐	☐

Q42 What strategies did you use to manage your workload as a teacher?

	Very useful (1)	Somewhat useful (2)	A little useful (3)	Not useful at all (4)	Did not participate (5)	Not offered (6)
Professional development provided by my district (1)	☐	☐	☐	☐	☐	☐
Collaborating with other teachers on TPEP issues (2)	☐	☐	☐	☐	☐	☐
Supports for working on my evaluation (3)	☐	☐	☐	☐	☐	☐

Q43 Please rate the usefulness of the following supports for TPEP, if available, in your last year as a teacher:

Q44 Would you have benefited from more professional development or support related to your TPEP teacher evaluation?

- ☐ Yes – respond to open-ended question 48 (1)
- ☐ No – skip to 47 (2)

Q45 What kind of training and support would have been most beneficial for you as a teacher”

Q46 How did TPEP impact your other responsibilities as a teacher while you were in the classroom?

	Extremely positive (1)	Somewhat positive (2)	Neither positive nor negative (3)	Somewhat negative (4)	Extremely negative (5)
Time to focus on instructional innovation (1)	0	0	0	0	0
Time to address issues of student discipline (2)	0	0	0	0	0
Time to interact with families (3)	0	0	0	0	0
Spending informal time with students (4)	0	0	0	0	0

Time to
respond to
paperwork,
email, and
other tasks
(5)

0 0 0 0 0

Q47 In your last year of teaching, how did TPEP impact your school in the following areas:

Extremely positive (1)	Somewhat positive (2)	Neither positive nor negative (3)	Somewhat negative (4)	Extremely negative (5)
---------------------------	--------------------------	---	--------------------------	---------------------------

Quality of
professional
development
(1)

0 0 0 0 0

Professional
conversations
about what
constitutes
effective
teaching (2)

0 0 0 0 0

Quality of
instruction (3)

0 0 0 0 0

Student learning outcomes (4)	0	0	0	0	0
Quality of the relationships between the district and the teachers association (5)	0	0	0	0	0
Teachers' skills to meet instructional needs of students from diverse cultural backgrounds (6)	0	0	0	0	0

Q48 Please indicate the extent to which you agree or disagree with the following statements.
Remember, this is from a teacher perspective:

	Strongly agree (1)	Somewhat agree (2)	Somewhat disagree (3)	Strongly disagree (4)
TPEP was useful but took up too much of my time (1)	0	0	0	0

I had better interactions with my colleagues because of TPEP (2)	☐	☐	☐	☐
TPEP helped me align the variety of improvement initiatives in my classroom (3)	☐	☐	☐	☐
I had the flexibility to implement TPEP in a way that works best in my classroom (4)	☐	☐	☐	☐
The majority of the time I spent on TPEP was useful (5)	☐	☐	☐	☐
TPEP made me consider leaving my position as a teacher (6)	☐	☐	☐	☐

Q49 What grades did you teach while using TPEP? Check all that apply.

- ☐ Elementary School (K–5) (1)
- ☐ Middle School (6–8 or 7–8) (2)

☐ High School (9–12) (3)

Q50 In total, how many years did you work as a teacher?

- ☐ 3 or less (1)
- ☐ 4–6 (2)
- ☐ 7–10 (3)
- ☐ 10 or more (4)

