

May 2015

Negative Impacts of Teacher Evaluation Reform: The Role of Federal Policy in Student Achievement and Teacher Quality

Allison Smith
University of Nevada, Las Vegas

Follow this and additional works at: <https://digitalscholarship.unlv.edu/thesesdissertations>



Part of the [Bilingual, Multilingual, and Multicultural Education Commons](#), [Education Policy Commons](#), and the [Teacher Education and Professional Development Commons](#)

Repository Citation

Smith, Allison, "Negative Impacts of Teacher Evaluation Reform: The Role of Federal Policy in Student Achievement and Teacher Quality" (2015). *UNLV Theses, Dissertations, Professional Papers, and Capstones*. 2431.
<http://dx.doi.org/10.34917/7646062>

This Dissertation is protected by copyright and/or related rights. It has been brought to you by Digital Scholarship@UNLV with permission from the rights-holder(s). You are free to use this Dissertation in any way that is permitted by the copyright and related rights legislation that applies to your use. For other uses you need to obtain permission from the rights-holder(s) directly, unless additional rights are indicated by a Creative Commons license in the record and/or on the work itself.

This Dissertation has been accepted for inclusion in UNLV Theses, Dissertations, Professional Papers, and Capstones by an authorized administrator of Digital Scholarship@UNLV. For more information, please contact digitalscholarship@unlv.edu.

NEGATIVE IMPACTS OF TEACHER EVALUATION REFORM: THE ROLE OF
FEDERAL POLICY IN STUDENT ACHIEVEMENT AND TEACHER QUALITY

By

Allison Leigh Smith

Bachelor of Arts in Communication
University of California, Santa Barbara
2002

Master of Education
California Lutheran University
2007

A dissertation submitted in partial fulfillment
of the requirements for the

Doctor of Philosophy -- Teacher Education

Department of Teaching and Learning
College of Education
The Graduate College

University of Nevada, Las Vegas
May 2015

Copyright by Allison Leigh Smith, 2015

All Rights Reserved



We recommend the dissertation prepared under our supervision by

Allison Leigh Smith

entitled

Negative Impacts of Teacher Evaluation Reform: The Role of Federal Policy in Student Achievement and Teacher Quality

is approved in partial fulfillment of the requirements for the degree of

Doctor of Philosophy - Teacher Education

Department of Teaching and Learning

Christine Clark, Ed.D., Committee Chair

Marilyn McKinney, Ph.D., Committee Member

Shaoan Zhang, Ph.D., Committee Member

LeAnn Putney, Ph.D., Graduate College Representative

Kathryn Hausbeck Korgan, Ph.D., Interim Dean of the Graduate College

May 2015

Abstract

The purpose of this sequential transformative study was to elucidate the negative experiences of teachers with performance evaluations and to juxtapose the intended use of current popular teacher evaluation reform movements to the evident implementation. One may quickly assume that negative experiences with evaluation are a result of unsatisfactory teaching practices. However, this may not accurately explain the negative experiences. This study focused on the negative experience of teacher evaluation to provide a broader understanding of the impact of new evaluation policy reform on student achievement and teacher quality. Five data sources were analyzed: PAR policy, a teacher survey, district PAR data, individual interviews, and one group interview. Five teachers, bound by one school district, participated in the interviews.

With a paucity of previous research focused on the negative impacts of teacher evaluation, this study addressed the following questions: (1) How does the PAR teacher evaluation process negatively impact teachers? (2) What, if any, parallel traits exist among those teachers who had negative experiences with the PAR evaluation system? (3) How does the intended use of the PAR teacher evaluation process compare to the evident use of PAR? Data were collected and analyzed using the Complementary Analysis Research Matrix Application (CARMA) as an orienting tool to juxtapose data collected in interviews, surveys, and public data sources. CARMA aligned with the overarching theoretical framework of critical multiculturalism by systematically collecting data and analyzing it in an organized approach that hoped to empower participants to share their experiences. The study aligned with critical multiculturalism by giving voice to a marginalized group and empowering underrepresented people.

Data revealed dissonance among intent and evident use of the evaluation policy. A disproportionate number of African Americans, women over the age of 55, and teachers higher on the pay scale were referred to PAR. Additionally, data indicated that through the PAR evaluation system, teachers experienced negative health impacts such as, high levels of stress, depression, weight gain, despondence, insomnia, hostility toward family, deteriorating relationships, high blood pressure, and overall distraction. Furthermore, vague policy language was suggested as the impetus for misuse, abuse, and biased implementation at the local level. This study suggests that policy makers and school district officials take heed of multiple perspectives and consider the negative impacts of teacher evaluation reform. Evaluation systems that prioritize teacher learning over accountability are integral to successfully improving student achievement.

Keywords: Peer evaluation, teacher evaluation, negative experiences of teachers, PAR, critical multiculturalism.

Acknowledgements

I am grateful to my advisor, Dr. Christine Clark, who challenged and encouraged me throughout my doctoral coursework and dissertation. With high expectations and a huge heart, Dr. Clark enabled me to see my potential. A critical moment happened when she invited me to dinner with Dr. Sonia Nieto, an amazing academic role model of mine, who gave the directive that Dr. Clark must be my committee chair. That is when everything seemed to progress in my research. Dr. Clark's critical perspective and passionate approach to social justice and multicultural education was an inspiration. I could not have finished this dissertation without her mentorship.

I am also hugely indebted to Dr. Alexandru Spataru who pushed me to finish my master's thesis, start my doctoral program, and finish my dissertation. Alex has been there the entire time demanding progress. Another huge supporter has been my dear friend Judith Raimondi. Judith has always shown her love for me and was the only person I trusted to take the first look at my writing. Her feedback and edits were critical to the progress of this study.

I am also thankful to my committee for their wonderful feedback and ideas. I thank all my participants for discussing the emotional topics in this study. Brian Crowell has also been extremely helpful throughout the entire study by providing feedback and data sources to consider. Conversations with Brian sparked many ideas for research and further research.

Lastly, I am forever grateful to KG for inspiring this entire line of research. Her evaluation methods made me realize the importance of studying this topic. Without KG, this research would have never happened.

Dedication

Dedicated to my loving mother, father, and son. They have always been my biggest supporters who motivated me to keep moving forward with strength, determination, and heart. Zico, my energetic and loving son, has been my inspiration, as I became Doctor Mommy. His happy soul gave me strength through the toughest times.

Table of Contents

| | |
|--|-----------|
| Abstract | ii |
| Acknowledgements | v |
| Dedication | vi |
| Table of Contents | vii |
| List of Tables | x |
| List of Figures | xi |
| Chapter 1: Introduction | 1 |
| Background | 2 |
| Problem Statement | 3 |
| Purpose | 3 |
| Definitions | 5 |
| Theoretical Framework | 6 |
| Brief Review of Sequential Transformative Design | 8 |
| Research Questions | 8 |
| Nature of the Study | 8 |
| Significance of the Study | 9 |
| Assumptions of the Study | 10 |
| Limitations | 12 |
| Chapter Summary and Transition | 12 |
| Chapter 2: Literature Review | 14 |
| Policy | 16 |
| Lead-up to Federal Policy Reform | 18 |
| No Child Left Behind Act | 20 |
| Policy Reform to Address NCLB and Teacher Evaluation | 21 |
| Waivers to NCLB | 21 |
| Competitive incentive grants | 25 |
| Connection to the Study | 26 |
| Scholarly Response to Government Reform | 27 |
| Teacher Quality | 33 |
| Positive Impacts of Teacher Evaluation | 39 |
| Negative Impacts of Teacher Evaluation | 44 |
| Research from Private Interest Groups | 46 |
| Summary Discussion: Shift in Accountability | 48 |
| Summary Implications for the Study: Marginalized Group of Teachers | 51 |
| Chapter 3: Methodology | 55 |
| Restatement of Problem | 55 |
| Restatement of Purpose | 56 |
| Review of Research | 56 |
| Research Questions | 58 |
| Data Sources | 58 |
| Data Analysis | 58 |
| Timeline | 58 |
| Research Questions | 59 |
| Sequential Transformative Design | 59 |
| Phase One | 61 |
| Phase Two | 62 |

| | |
|--|-----|
| Sampling and Participants | 62 |
| Data Collection | 63 |
| Data Collection using CARMA | 64 |
| Data Sources | 65 |
| Teacher survey | 66 |
| District data | 66 |
| California Peer Assistance and Review (PAR) policy | 67 |
| Interviews | 67 |
| Data Analysis | 68 |
| Data Analysis using CARMA | 69 |
| Category Construction | 69 |
| Timeline | 70 |
| Role of the Researcher | 70 |
| Trustworthiness | 71 |
| Chapter Summary and Transition | 73 |
| CHAPTER 4: Findings of the Study | 74 |
| Introduction of Emergent Themes | 75 |
| Alignment to Research Questions | 78 |
| Phase One | 78 |
| CARMA Matrix for Phase One | 79 |
| California Peer Assistance and Review (PAR) Policy | 81 |
| Unclear expectations | 83 |
| District Data | 86 |
| Teacher Survey | 88 |
| Phase Two | 91 |
| Individual Interviews | 92 |
| Reggie's interview | 92 |
| Tina's interview | 99 |
| Ana's interview | 104 |
| Dawn's interview | 109 |
| Missing data | 113 |
| Group Interview | 113 |
| Emotional overload | 115 |
| Privatization and standardization of education | 116 |
| Inconsistent and unclear expectations | 117 |
| Outliers | 120 |
| Major Themes | 120 |
| The Scarlet P (PAR) | 120 |
| Kangaroo Court | 122 |
| The Salem Witch Trials | 124 |
| Discrepant Data | 126 |
| Discrimination Suggested Through CARMA | 127 |
| Making Connections | 128 |
| Research Questions | 129 |
| Theoretical Framework | 131 |
| Teacher Education | 133 |
| Missing Data | 134 |
| Collaboration | 135 |
| Conclusion | 136 |
| Chapter 5: Implications | 138 |

| | |
|--|-----|
| Review of Research Questions | 140 |
| Interpretations of Findings..... | 140 |
| Major Themes | 145 |
| Theme One: The Scarlet P..... | 145 |
| Theme Two: The Kangaroo Court | 146 |
| Theme Three: The Salem Witch Trials | 146 |
| Theme Four: Nonconfirming Data | 147 |
| Unexpected Results..... | 147 |
| Implications in Teacher Education | 149 |
| Limitations | 151 |
| Recommendations for Further Study | 152 |
| Researcher Reflections | 156 |
| Conclusion | 157 |
| References..... | 159 |
| Appendix A: Complementary Analysis Research Matrix Application (CARMA) | 166 |
| Appendix B: 2012-2013 Teacher Survey Questions | 167 |
| Appendix C: District PAR Data..... | 168 |
| Appendix D: Interview Questions | 171 |
| Appendix E: Informed Consent | 173 |
| Appendix F: Thematic Analysis Tables..... | 175 |
| Appendix G: CARMA Worksheets | 190 |
| Curriculum Vitae | 195 |

List of Tables

| | |
|---|----|
| Table 1. Review of the Research Process | 58 |
| Table 2. Integration of CARMA Matrix and Sequential Tranformative Design..... | 80 |
| Table 3. Education Code Analysis..... | 85 |
| Table 4. Binomial Probability..... | 88 |

List of Figures

| | |
|---|-----|
| Figure 1. Theoretical framework, methodology, and the topic of teacher evaluation | 57 |
| Figure 2. Phases of a sequential transformative design | 60 |
| Figure 3. Data sources..... | 66 |
| Figure 4. Emergent themes and their major attributes | 77 |
| Figure 5. Major findings and their alignment to the research questions..... | 141 |

Negative Impacts of Teacher Evaluation Reform: The Role of Federal Policy in Student Achievement and Teacher Quality

Chapter 1: Introduction

This study aimed to increase understanding about current teacher evaluation reform and the potentially negative impacts that new methods of evaluation could have on teachers. Recently, federal policy focused on teacher evaluation to address low student achievement. Many states chose to aggressively reform their previous ways of evaluating teachers to adhere to federal incentives and awards. However, current research on education does not indicate that teacher evaluation is the best means to address student achievement. Education scholars have identified other empirically supported ways of improving student achievement, such as funding after-school tutoring, equalizing resources, and improving professional development for teachers (references).

Overall, I noted an absence of existing research focused on teacher evaluation within the current political climate. Among the limited existing research, studies had only focused on the positive aspects of evaluation methods (Goldstein, 2004; SRI, 2012). This study sought to address the limited perspective by elucidating the negative experiences of teachers in the evaluation processes promoted by current political reform.

One may quickly assume that negative experiences with evaluation are a result of unsatisfactory teaching. However, this may not accurately explain the negative experiences. This study focused on the negative experience to provide a broader understanding of the impacts of new evaluation policy reform.

This chapter provides the background, problem statement, purpose, and significance. Additionally, key terms are defined, the theoretical framework is explained,

the research questions are stated, and a brief overview of the methodology is explained. Chapter two reviews previous literature about teacher evaluation. Chapter three explains the methodology. Chapter four includes the findings and chapter five concludes with the interpretation and implications of the results.

Background

In 2001 the newly enacted No Child Left Behind Act (NCLB) changed the education system in the United States by setting strict student achievement levels and increasing school accountability. NCLB mandates stated that by the 2013-2014 school year, all students would be proficient in math and English. With these strict mandates not achieved, the federal government responded by shifting policy reform to teacher accountability through increased teacher evaluation as the means to improve student achievement. Most teacher evaluation methods shifted from the school principal as the primary evaluator, to peer evaluation, such as the Peer Assistance and Review (PAR) system (National Council on Teacher Quality [NCTQ], 2012). This shift in accountability, from school to teacher and principal to peer evaluation, was a newer phenomenon in education reform that prominent education scholars suggest necessitates further empirical research within teacher education (Darling-Hammond, 2013).

Education scholars have conducted powerful and insightful research focused on improving student achievement and teacher quality (Cochran-Smith & Lytle, 2006; Darling-Hammond, 2013; Hill & Grossman, 2013; Ladson-Billings, 2005; Nieto, 2006; Noguera, 2003; Ravitch, 2010). However, their findings and suggestions have not been central to policy implementation. This is evidenced by the movement toward increased teacher evaluation without a solid research foundation to support the ways evaluation

systems were reformed and implemented.

Schools, once held under strict accountability, shifted accountability to teachers, to appeal to federal initiatives. Evaluation systems, such as PAR, were implemented to meet federal requirements for waivers and grants because of the perceived benefits of a peer-assessment design. PAR has received positive reviews from researchers, private research groups, policymakers, and community members. However, research remains limited in this area and previous studies only focused on districts where PAR was successful (Goldstein, 2004; SRI, 2012).

Problem Statement

The United States has consistently underperformed on international standardized exams, such as TIMSS and PISA, and state and national exams (Organization for Economic and Co-operation and Development [OECD], 2012; and Trends in International Mathematics and Science Study [TIMSS] (2011). Education reform policy began to focus on teacher evaluation to address low student achievement. However, there was limited research about potential impacts of teacher evaluation. Within existing research about teacher evaluation, studies focused on positive experiences with evaluation reform initiatives. For this study, the problem of not having a broad perspective or research base to inform the ways policy surrounding teacher evaluation reform were implemented was central.

Purpose

The purpose of this study was two-fold and exploratory in nature due to the lack of existing empirical research. Firstly, I aimed to elucidate the negative evaluation experience to broaden the understanding of potentially negative impacts of evaluation

reform policy. This study intended to provide more research about teacher evaluation in the current political climate to better inform policymakers and teacher preparation programs.

Secondly, the study juxtaposed the intended use of a current teacher education reform movement, Peer Assistance and Review (PAR), to the evident use in schools through an exploration of teachers' negative evaluation experiences. Comparing and contrasting the intended and evident use of policy revealed areas of concern and areas that necessitate further reform and/or research.

Through this two-fold purpose, the study aimed to inform teacher preparation programs that educate pre-service teachers. Connecting preparation program design to in-service performance evaluation could improve successful entry and retention in the field of education. Currently, among the research, the connection between teacher evaluation reform and teacher preparation programs is not clear.

This research utilized a sequential transformative method design to support an overarching critical multiculturalist theoretical framework. In a sequential transformative study, participants engage in co-constructing data analysis and results, to ensure accuracy and promote empowerment. Likewise, a critical multiculturalist framework aimed to give voice to marginalized groups and empower underrepresented people.

As a teacher who fell victim to the misuse of the teacher evaluation process, I was passionate about exploring this topic and providing policymakers with more insight about contrastive perspectives of teacher evaluation. As a passionate educator, I resigned from a school district that used the PAR evaluation system for punitive purposes. Initially, I internalized the demoralizing nature of the evaluation process, but later found out that I

was not alone in my experience. Many dedicated, quality teachers were also feeling harassed through the district's teacher evaluation system. This sequential transformative study took an in-depth analysis of the negative evaluation experiences of teachers in a school district in Northern California that has utilized the PAR evaluation system for over ten years. Potential insider bias is later discussed in the methodology section because research design intentionally worked to mitigate bias.