End of Block: Default Question Blo

Appendix B - CEL 5D+ Framework

5 Dimensions of Teaching and Learning™

Instructional Framework Version 4.0

5D™	Subdimension	The Vision	Guiding Questions
Purpose	Standards	- The lesson is based on grade-level standards, is meaningful and relevant beyond the task at hand (e.g., relates to a broader purpose or context such as problem-solving, citizenship, etc.), and helps students learn and apply transferable knowledge and skills. - The lesson is intentionally linked to other lessons (previous and future) in support of students meeting standard(s).	- How do the standard and learning target relate to content knowledge, habits of thinking in the discipline, transferable skills, and students' assessed needs as learners (re: language, culture, academic background)? - How do the standard and learning target relate to the ongoing work of this classroom? To the intellectual lives of students beyond this classroom? To broader ideals such as problem-solving, citizenship, etc.? - What is the learning target(s) of the lesson? How is it meaningful and relevant beyond the specific task/activity? - Is the task/activity aligned with the learning target? How does what students are actually engaged in doing help them to achieve the desired outcome(s)? - How are the standard(s) and learning target communicated and made accessible to all students? - How do students communicate their understanding about what they are learning and why they are learning it? - How does the learning target clearly communicate what students will know and be able to do as a result of the lesson? What will be acceptable evidence of student learning? - How do teaching point(s) support the learning needs of individual students in meeting the learning target(s)?
	Learning Target and Teaching Points	- The learning target is clearly articulated, linked to standards, embedded in instruction, and understood by students. - The learning target is measurable. The criteria for success are clear to students and the performance tasks provide evidence that students are able to understand and apply learning in context. - The teaching points are based on knowledge of students' learning needs (academic background, life experiences, culture and language) in relation to the learning target(s).	
Student Engagement	Intellectual Work	- Students' classroom work embodies substantive intellectual engagement (reading, thinking, writing, problem-solving and meaning-making). - Students take ownership of their learning to develop, test and refine their thinking.	- What is the frequency of teacher talk, teacher-initiated questions, student-initiated questions, student-to-student interaction, student presentation of work, etc.? - What does student talk reveal about the nature of students' thinking? - Where is the locus of control over learning in the classroom? - What evidence do you observe of student engagement in intellectual, academic work? What is the nature of that work? - What is the level and quality of the intellectual work in which students are engaged (e.g. factual recall, procedure, inference, analysis, meta-cognition)? - What specific strategies and structures are in place to facilitate participation and meaning-making by all students (e.g. small group work, partner talk, writing, etc.)? - Do all students have access to participation in the work of the group? Why/why not? How is participation distributed? - What questions, statements, and actions does the teacher use to encourage students to share their thinking with one another, to build on one another's ideas, and to assess their understanding of one another's ideas?
	Engagement Strategies	- Engagement strategies capitalize on and build upon students' academic background, life experiences, culture and language to support rigorous and culturally relevant learning. - Engagement strategies encourage equitable and purposeful student participation and ensure that all students have access to, and are expected to participate in, learning.	
	Talk	- Student talk reflects discipline-specific habits of thinking and ways of communicating. - Student talk embodies substantive and intellectual thinking.	

5D™	Subdimension	The Vision	Guiding Questions
Curriculum & Pedagogy	Curriculum	- Instructional materials (e.g., texts, resources, etc.) and tasks are appropriately challenging and supportive for all students, are aligned with the learning target and content area standards, and are culturally and academically relevant. - The lesson materials and tasks are related to a larger unit and to the sequence and development of conceptual understanding over time.	- How does the learning in the classroom reflect authentic ways of reading, writing, thinking and reasoning in the discipline under study? (e.g., How does the work reflect what mathematicians do and how they think?) - How does the content of the lesson (e.g., text or task) influence the intellectual demand (e.g. the thinking and reasoning required)? How does it align to grade-level standards? - How does the teacher scaffold the learning to provide all students with access to the intellectual work and to participation in meaning-making? - What does the instruction reveal about the teacher's understanding of how students learn, of disciplinary habits of thinking, and of content knowledge? - How is students' learning of content and transferable skills supported through the teacher's intentional use of instructional strategies and materials? - How does the teacher differentiate instruction for students with different learning needs—academic background, life experiences, culture and language?
	Teaching Approaches and/or Strategies	- The teacher makes decisions and utilizes instructional approaches in ways that intentionally support his/her instructional purposes. - Instruction reflects and is consistent with pedagogical content knowledge and is culturally responsive, in order to engage students in disciplinary habits of thinking. - The teacher uses different instructional strategies, based on planned and/or in-the-moment decisions, to address individual learning needs.	
	Scaffolds for Learning	The teacher provides scaffolds for the learning task that support the development of the targeted concepts and skills and gradually releases responsibility, leading to student independence.	
Assessment for Student Learning	Assessment	- Students assess their own learning in relation to the learning target. - The teacher creates multiple assessment opportunities and expects all students to demonstrate learning. - Assessment methods include a variety of tools and approaches to gather comprehensive and quality information about the learning styles and needs of each student (e.g., anecdotal notes, conferring, student work samples, etc.). - The teacher uses observable systems and routines for recording and using student assessment data (e.g., charts, conferring records, portfolios, rubrics). - Assessment criteria, methods and purposes are transparent and match the learning target.	- How does the instruction provide opportunities for all students to demonstrate learning? How does the teacher capitalize on those opportunities for the purposes of assessment? - How does the teacher gather information about student learning? How comprehensive are the sources of data from which he/she draws? - How does the teacher's understanding of each student as a learner inform how the teacher pushes for depth and stretches boundaries of student thinking? - How do students use assessment data to set learning goals and gauge progress to increase ownership in their learning? - How does the teacher's instruction reflect planning for assessment? - How does the teacher use multiple forms of assessment to inform instruction and decision-making? - How does the teacher adjust instruction based on in-the-moment assessment of student understanding?
	Adjustments	The teacher uses formative assessment data to make in-the-moment instructional adjustments, modify future lessons, and give targeted feedback to students.	
Classroom Environment & Culture	Use of Physical Environment	- The physical arrangement of the room (e.g., meeting area, resources, student seating, etc.) is conducive to student learning. - The teacher uses the physical space of the classroom to assess student understanding and support learning (e.g., teacher moves around the room to observe and confer with students). - Students have access to resources in the physical environment to support learning and independence (e.g., libraries, materials, charts, technology, etc.).	- How does the physical arrangement of the classroom, as well as the availability of resources and space to both the teacher and students, purposefully support and scaffold student learning? - How and to what extent do the systems and routines of the classroom facilitate student ownership and independence? - How and to what extent do the systems and routines of the classroom reflect values of community, inclusivity, equity and accountability for learning? - What is the climate for learning in this classroom? How do relationships (teacher-student, student-student) support or hinder student learning? - What do discourse and interactions reveal about what is valued in this classroom? - What are sources of status and authority in this classroom (e.g., reasoning and justification, intellectual risk-taking, popularity, aggressiveness, etc.)?
	Classroom Routines and Rituals	- Classroom systems and routines facilitate student responsibility, ownership and independence. - Available time is maximized in service of learning.	
	Classroom Culture	- Classroom discourse and interactions reflect high expectations and beliefs about all students' intellectual capabilities and create a culture of inclusivity, equity and accountability for learning. - Classroom norms encourage risk-taking, collaboration and respect for thinking.	

Appendix C - Marzano Framework

Marzano Instructional Framework Aligned to Washington State Teacher Evaluation Criteria

Domain 1: Observable Classroom Strategies and Behaviors

Lesson Segments Involving Routines

#1 HIGH EXPECTATIONS FOR STUDENT ACHIEVEMENT

- 1.1 Provides clear learning goals and scales **0Q1**
- 1.2 Celebrates student success **0Q1**

#6 STUDENT DATA

- 6.3 Provides opportunities for students to self-reflect and track progress toward learning goals **0Q1**

#5 SAFE, POSITIVE LEARNING ENVIRONMENT

- 5.1 Organizes a safe physical layout of the classroom to facilitate movement and focus on learning **0Q6**
- 5.2 Reviews expectations regarding rules and procedures to ensure their effective execution **0Q6**

Monitoring for the Desired Effect

How do you know the majority of students are getting it?

The Teacher

- ☐ Uses progress checks, i.e., exit tickets, clickers, white boards, etc.
- ☐ Monitors verbal and nonverbal responses
- ☐ Uses a system to record and rate important types of student cues, responses or participation
- ☐ Adjusts instruction to student cues, student participation, and/or progress

The Majority of Students

- ☐ Attend and respond to instruction

September 3, 2013

Lesson Segments Addressing Content

#2 EFFECTIVE TEACHING PRACTICES

- 2.1 Helps students effectively interact with new knowledge **0Q2**
 - 2.1.1 Identifies critical information
 - 2.1.2 Organizes students in small groups to facilitate the processing of new knowledge
 - 2.1.3 Helps students to link prior knowledge to new content
 - 2.1.4 Chunks content into "digestible bites"
 - 2.1.5 Breaks presentation of content and engages students in processing new information
 - 2.1.6 Through questions or activities, students elaborate on new information
 - 2.1.7 Students record and represent knowledge in linguistic and/or nonlinguistic ways
 - 2.1.8 Students reflect on their learning and the learning process

2.2 Helps students to practice and deepen knowledge **0Q3**

- 2.2.1 Reviews content, highlights critical information
- 2.2.2 Organizes students in groups to practice and deepen knowledge
- 2.2.3 Uses homework when appropriate (not routinely)
- 2.2.4 Students examine similarities and differences
- 2.2.5 Students examine errors in their own reasoning or the logic of information presented
- 2.2.6 Students practice skills, strategies, and/or processes
- 2.2.7 Students revise previous knowledge

2.3 Organizes students for cognitively complex tasks (transfer and application) **0Q4**

2.7 Uses and applies Academic Vocabulary

Lesson Segments Enacted on the Spot

#2 EFFECTIVE TEACHING PRACTICES

- 2.6 Engages Students **0Q5**
 - 2.6.1 Notices when students are not engaged
 - 2.6.2 Uses academic games
 - 2.6.3 Manages response rates
 - 2.6.4 Uses physical movement
 - 2.6.5 Maintains a lively pace
 - 2.6.6 Demonstrates intensity and enthusiasm
 - 2.6.7 Uses friendly controversy
 - 2.6.8 Provides opportunities for students to talk about themselves
 - 2.6.9 Presents unusual or intriguing information

#5 SAFE, POSITIVE LEARNING ENVIRONMENT

- 5.3 Demonstrates awareness of classroom environment at all times (withitness) **0Q7**
- 5.4 Applies consequences for lack of adherence to rules and procedures **0Q7**
- 5.5 Acknowledges adherence to rules and procedures **0Q7**
- 5.6 Builds positive relationships with students by displaying objectivity and control **0Q8**

#1 HIGH EXPECTATIONS FOR STUDENT ACHIEVEMENT

- 1.3 Understanding students' interests and backgrounds (positive relationships) **0Q8**
- 1.4 Demonstrates value and respect for all, including typically underserved students **0Q9**