Definitions

This study used educational terms in specific ways for the topic of teacher evaluation. The operational definitions are:

- Evaluation: References to evaluation include in-service assessments of teacher performance, significance, value, and worth. Typically evaluation is done through thoughtful consideration and careful appraisal. However, for this study, that was not a component of the definition. The degree, or presence, of thoughtful consideration and careful appraisal shall be discussed in the results of this study and future research focused on teacher evaluation (Merriam-Webster, n.d.).
- Highly Qualified Teacher (HQT): This term is used with the same definition as noted in the No Child Left Behind act of 2001, "To be deemed highly qualified, teachers must have: 1) a bachelor's degree, 2) full state certification or licensure, and 3) prove that they know each subject they teach" (USDOE, 2005).
- Quality Teacher: This term was used through the lens of a critical multiculturalist theoretical framework. A quality teacher is one who embodies

social justice teaching. Rethinking Schools further defines this as teaching that is, “grounded in the lives of our students, critical, multicultural, anti-racist, pro-justice, participatory, experiential, hopeful, visionary, activist, academically rigorous, and culturally and linguistically sensitive” (2006).

- Student Achievement: In this study, student achievement refers to how well students perform on standardized and school tests. This definition contrasts the critical multiculturalist framework. However, it is in line with the federal mandates of NCLB and the movement toward Common Core State Standards.

Theoretical Framework

As previously mentioned, this study was framed with the perspective of empowering marginalized groups through critical multiculturalism. McDowell and Fang (2007) define critical multicultural research as, “research that is (a) informed by critical, feminist, and multicultural theories; (b) supportive of equity and inclusion; and (c) centered on the concerns of those inhabiting traditionally marginalized and oppressed social locations” (p. 551). This study amplified the negative experience of teachers who had negative experiences with the PAR evaluation system.

“Critical multiculturalism draws from critical and multicultural perspectives as well as feminist discourses on identity politics and social location to support racial, ethnic, and cultural equity” (McDowell & Fang, 2007, p. 552). This study investigated parallel attributes of those who have been negatively impacted by PAR and the potential dissonance between policy and action. Furthermore, the ways in which the inequalities of power within the school system impact the teacher evaluation process, and the inter-

group relation between administrators and teachers was also a focus of this study. These goals were guided by the overarching critical multiculturalism theoretical framework.

Additionally, the study used Darling-Hammond's (2013) most recent publication as a pillar among existing empirical research about teacher evaluation. Darling-Hammond's guidelines about the problems with the current evaluation approaches and guidelines to the ideal evaluation system were initially used as a guide to data collection and analysis as I conducted an in-depth analysis of the negative experiences of teachers through evaluation processes. However, as data were analyzed, the study diverged from Darling-Hammond's guidelines and critical multiculturalism became the sole theoretical framework.

Darling-Hammond (2013) provided six guidelines about variables that were innate problems with current evaluation systems. These guidelines included:

- Lack of consistent, clear standards of good practice
- No focus on improving practice
- Inadequate time and staff for effective evaluations
- Little or no consideration of student outcomes
- Cookie-cutter procedures that don't consider teacher needs
- Detachment of evaluations from professional development

Darling-Hammond (2013) suggested that the components of PAR were the ideal evaluation system. Her guidelines, and support of PAR framed the initial data collection and analysis. However, a shift from her guidelines took place later in the study due to divergent data and is explained in chapter four.

Brief Review of Sequential Transformative Design

This study utilized a sequential transformative design intended to empower participants to be agents of change through a quantitative phase, followed by a qualitative phase (Creswell, 2009). This methodology supported a critical multiculturalist framework with the dual focus on empowerment. Data sources in the first quantitative phase of the study were district data, an existing anonymous teacher survey, and state PAR policy. The data sources in the second qualitative phase were intended to be six individual interviews and one group interview. However, due to participant drop out, the data consisted of four individual interviews and one group interview with three participants. As stated in the sequential transformative design, data from both phases were analyzed together and co-constructed with participants to formulate results and recommendations. The goal of this design was to empower participants to change their current situation (Creswell, 2009).

Research Questions

Three research questions guided this study: (1) How does the PAR teacher evaluation process negatively impact teachers? (2) What, if any, parallel traits exist among those teachers who had negative experiences with the PAR evaluation system? (3) How does the intended use of the PAR teacher evaluation process compare to the evident use of PAR? My research aimed to elucidate marginalized perspectives to better inform policy makers about a topic that was central to current education reform and ultimately impacted students.

Nature of the Study

This study, focusing on a sub-group of teachers who had negative experiences

with the evaluation process, utilized a sequential transformative design (Creswell, 2009). This study occurred in a school district in Northern California that utilized the PAR teacher evaluation system. The school district policies and procedures about PAR were the same for all participants and data analysis was focused on the intended and evident use of PAR, as well as the experiences of the participants. This study utilized the Complementary Analysis Research Matrix Application (CARMA) (Putney, Wink, & Perkins, 2006) as an orienting tool to juxtapose data collected in interviews, surveys, and public data sources. Triangulation of these data sources increased reliability and provided depth to the findings. CARMA was utilized throughout data collection and analysis. CARMA aligned with the overarching theoretical framework of critical multiculturalism by systematically collecting data and analyzing it in an organized approach that hoped to empower participants to share their experiences and “use the research to improve their educational space” (Putney et al., 2006). For this study, CARMA was utilized in the adapted format to better fit the research questions and methodology. Originally, CARMA was referred to as the Critical Action Research Matrix Application. This study was not action research and therefore the adapted form of CARMA, as the Complementary Analysis Research Matrix Application, was more appropriate.

Significance of the Study

Overwhelming attention aimed at improving student achievement, focused primarily on teacher evaluation Federal waivers to No Child Left Behind and other financial incentives hinged upon teaching evaluation reform at the state level. It was evident that the way teachers were evaluated and overall teacher quality was of great concern to our nation. People were demanding a better education system and

policymakers were looking at teacher quality as the primary source of improving student learning. Education researchers have been attentive to the situation surrounding teacher quality and evaluation. Most of the prominent teacher education and multicultural education scholars reported major deficiencies in, and resulting from, NCLB. However, there still remained limited empirical research about teacher evaluation.

This study provided a broader understanding of the impacts of teacher evaluation reform. Within the limited body of research about teacher evaluation, policymakers did not have a good working knowledge, or foundation, to make decisions that have significant impact, positive or negative, on students, teachers, teacher preparation programs, school systems, and the teaching profession as a whole.

Within the field of teacher education, this study addressed the dissonance between teacher evaluation policy reform and teacher preparation program design. The current reform movements did not address or incorporate the necessary connection between teacher education at the pre-service level and in-service teacher evaluation. Teacher education programs could benefit by aligning with teacher evaluation systems, or vice-versa, to prepare teachers to succeed within the education system and teaching profession. This study provided vital information to help inform teacher education programs and policymakers about preparing teachers and using evaluation effectively to increase student achievement.

Assumptions of the Study

This study operated within a specific context, with numerous elements that were understood to be true. This section articulates the major assumptions about the conceptual framework, phenomenon, methodology, participants, and anticipated results.

The first major assumption was that current reform movements and policy centered on teacher evaluation were imperfect. I framed the study by using Linda Darling-Hammond's most recent publication about teacher evaluation, *Getting Teacher Evaluation Right: What Really Matters for Effectiveness and Improvement* (2013). It was assumed that Darling-Hammond was a lead scholar within teacher education and her research was grounded in years of scholarly work within education. Using her research as a framework to guide this study was assumed to be an appropriate foundation.

In addition, I used critical multiculturalist theory to frame this study. A major assumption was that teachers who had negative experiences with teacher evaluation were marginalized and voiceless. As an aim to elucidate their unique experiences, I assumed that this research worked toward a more just educational system that supported teachers and students of diverse backgrounds.

This study utilized a sequential transformative design and the major assumption was that this methodology delved deeper in to the experiences of teachers and added to the body of research. I assumed that a sequential transformative design would reveal more detail than previously utilized methodology, such as survey method, that existed within current research about teacher evaluation.

It was also assumed that the teachers that were chosen to participate in the study expressed experiences that help the readers understand the perspective of a marginalized group who had negative experiences with teacher evaluation. The results were assumed to better inform policy makers, administrators, and districts to create more appropriate evaluation systems that maintained focus on student achievement.

Limitations

Previous research in this area utilized survey method and has not accounted for multiple perspectives of teachers who have had negative experiences through the teacher evaluation process. Researchers have analyzed data as a whole and not disaggregated variables during analysis. Their research has produced findings that do not express multiple perspectives because they are represented as one unified experience (Goldstein, 2004). My research addressed this limitation by focusing on the negative experience only and elucidating this specific experience.

However, limitations were evident in this study. This study was bound by one district. More insight to the negative experiences of teacher evaluation may be achieved with a broader reach to multiple districts. The data analysis procedures may have also been a limitation. As the only researcher, my interpretations and analysis of the interviews, surveys, and district data may be subjective. I addressed this limitation by including the participants' feedback with the data analysis tool, CARMA.

Lastly, saturation may have been an issue. By design, this study focused on four teachers through in-depth interviews. More interviews and additional questions may have been necessary if enough data were not collected in the initial interviews. More participants may have increased significance of the findings.