#2 EFFECTIVE TEACHING PRACTICES

- 2.4 Asks questions of typically underserved students with the same frequency and depth as other students **0Q9**
- 2.5 Probes typically underserved students' incorrect answers **0Q9**

Appendix D - Danielson Framework

Charlotte Danielson's FRAMEWORK FOR TEACHING

DOMAIN 1: Planning and Preparation

1a Demonstrating Knowledge of Content and Pedagogy

- Content knowledge
- Prerequisite relationships
- Content pedagogy

1b Demonstrating Knowledge of Students

- Child development
- Learning process
- Special needs
- Student skills, knowledge, and proficiency
- Interests and cultural heritage

1c Setting Instructional Outcomes

- Value, sequence, and alignment
- Clarity
- Balance
- Suitability for diverse learners

1d Demonstrating Knowledge of Resources

- For classroom
- To extend content knowledge
- For students

1e Designing Coherent Instruction

- Learning activities
- Instructional materials and resources
- Instructional groups
- Lesson and unit structure

1f Designing Student Assessments

- Congruence with outcomes
- Criteria and standards
- Formative assessments
- Use for planning

DOMAIN 2: The Classroom Environment

2a Creating an Environment of Respect and Rapport

- Teacher interaction with students
- Student interaction with students

2b Establishing a Culture for Learning

- Importance of content
- Expectations for learning and achievement
- Student pride in work

2c Managing Classroom Procedures

- Instructional groups
- Transitions
- Materials and supplies
- Non-instructional duties
- Supervision of volunteers and paraprofessionals

2d Managing Student Behavior

- Expectations
- Monitoring behavior
- Response to misbehavior

2e Organizing Physical Space

- Safety and accessibility
- Arrangement of furniture and resources

DOMAIN 4: Professional Responsibilities

4a Reflecting on Teaching

- Accuracy
- Use in future teaching

4b Maintaining Accurate Records

- Student completion of assignments
- Student progress in learning
- Non-instructional records

4c Communicating with Families

- About instructional program
- About individual students
- Engagement of families in instructional program

4d Participating in a Professional Community

- Relationships with colleagues
- Participation in school projects
- Involvement in culture of professional inquiry
- Service to school

4e Growing and Developing Professionally

- Enhancement of content knowledge / pedagogical skill
- Receptivity to feedback from colleagues
- Service to the profession

4f Showing Professionalism

- Integrity/ethical conduct
- Service to students
- Advocacy
- Decision-making
- Compliance with school/district regulation

DOMAIN 3: Instruction

3a Communicating With Students

- Expectations for learning
- Directions and procedures
- Explanations of content
- Use of oral and written language

3b Using Questioning and Discussion Techniques

- Quality of questions
- Discussion techniques
- Student participation

3c Engaging Students in Learning

- Activities and assignments
- Student groups
- Instructional materials and resources
- Structure and pacing

3d Using Assessment in Instruction

- Assessment criteria
- Monitoring of student learning
- Feedback to students
- Student self-assessment and monitoring

3e Demonstrating Flexibility and Responsiveness

- Lesson adjustment
- Response to students
- Persistence

Appendix E - IRB Authorization



TO: Jeffrey Zacharakis
Manhattan, KS 66506

Proposal Number: IRB-10789 Educational Leadership

FROM: Rick Scheidt, Chair
Committee on Research Involving Human Subjects

DATE: 08/10/2021

RE: Proposal titled "Changes in Perspective Regarding Washington State Teacher Principal Evaluation Protocol (TPEP)
/EDACE999"

The Committee on Research Involving Human Subjects/Institutional Review Board (IRB) for Kansas State University has reviewed the proposal identified above and has determined that it is EXEMPT from further IRB review. This exemption applies only to the proposal as written and currently on file with the IRB. Any change potentially affecting human subjects must be approved by the IRB prior to implementation and may disqualify the proposal from exemption.

Based upon information provided to the IRB, this activity is exempt under the criteria set forth in the Federal Policy for the Protection of Human Subjects, **45 CFR §104(d), category: Exempt Category 2 Subsection ii.**

Certain research is exempt from the requirements of HHS/OHRP regulations. A determination that research is exempt does not imply that investigators have no ethical responsibilities to subjects in such research; it means only that the regulatory requirements related to IRB review, informed consent, and assurance of compliance do not apply to the research.

Any unanticipated problems involving risk to subjects or to others must be reported immediately to the Chair of the Committee on Research Involving Human Subjects, the University Research Compliance Office, and if the subjects are KSU students, to the Director of the Student Health Center.

Electronically signed by Rick Scheidt on 08/10/2021 2:21 PM ET

203 Fairchild Hall, Manhattan, KS 66502 | (785) 532-3224 | fax: (785) 532-3278