Chapter Summary and Transition

This chapter introduced the current education reform focus on teacher evaluation to address low student achievement. Chapter one has also provided a rationale for exploring a broader perspective of teacher evaluation reform policy and the implications it has on teaching and learning through a sequential transformative design.

Chapter two will review the existing research about teacher evaluation and will identify themes that emerged through a thorough review of literature. The review discusses policy build-up to and results of the No Child Left Behind Act that served as the entry point to discussing teacher evaluation in this study. Chapter two furthers the review with literature that supported teacher evaluation, literature that focused on the negative impacts of teacher evaluation reform, and research from private research groups. Chapter three provides specific details about the methodology design of the study and the use of the CARMA data collection and analysis tool that supported a critical multiculturalist theoretical framework. Chapters four and five report the findings and discuss the interpretations and implications.

Chapter 2: Literature Review

The education system in the United States focused on teacher evaluation as the primary means to address low student achievement nationwide. However, research to support the connection between teacher evaluation and student achievement was limited. The limited perspective and knowledge about teacher evaluation left a group of teachers who had negative teacher evaluation experiences marginalized. The teachers who had negative experiences were assumed to be bad teachers and they remained voiceless in research and policy.

This study elucidated negative evaluation experiences and juxtaposed the intended use of the policy with the evident experiences. The Peer Assistance and Review (PAR) evaluation policy was central to this study because teacher evaluation through peer review was favored in current federal policy reform (USDOE, 2009; USDOE, 2012, p. 3; USDOE, 2013). Ultimately, this study aimed to better inform policymakers, teacher preparation programs, and empower participants to change their negative experiences to positive.

Three research questions guided this study: (1) How does the PAR teacher evaluation process negatively impact teachers? (2) What, if any, parallel traits exist among those teachers who had negative experiences with the PAR evaluation system? (3) How does the intended use of the PAR teacher evaluation process compare to the evident use of PAR? These questions aligned with the emerging themes from the review of literature through a focus on a marginalized perspective that was missing in the existing research and broadening the overall body of research focused on teacher evaluation.

The review of literature revealed dissonance between major federal reform policy

that implicated teacher evaluation as the primary means to improve student achievement and what research and leading scholars revealed about ways to improve teacher quality and student achievement. The research questions for this study took a closer look at the PAR evaluation policy and addressed the perspective of the negative experiences of teacher evaluation, which tied closely with the critical multiculturalist framework to give voice to marginalized groups.

This study was approached through a critical multiculturalist theoretical framework, with the intent to focus on illuminating the experiences of a marginalized group. Specifically, this study gave voice to teachers who had negative experiences with teacher evaluation. The study was also partially guided by Darling-Hammond's (2013) guidelines to improving teacher evaluation. As a prominent teacher education scholar, her work on teacher evaluation set the standard for best practices from the perspective of teacher education. In reviewing the literature, this perspective aligned with the synthesis and analysis of existing sources, as well as the inclusion of the perspectives of prominent multicultural education scholars.

In a massive search of empirical research, very few articles in peer-reviewed journals were found that addressed teacher evaluation in the current political climate. Through an online search in the education database of scholarly and popular literature, ERIC, 19,080 search results appeared with the search phrase 'teacher evaluation'. Despite many articles listed, most articles did not relate specifically to this topic or they were not from peer-reviewed journals. It did not appear that any of the articles focused on the negative experiences of teacher evaluation. By further limiting sources to peer reviewed journals from the previous five years, 961 articles appeared. Among these articles, I

chose studies that were aligned closest with my research topic. Other articles were not chosen for review for various reasons, including limitation in scope (subject-specific or focused on a specific system of evaluation), not set in the context of the United States, or focused on mentoring or professional development. Articles from private interest research groups dominated the search results. Also, many articles focused on evaluating students despite a search for ‘teacher evaluation’. With an ERIC search using the phrase ‘peer evaluation’, 4,287 results appeared. However, results ranged from online peer evaluation to peer evaluation in higher education. These results did not connect with my research topic. Other combinations of key words included ‘teacher quality evaluation’, ‘teacher quality’, and ‘Peer Assistance and Review’. These searches revealed studies that were off topic and had a more international focus rather than teacher evaluation within the current political climate of the United States. Consequently, to expand this review of literature, I chose to include a broader scope of articles and organize them by emerging themes. The emerging themes included: policy, scholarly response to government reform, teacher quality, positive impact of teacher evaluation, negative impact of teacher evaluation, and research from private interest groups.

Policy

Over a decade ago, the federal government amended the Elementary and Secondary Education Act (ESEA) of 1965 to address the continued and unequal access to education. The controversial No Child Left Behind Act amended ESEA and prioritized accountability in the quest to improve student achievement, locally and globally. However, educational accountability was not new to the education reform scene, as teacher evaluation and teacher quality have long been central issues in the field of

education for researchers, communities, and the federal government (Darling-Hammond, Wise, & Pease, 1983). With wide-spread reports about the failing American education system ultimately prompting NCLB, from sources such as The National Commission on Teacher and America's Future (1996), The Goals 2000: Educate America Act (USDOE, 1994), and the National Commission on Excellence in Education published the report, A Nation at Risk (1983), it has been evident that education reform in the United States was necessary. This was especially true when the international test scores of the United States, such as TIMSS and PISA, were compared to successful turnaround countries such as Finland, who out-performed all other countries in the world with their high student achievement.

For the United States, teacher evaluation was the central focus of reform to address low student achievement. The federal government rewarded states that increased and changed teacher evaluation practices (McGuinn, 2012). Teacher education and multicultural education scholars confirmed that this may be a means to improve student achievement, as long as the evaluation systems stimulated teaching practices that promoted student achievement and developed teacher competence (Darling-Hammond, 2013). Unfortunately, empirical research about current trends in teacher evaluation, including research that represents multiple perspectives of those directly impacted by policy changes surrounding teacher evaluation, was missing. Consequently, it was necessary to conduct more research to ensure evaluation practices were used to promote student achievement and foster high quality teaching.

Teachers who had negative experiences with teacher evaluation have been marginalized. Their perspective was missing from the research and their experiences may

have been assumed to be connected to unsatisfactory teaching. Exploring their perspective and the potential misuse of new evaluation practices could improve evaluation practices. The following section outlines the policy reform and national reports that led up to focusing on teacher evaluation, as well as anecdotal scholarly criticisms to the current policy reform focused on teacher evaluation. I further this review of literature by examining the existing empirical research that focused on teacher quality, the positive impact of teacher evaluation, and the negative impact of teacher evaluation.

Lead-up to Federal Policy Reform

Many researchers have found numerous variables that directly and indirectly impacted student learning, teacher effectiveness, and evaluation systems (Adamson & Darling-Hammond, 2012; Cochran-Smith & Lytle, 2006; Nieto, 2006; Noguera, 2011; Ravitch, 2010; Rowley & Wright, 2011). However, federal mandates have been the driving force behind a number of aggressive reform movements focused on teacher evaluation. Policy was focused on improving student achievement, but many education researchers questioned the policy changes and have indicated the pitfalls of teacher evaluation systems and policies that directly and indirectly impacted students. Regardless of the various criticisms, most have agreed that teachers do play a large part in student achievement and all students deserve to have effective teachers (Cochran-Smith & Lytle, 2006).

As a catalyst to heightened education reform, the National Commission on Excellence in Education published the report, *A Nation at Risk* (1983). This report initiated an increase in school accountability for student achievement. *A Nation at Risk* outlined deficiencies, low literacy rates, and the frustrations surrounding the loss of

international achievement/competition. This report was an alarm, or wake-up call, to all communities and interest groups in the United States to initiate education reform to better serve the whole population and increase international performance. Reform was suggested at radical rates and an increase in collective accountability among all groups was emphasized (NCEE, 1983). Using the nuclear arms race as the attention-grabber, the nation was made aware of the ‘education crisis’.

The subsequent influential report in national education reform, leading to NCLB (as the point of entry for this research), was The National Commission on Teaching and America’s Future (1996). This report was written to support reform that ensured every child had access to quality teaching. The report was designed to be a blueprint for “recruiting, preparing, and supporting excellent teachers in all of America’s schools” (NCTAF, 1996, p. 10). As a continued movement toward educational excellence across America with an emphasis on collective accountability, the report proposed an audacious goal of providing all students with “competent, caring, qualified teaching in schools organized for success” (NCTAF, 1996, p. 10).

The next influential report was the Goals 2000: Educate America Act that emphasized similar aims as the previous two reports. Goals 2000 focused on improving learning and teaching, “by providing a national framework for education reform” (USDOE, 1994, para. 1). The movement was to gain federal control of schools that were always under state control and legislation. This shifted reform to national systems working to improve student achievement and prompted the infamous No Child Left Behind legislation.

No Child Left Behind Act

In 2001, President George W. Bush enacted an amendment to the Elementary and Secondary Education Act (ESEA, 1965) to address low student achievement. The amendment was entitled No Child Left Behind (NCLB) and claimed to be “a landmark in education reform” (USDOE, 2002, p. 9). The intent was to completely overhaul education in the United States with aggressive goals and mandates centered on accountability. NCLB outlined ten major sections (titles) of reform that integrated four major principles throughout: 1) accountability, 2) flexibility and local control, 3) enhanced parental choice, and 4) focusing on what works.

The “No Child Left Behind: A Desktop Reference 2002” states that:

Federal policy has had a significant impact on America’s schools and children ever since ESEA was enacted in 1965. Yet, despite hundreds of billions of dollars invested during the last generation, American students still lag behind many of their fellow foreign students and the academic achievement gap in this country between rich and poor, white and minority students, remains wide. (USDOE, 2002, p. 9)

NCLB was designed to address these issues. It was the self-proclaimed cornerstone to former President Bush’s administration and referenced the Nation at Risk report (1983) as a catalyst for his strong push for education reform. The theme of enhancing the quality of the nation’s teachers and holding schools accountable was a major focus throughout the NCLB amendment. NCLB defined highly qualified teachers as: having a bachelor’s degree, a teaching credential, and proven content knowledge. Other factors important to the determination of teacher quality were not included in the

legislation, and were the cause of overwhelming criticism from education scholars (Bode & Nieto, 2003; Darling-Hammond, 2012; Rowley & Wright, 2011).

Policy Reform to Address NCLB and Teacher Evaluation

Despite a rather simplistic definition of highly qualified teachers, NCLB outlined comprehensive goals and clear expectations to improve student achievement (Cochran-Smith & Lytle, 2006). The deadline for meeting those goals was set for the 2013-2014 school year. As this deadline quickly approached, states were finding that the majority of schools were not meeting student achievement requirements. Specifically, all students were not proficient in English and math. As a result, a paradigm shift has occurred in student achievement from school accountability to teacher accountability, where the public sector and policymakers were critically examining the quality and effectiveness of the nation's teachers.

Waivers to NCLB. In response to the nation failing to meet NCLB goals, the federal government devised a waiver program that released states from the student achievement requirements, expectations that were once the core of NCLB, in exchange for implementing teacher evaluation systems. Pedro Noguera, a prominent scholar in education reform, stated “it is time for the federal government to go further than to simply allow waivers under the law” (2011).

The NCLB waiver system was the most prominent education reform movement, at the time of this study, and focused on teacher evaluation. This revision to the NCLB Act (2001) was done as an amendment rather than through a complete policy change. President Barack Obama and the Secretary of Education Arne Duncan promoted these waivers of the NCLB requirements to the states instead of rewriting the amendment for

congressional approval. The waiver system was designed as a “choice” for states to make teacher evaluation reform. However, if a state chose not to develop a federally approved teacher evaluation system, then they did not receive a waiver. The waivers released states from the central requirements of NCLB in exchange for other contingencies such as teacher evaluation policy. Most states would rather have the alternative than to uphold the expectations of NCLB, as evidenced by the number of states that applied for waivers.

According to the U.S. Department of Education (USDOE, 2013),

The U.S. Department of Education has invited each State Educational Agency (SEA) to request flexibility regarding specific requirements of the No Child Left Behind Act of 2001 (NCLB) in exchange for rigorous and comprehensive State-developed plans designed to improve educational outcomes for all students, close achievement gaps, increase equity, and improve the quality of instruction. (para. 1)

Specifically, this referred to increasing teacher evaluation policy. If states were approved for the waiver, they were released from the ten provisions outlined in NCLB (2001). The NCLB expectations, that were once determined critical to student achievement, were now either not required or allowed flexibility. These items included:

- Schools would not have to follow Adequate Yearly Progress (AYP) requirements starting in the 2013-2014 school year.
- Schools would not need to identify or implement school improvement actions and they would not need to report anything about school improvement.
- Schools would not be required to identify or take corrective action for schools that do not meet AYP for two years or more and they would not need to report anything about improving schools.
- Rural schools would have flexibility to use funds regardless of AYP.

- Schools would have the option to operate school-wide programs to enhance the entire educational program.
- Schools would have more flexibility in the use of funds to support school improvement.
- Schools would be able to use funds to reward schools financially.
- Schools would not have to meet the Highly Qualified Teacher (HQT) targets or develop improvement plans.
- Districts would be able to transfer monies and not be required to notify state prior to transferring money.
- States would be able to award grant money to any school of its choice.

These were the ten main NCLB requirements that were allowed flexibility under the waivers. However, contingencies were enacted for states to receive these waivers.

The major contingency item in this legislation was “supporting effective instruction and leadership” (USDOE, 2012, p. 3). States must, “commit to develop, adopt, pilot, and implement, with the involvement of teachers and principals, teacher and principal evaluation and support systems” (USDOE, 2012, p. 3). The systems must specifically ensure the following:

- evaluation is used for improvement of instruction.
- differentiate performance using at least three performance levels.
- use multiple valid measures in determining performance levels.
- evaluate teachers and principals on a regular basis.
- provide clear, timely, and useful feedback that guides professional development.
- be used to inform personnel decisions.

Once states developed a teacher evaluation process, with the above items addressed, the federal government granted approval (or required revisions) and the states implemented their plan so they could receive the waiver of flexibility on the ten items previously mentioned. As of July 2013, “47 states, the District of Columbia, Puerto Rico, and the Bureau of Indian Education submitted requests for ESEA flexibility” (USDOE, 2013). Based on these numbers, it was clear that the government had been successful at inviting schools to apply for the waiver to NCLB. With the popularity of the waiver, it also appeared that the original NCLB amendment (2001) to the ESEA (1965) had not reached the success it had intended and the current federal administration had different objectives of education than the Bush administration.

The primary issue with having teacher evaluation programs as the bargaining tool to be released from NCLB requirements was the lack of research to support a connection between teacher evaluation, teacher quality, and student achievement. It was not clear what the motives were for a shift toward teacher evaluation. While it was beyond the scope of this study, it is worth noting that some scholars suggest ulterior motives to teacher evaluation reform, such as dismantling unions, promoting the Common Core movement, and the overall privatization of public education (Ravitch, 2013, 2014; Youngs, 2013).

In addition to teacher evaluation as a contingency item for states to be released from NCLB requirements, the waivers also had a consultation component. The U.S. Department of Education required states to “solicit input from stakeholders representing diverse perspectives, experiences, and interests, including those that will be impacted by...and will strengthen its request [for a waiver] by revising it based on this input”

(USDOE, 2012, p. 9). This provided a foundation for my research. If states needed to hear from people with diverse perspectives, experiences, and interests, it was necessary to hear from the teachers who had negative experiences in the teacher evaluation process. Their voice was minimally heard and often times completely missing in policy and research about teacher evaluation. It may be assumed that their voice represented the underperforming teachers, but it was also imperative to learn about potential misuse of the evaluation process. As policy makers, administrators, and the general public focused on teacher quality as the primary variable that determined student success, teacher evaluation continued to be the central component of reform. This was evidenced by the policy changes that have happened over the last decade.

In addition to NCLB waivers, the federal government had recently implemented policies and grants to forcefully encourage states to increase teacher evaluation. The majority of new legislation targeted teacher evaluation to improve student achievement, including Race to the Top and the Teacher Incentive Fund.

Competitive incentive grants. In November 2009, President Obama challenged states to a competition he called Race to the Top that, “initiated an unprecedented wave of state teacher-evaluation reform across the country” (McGuinn, 2012, p.1). Former President Bush’s call for educational reform did not meet the expectations of many educational researchers, community groups, and government officials (USDOE, 2009). According to President Obama, “It’s time to stop just talking about education reform and start actually doing it. It’s time to start making education America’s national mission” (USDOE, 2009, p. 2). This quote brought to call a challenge designed to reward states that have demonstrated success in improving student achievement and have a future plan

to increase teacher evaluation. Race to the Top specifically implicated teacher evaluation as the way to improve teacher quality and student achievement. Through a two-phase application system, states were challenged to apply and receive funds made available through the American Recovery and Reinvestment Act of 2009. Education was one component of this law enacted to stimulate the economy. The long-term vision of the potential gains of this reform was financially motivated to ensure that students, America's future workers, were productive and effective (USDOE, 2009). This type of policy that challenges states to be the best left some states behind, or "losing". This competitive structure appeared to be a potentially disastrous way to collectively improve the educational experience of all students.

Another competitive grant from the federal government was the Teacher Incentive Fund. The grant program encouraged states to move toward incentive-based pay systems to reward teachers and principals who scored high on performance-based evaluations (USDOE, 2012). Similar to Race to the Top, money was the motivating factor to improve student achievement. The major goals were to improve student achievement by increasing teacher effectiveness, reform teacher compensation systems, increase the number of effective teachers, and create performance-based compensation systems. This grant also rewarded schools that used student test scores to define an effective teacher, as well as multiple classroom observations (USDOE, 2012). Overall, this funding was rather vague and did not clearly define how the multiple components of teacher evaluation would be weighted and used to determine how "effective" a teacher actually was or how it would be used in personnel decisions.

Connection to the Study

The study was built on this body of research focused on policy by looking more closely at the impact of teacher evaluation policy decisions and provided more empirical research for policymakers to consider. This was done by juxtaposing the intended use of evaluation reform policy to the evident implementation. Additionally, this study aimed to broaden the perspectives to be considered with teacher evaluation policy reform.

Scholarly Response to Government Reform

Various scholars in education research suggested major shortcomings of NCLB and the subsequent legislation focused on teacher evaluation (Adamson & Darling-Hammond, 2012; Nieto, 2006; Rowley & Wright, 2011). Scholarly criticism of NCLB suggested that the legislation was unfunded, reactionary, and not based in educational research (Adamson & Darling-Hammond, 2012; Nieto, 2006; Rowley & Wright, 2011). Policymakers took on the challenge of overhauling an education system that was fragmented and disjointed from state to state, district to district, and school to school. However, despite major reform and monetary investment, huge inequalities remained in our education system (Bode & Nieto, 2003).

In a book recently published, within the current political context surrounding teacher evaluation, Linda Darling-Hammond (2013) clearly explained the deficiencies and explained what she felt was necessary in reform movements. Her book analyzed the reform movement focused on teacher evaluation through the lens of teacher education. Her work highlighted the significance and impact that teacher evaluation reform had in the field of teacher education. Darling-Hammond (2013) posited that, “existing [teacher evaluation] systems rarely help teachers improve or clearly distinguish those who are succeeding from those who are struggling” (p. 24). One of the major issues Darling-

Hammond (2013) had with teacher evaluation reform was no clear connection existed between curriculum in teacher preparation programs and evaluation for in-service teachers. “We will not really improve the quality of the profession if we do not also cultivate an excellent supply of good teachers who are well prepared and committed to career-long learning” (p. 26).

Darling-Hammond (2013) identified six problems with the current evaluation systems: lack of consistent, clear standards of good practice; no focus on improving practice; inadequate time and staff for effective evaluations; little or no consideration of student outcomes; cookie-cutter procedures that don’t consider teacher needs; and detachment of evaluations from professional development. She concluded that, “changing on-the-job evaluation will not, by itself, transform the quality of teaching” (Darling-Hammond, 2013, p. 26). Current education reform targeting teacher evaluation had yet to follow these comprehensive guidelines published by Darling-Hammond, a leading scholar in teacher education. My research initially used Darling-Hammond’s guidelines as a foundation for the study design, data collection, and data analysis.

In another article focused on more effective ways to improve student achievement (other than teacher evaluation reform), Adamson and Darling-Hammond (2012) criticized the effectiveness of the NCLB waiver amendment because of the continued large funding disparity and inequitable distribution of quality teachers. Their research focused on teacher pay, working conditions, and unequal distribution of school resources under NCLB. In Adamson and Darling-Hammond’s study, they also found that, as a result of alternative licensure, teacher preparation programs were diverse and decentralized, resulting in major impacts on teacher effectiveness and student learning. This was just

one more example of other areas the federal government could have focused reform to improve student achievement. Adamson and Darling-Hammond concluded that equalizing resources, specifically ensuring that high quality teachers are in every school (not just the schools that are predominately White and upper-middle class) was necessary to improve student achievement.

In a study by Rowley and Wright (2011), the authors criticized the continued inequalities between black and white students, as well as low-income and middle-income students, despite the enactment of NCLB. Results from their study indicated that the teacher factors that impacted test scores most were teacher-student ratios and teacher perceptions of students. This further implied that there were other ways to address low student achievement that did not involve teacher evaluation reform.

Pedro Noguera, another leading scholar, publicly denounced NCLB as a “train wreck” and further criticized the use of waivers to NCLB by stating that, “10 years of test-based accountability ‘reform’ has delivered no significant progress for students” (Noguera, 2011, para. 2). Noguera (2011) concluded that:

The federal government must embrace a broader, bolder approach to education that includes high-quality early education to narrow large gaps in school readiness, health and nutrition supports to keep children in class and alert, and enriching afterschool and summer activities to build on school-year gains resulting from the work of those great teachers. Anything less will keep us from achieving the educational progress our society so desperately needs. (para. 6)

Diane Ravitch, a prominent educational historian, changed her mind about what she had once perceived as NCLB’s positive impact on student achievement. When NCLB

was conceived, Ravitch supported the amendment and applauded the reform efforts. Ten years later, Ravitch had issue with many aspects of NCLB, including teacher evaluation. In an article about why she changed her mind about NCLB she stated:

The current emphasis on accountability has created a punitive atmosphere in the schools. The Obama administration seems to think that schools will improve if we fire teachers and close schools. They do not recognize that schools are often the anchor of their communities, representing values, traditions and ideals that have persevered across decades. They also fail to recognize that the best predictor of low academic performance is poverty—not bad teachers. (Ravitch, 2010, para. 15)

Sonia Nieto, a leading authority in multicultural education, contested NCLB's definition of what it was to be a Highly Qualified Teacher in her book, *Why We Teach* (2006). Nieto (2006) suggested that by increasing teacher moral, resources, higher respect, and pay, student learning would improve. Furthermore, Nieto (2006) contends that No Child Left Behind is "in fact, leaving many children behind, particularly those that the legislation was suppose to help" (p. 459). NCLB initiated a shift to high-stakes testing that held schools accountable for student learning, but resulted in "limiting the kinds of pedagogical approaches that teachers use, as well as constricting the curriculum, especially in classrooms serving the most educationally disadvantaged students" (Nieto, 2006, p. 460). Nieto's suggestions did not indicate that teacher evaluation was a means to improve student achievement. If it was true that a movement toward high-stakes testing limited pedagogical approaches, it was possible that NCLB had a major influence in the perceived lower performance of teachers. This was the point when teacher education

departments and researchers could have conducted more empirical research focused on various aspects/impacts of teacher evaluation reform to ensure student achievement remained central to reform movements. Empirical research did not suggest that rigorous teacher evaluation systems led to increased student achievement.

Gloria Ladson-Billings, also a leading scholar in teacher education, expressed issues with NCLB in her book, *The Fragmented World of the Social: Essays in Social and Political Philosophy* (Ladson-Billings & Jackson, 2007). One of her main issues was that the policy discouraged African Americans from joining the teaching profession. She stated that, “the proliferation of states and federal education reforms such as the No Child Left Behind Act (NCLB) undermines the goals and aspirations that have historically attracted African Americans into the profession” (Ladson-Billings & Jackson, 2007, p. 48). From a critical multicultural perspective, this was a major issue within the teaching profession. Students learn better when they see themselves in the curriculum. Curriculum was not limited to the textbooks and exams. It transcends all aspects of the schooling process, which included the teacher (Ladson-Billings, 2005). Ladson-Billings provided critical feedback to improve the success of marginalized students; this did not include teacher evaluation reform. Her research suggested that teacher evaluation systems negatively impacted students and teachers of color. My study addressed this through methodology that embodied a critical multiculturalist framework.

Marilyn Cochran-Smith and Susan Lytle (2006) added to the myriad of critiques of NCLB and reform focused on teacher evaluation. Their primary concern was what NCLB assumed and implied about teaching and the role of the teacher through this education reform movement. They applauded NCLB for its comprehensive nature;

however, they stated that, “the act is couched in the unassailable and persuasive language of fairness, equity, and high standards for America’s schools, its teachers, and its children. This language masks some of the most negative consequences of NCLB” (Cochran-Smith & Lytle, 2006, p. 669). Their article revealed that:

NCLB’s conceptions of teachers and teaching are flawed -- linear, remarkably narrow, and based on a technical transmission model of teaching, learning, and teacher training that was rejected more than two decades ago and that is decidedly out of keeping with contemporary understandings of learning. (Cochran-Smith and Lytle, 2006, p. 669)

Cochran-Smith and Lytle (2006) indicated major flaws in NCLB, some of which seemed to be replicated in current teacher evaluation reform. The negative consequences of NCLB were masked and I perceived that to be true with teacher evaluation reform legislation. My research addressed this by focusing on the negative evaluation experiences of teachers.

This study builds on the scholarly responses to the government reform body of research by further exploring diverse perspectives that were missing. The study also provided a space to further explore, through empirical research, issues that prominent education scholars expressed. Interestingly, among all of the critiques of teachers and teaching under NCLB, none of the prominent education scholars alluded to teacher evaluation as the key to student achievement. They cited many pitfalls to NCLB with respect to teachers: the definition of highly qualified teacher, limits to pedagogical approaches, unequal distribution of quality teachers, a broader/bolder approach, and the punitive atmosphere of schools focused on accountability. Building on the scholarly

responses to NCLB, the following section outlines the purpose of my study and its significance within the field of teacher education.

Teacher Quality

The impetus of school reform efforts was to increase student achievement. To this end, the federal government prioritized teacher evaluation as the primary variable impacting student achievement. This section examined the empirical research that focused on student achievement and included the impact of teachers as a variable.

Rowley and Wright (2011) examined the racial inequalities in education, based on a gap in student achievement, by looking at the difference between Black and White students. They chose to conduct a mixed method study using the Educational Longitudinal Study of 2002, by the United States Department of Education, to examine the math and English scores of 8,315 tenth grade students. They used the participants from the longitudinal study in 2002 to administer a survey to the students, as well as teachers, parents, and administrators spanning 752 schools. After the survey, they administered two tests in math and English. Finally, they were able to inspect the school in order to gain information about the physical space. They took a holistic approach to data analysis to examine direct and indirect factors that did impact student success among Black and White students. One of the indirect factors coded was the impact of teachers.

Rowley and Wright (2011) found that most of their variables did not have meaningful effects on student achievement among Black and White students. The factors they did find to have statistical significance were those that differentiated between Black and White students. Black student achievement was shown to be most influenced by the socio-economic status of the family, time spent on homework, stereotype threat, and

peers that had dropped out of school. Among White students, significant factors included learning disabilities and contact with peers that had dropped out of school (but to a lesser degree than Black students). Both groups were impacted heavily by the level of bullying in the school.

Rowley and Wright (2011) concluded the study with suggestions for educational changes to improve student achievement. This conclusion was based on their findings in this study spanning public, catholic, and private schools; and suburban, urban, and rural schools; with an equal distribution between males and females. They suggested allocating resources to after-school tutoring, programs like Advancement Via Individual Determination (AVID) that support students who are the first in their family to go to college, early childhood education programs, and anti-bullying programs.

This extensive study was based on the foundation that teachers play a large role in student achievement; however, they did not determine that teacher evaluation was a primary way to improve student achievement goals. Instead, the teacher-factors they mentioned that impacted student achievement the most were teacher perceptions and expectations of students, and their ability to address cultural diversity.

Rowley and Wright (2011) did a good job of analyzing numerous variables that do impact student achievement through a large participant base, however, analyzing teacher-factors in greater depth may have been more complete. I found that the survey method was only able to provide surface details about the impacts on student achievement. Interviews may have been a better choice in methodology, since most of the factors they analyzed showed no statistical significance. In addition, the teacher perspective was missing in the findings and concluding discussions. The impact and

experience of teachers in this study was a tertiary variable to student-role performance and family.

In a large-scale study conducted by the U.S. Department of Education, researchers used two studies, the Study of State Implementations of Accountability and Teacher Quality Under NCLB (SSI-NCLB) and the National Longitudinal Study of NCLB (NLS-NCLB) and integrated the findings for this study with particular focus on teacher quality (Birman et al., 2009). SSI-NCLB was an, “analysis of school performance data and state documents (including Web sites and consolidated applications and reports), and telephone interviews with state officials” (Birman et al., 2009, p. 9). The other study, NLS-NCLB, utilized survey data from a nationally representative sample of 300 districts, including 1,500 elementary, middle, and high schools. Surveys were given to twelve general education teachers, one special education teacher, and one paraprofessional per school. The teachers who participated in the survey did not indicate if they had a positive or negative experience with teacher evaluation.

Birman et al. (2009) concluded that the definition of and means to obtain the distinction of NCLB’s definition of a “highly qualified teacher” varied dramatically from state to state. Entrance requirements in to the profession of teaching were completely different in each state. The study also found that there were a higher percentage of teachers who did not meet the state “highly qualified teacher” requirements in special education classes, middle schools, high poverty schools, and high-minority schools. This high level of inconsistency contributed largely to low student achievement, especially among minority students (Ladson-Billings & Jackson, 2007; Nieto, 2006).

These findings were troubling on multiple levels. Special education classes, middle school classes, and urban schools were not given adequate resources. Qualified quality teachers were migrating out of urban schools (Birman et al., 2009). In the special education classes, research has shown there was an overpopulation of Black male students, which has also been found to be a component of the pipeline to prison. Students and teachers were systematically routed out of these schools (Noguera, 2003). It seemed that to improve student achievement, and have a socially just system, funds needed to be allocated to ensure there were highly qualified teachers in every class. Reform focused on these findings was not a priority for policymakers, as the current waiver to NCLB program did not require schools to adhere to a standard measurement of teacher quality, or report which teachers were highly qualified teachers.

In addition to special education classes, middle schools were also heavily impacted by a low percentage of highly qualified teachers. In a review of literature, Aughinbaugh (2001) indicated that the impact of early childhood education to address education gaps between students based on ethnicity, such as Head Start, loses its impact once students reach the eighth grade. Based on the findings of Birman et al. (2009) and the review of literature by Aughinbaugh (2001), middle school seemed to be a critical time to ensure that qualified teachers were in the classrooms to help improve student achievement for ALL students. When using empirical research to inform education reform, teacher evaluation did not appear to be the proper reform movement. The imperative appeared to be ensuring funding is in place to have highly qualified teachers in the classrooms at every level of education.

Birman et al. (2009) also found that schools with high poverty and larger populations of minority students were likely to have a greater percentage of teachers not highly qualified. Results showed that of the highly qualified teachers in these high poverty, high minority schools, those teachers were either teaching out of their subject area, or were less experienced than the teachers in more affluent, White schools. These results showed a clear disparity in qualified teacher allocation. This finding was in line with the sentiments from prominent education scholars (Ladson-Billings & Jackson, 2007; Nieto, 2006).

As outlined in this study (Birman et al., 2009), teacher recruiting and retention were major issues impacting student achievement. The federal government's move toward increased teacher accountability, through rigorous teacher evaluation, seemed misaligned with the goal of increasing student achievement. "Between one-third and one-half of districts reported encountering workforce barriers to improving teacher qualifications in 2006-07" (Birman et al., 2009, p. xxvi). Under the recent federal policies, improving workforce barriers, such as quality teacher recruitment and retention, was not a focus. Instead, holding teachers accountable for student achievement based on student standardized test scores was the premise of federal policy reform. As Diane Ravitch (2010) stated, "the current emphasis on accountability has created a punitive atmosphere in the schools" (para. 15). How will this type of school climate appeal to highly qualified teacher candidates and attract them to the field of education?

Birman et al. (2009) concluded their study with four very clear concerns. It is important to note that this study was prepared specifically for the U.S. Department of Education because of the disconnect between research findings and policy

implementation. Birman et al. (2009) state, “If the goal of having an improved teaching workforce and better-served students is to be fully realized, several issues warrant attention” (p. xxxiii). These issues included:

- The variation in policy between states about what they accept as a “highly qualified” teacher.
- The variation in distribution of highly qualified teachers by subject area and school.
- The insufficient communication with teachers about how to become highly qualified.
- The low proportion of teachers participating in professional development and professional development focused on teaching methods and not content knowledge.

This government study came to very clear conclusions and gave clear directives to improve student achievement and teacher quality. However, federal policy has prioritized teacher evaluation as the primary variable to improve student achievement. In Birman et al.’s (2009) report to the Department of Education, they stated: “The potential for the NCLB provisions to effect positive change in the nation’s teaching workforce depends, in part, on addressing these issues” (p. xxxiii). This left me to wondering what political agenda was driving the current reform.

This study builds on the body of research focused on teacher quality by exploring the connection between teacher quality and teacher evaluation, through the perspective of teachers who had negative evaluation experiences. This perspective was missing from the body of research, and the connection between teacher quality and teacher evaluation was not explicated. Additionally, the sequential transformative methodology of this study has not been utilized within the body of research.

Positive Impacts of Teacher Evaluation

The previous studies analyzed teacher quality and the impact on student achievement and did not suggest teacher evaluation was a variable toward improvement. The studies in this section also looked at teacher quality, but found a positive impact from teacher evaluation on student achievement. Prominent education researchers began to examine teacher evaluation under the new federal push to increase teacher evaluation; however, empirical research was limited.

In an article by Adamson and Darling-Hammond (2012), a look at teacher quality through the scope of distribution of funds yielded multiple suggestions for policy makers, including strengthening evaluation for teachers and school leaders. The researchers examined district funding and student achievement. They used both descriptive statistics and ordinary-least squares regression to analyze state data on funding during the 2008-2009 school year. Adamson and Darling-Hammond (2012) looked at where and how money was distributed to schools in New York and California and the impact on student achievement.

Adamson and Darling-Hammond (2012) found that, “achievement has not kept pace among poor and minority students who are, once again, more likely to be taught by less-experienced and less-qualified teachers under less-supportive teaching conditions than their more affluent peers” (p. 35). This supported the findings from Birman et al. (2009) and showed a continued inequality in the education system. From a critical multicultural perspective, working towards equity was necessary and must be at the heart of education reform. Unlike Birman et al. (2009), who focused on improving the disparity

of teacher distribution through systematic approaches to entrance requirements, Adamson and Darling-Hammond (2012) suggested that districts improve school funding by using a weighted student formula allocation, that ensures all schools are funded fairly, and adjusts for cost-of-living differentials. They found that this reform would yield the most benefit to student achievement (Adamson & Darling-Hammond, 2012). They also found that improving preparation and licensing standards enforcement was also necessary simultaneously with funding adjustments. “Investments in more qualified teachers lead to greater achievement gains than other uses of education dollars” (p. 36). The investment suggested was to increase salaries for teachers. Their tertiary finding was that improving teacher evaluation, professional development programs, and mentoring and performance based induction programs could also impact student achievement.

Adamson and Darling-Hammond’s (2012) study provided support for the federal government’s efforts to increase focus on teacher accountability through teacher evaluation. Adamson and Darling-Hammond (2012) did find that improving teacher evaluation could improve student achievement by increasing accountability. However, they found other means to achieve this as well, and teacher evaluation was only briefly noted as a variable to improve learning. Adamson and Darling-Hammond (2012) emphasized that, “federal investments should be tied to each state’s movement toward equitable access to education resources” (p. 37).

Among the positive sentiments about teacher evaluation, was a movement toward Peer Assistance and Review (PAR). PAR was a support and evaluation program that shifted leadership responsibility for teacher evaluation. This was the reform trend and many states were working toward models of evaluation with a similar design. In this

design, teachers were given increased responsibility in mentoring and evaluation, thus alleviating principals from the primary responsibility of evaluating teachers. Research has indicated positive results and attributes of the PAR evaluation program. Goldstein (2004) conducted a study about PAR that was more extensive than any other at the time of publication. In her article, she explained the PAR process and the general sentiments that participating teachers, consulting teachers, and principals had about the newly implemented process (as of 2004). She found that across stakeholders, there was a positive view of the new collaborative process. Goldstein (2004) also found many challenges to implementing a program like PAR. These challenges included: education's hierarchical norms, difficulty in conducting evaluations, district leadership, and program ambiguity (Goldstein, 2004). Through my review of research, I was unable to find an empirical study that addressed these challenges or included multiple perspectives to better understand them.

In a more recent study, conducted by SRI International, the research group found that "PAR programs offer a rigorous and comprehensive way to evaluate teachers" (SRI, 2012, para. 1). Researchers examined two districts in California, Poway Unified and San Juan Unified, where the PAR program had been characterized as exemplary. They conducted site visits and interviews with district and union officials, consulting teachers, principals, and participating teachers. They also analyzed evaluation files, observed classrooms, and attended PAR governance board meetings. One of their main findings was that, despite common assumptions, having the same person in charge of improving teachers also in charge of evaluating them was not an issue. SRI International claimed that PAR fosters, "better collaboration and relations between district and union officials"

(SRI, 2012, para. 6). Unfortunately, the complete picture was not available because the study was conducted in districts where it had positive reviews by all parties. If this study included negative experiences, results may have varied.

Papay and Johnson (2012) added to these positive assessments of PAR. Their research suggested that “PAR encouraged a culture of collaboration not only among teachers but also between labor and management at both the school and district levels” (Papay & Johnson, 2012, p. 723). They also found positive effects on the roles of Principal and consulting teachers. Their final major finding was that despite PAR being an expensive program, it was worth the investment. They conducted 155 interviews with principals, district officials, union officials, consulting teachers, and members of the PAR panel, as well as conducting a document analysis, and site visits. Participating teachers were not part of the participant groups.

Among the research in support for PAR, I found numerous research limitations and implications for future research. Papay and Johnson (2012), designed their study to include participants that represented all of the interested parties, except the teachers who had a negative experience with the evaluation process. Their voice was unheard. They also undermined the entire group of education researchers focused on teacher evaluation and teacher quality. When they reviewed the literature and introduced the study, they cited the support of PAR to include, “educational observers and policy makers, including President Barack Obama, Secretary of Education Arne Duncan, and American Federation of Teachers President Randi Weingarten” (Papay & Johnson, 2012, p. 697). They did not mention the support of educational researchers or teacher educators or address why that group was missing from their review. Supporting the continued political reliance on

outside influence, rather than educational research was doing the field a disservice.

Empirical education research focused on the marginalized voices was needed to fill this research gap. As the research gap continued, policy also continued to be enacted based on personal and political agendas that were not rooted in empirical research.

As seen with the research by SRI (2012), none of the research available about PAR examined or mentioned the possible negative experiences that teachers had when going through the PAR program. Without mention, one could quickly assume that those teachers who were evaluated as unsatisfactory, sent to PAR, and then quit or had a negative experience must have been “bad” teachers. As a stigmatized group, rated as underperforming, they had no voice in policy or research thus far. Until their experiences are heard, through empirical research, we will not fully understand the impacts of new teacher evaluation systems that are so highly revered by the government and district officials. As PAR, and other evaluation programs, grew in popularity, it was necessary to address these gaps in research.

With PAR, I had specific concern about the lack of “checks and balances.” SRI (2012) recently reported no negative impact resulting from having the same person in charge of improving a teacher and evaluating a teacher, but the researchers did not gain perspective from the group that was directly impacted by the lack of “checks and balances.” The participating teachers who were going through the program could provide valuable information about the program. I was also concerned about the possibility of the PAR governance board working together and systematically removing teachers based on biases. My research addressed these issues by interviewing teachers who had negative experiences in the evaluation and PAR programs. Hearing their perspective could provide

a more complete picture, or could reveal shortcomings of the evaluation programs that were quickly coming to fruition. In the following section I review the empirical research that indicates negative aspects of teacher evaluation programs.

Negative Impacts of Teacher Evaluation

The foregoing has shown multiple findings about teacher evaluation and its impact on student achievement. This section includes research that specifically found negative impacts of implementing teacher evaluation to improve student achievement. Although teachers were found to be the main factor in student achievement, researchers have different findings about how teacher evaluation reform actually improves student achievement.

Among the research noting the negative impacts of an increased focus on teacher evaluation, Donaldson (2012) conducted a small-scale study to help inform policymakers in the debate about teacher evaluation. Donaldson (2012) addressed teacher evaluation through a qualitative approach, conducting interviews with 92 teachers and administrators in one district. Half of the schools she chose as study sites reported a positive experience with new teacher evaluation methods according to district surveys and the other half were schools that reported a negative experience.

The main findings of this study were both negative and positive, however, I placed this study in the negative experience section of this paper because the majority of the findings were negative and the study was based on Donaldson's (2012) statement that, "There is scant evidence that evaluation has improved the quality of teachers' classroom instruction or led to the dismissal of underperforming teachers" (p. 1) and that, "we have little systematic evidence regarding how teachers are responding to these

changes and whether their experiences with reform differ by level of teacher performance” (p. 41). Overall, Donaldson (2012) found major deficiencies in teacher evaluation as a means to improve student achievement. Her study focused on the difference of views and experience teachers had with the evaluation depending on how they were rated in their evaluation on a scale of one to five, with one being the lowest performing and five being rated as the highest performing.

Donaldson (2012) found that evaluations did not affect instructional practice but did change how teachers plan and prepare to teach. Teachers reported that evaluation reform was necessary, but expressed views that the system was not fair or objective, and created high anxiety among teachers. Surprisingly, despite the negative results, three-quarters of participants reported that they would recommend the new evaluation system.

In accordance with critical multiculturalism, that left one-quarter of teachers who did not recommend the teacher evaluation process and had become a marginalized group, as compared to the majority who recommended the changes. In line with the recommendation of teachers who supported teacher evaluation reform, teacher evaluation had become a reform movement in nearly every state. It was the minority group of educators that had a non-dominant opinion and who deserved some attention that so far had been neglected. Research focused solely on their experiences could provide more depth to the surface findings of Donaldson (2012), in terms of the teachers who had negative experiences. Donaldson (2012) thoroughly interviewed a large participant base, but findings were not aggregated by positive and negative experiences. Analyzing the data by grouping positive and negative experiences could have revealed more detail about the teacher evaluation process. Of particular interest is Donaldson’s (2012) finding that

teachers reported fear of administration and reported a perceived increase in the ability to fire teachers under the new evaluation process. Donaldson (2012) also mentioned that the evaluation program was perceived as a punishment tool with inconsistencies and varying levels of rigor. Research taking a deeper look in to these findings provides a more complete picture of the impact of teacher evaluation by giving voice to the minority group of teachers who had negative experiences with the teacher evaluation process.

In a timely article published in the Harvard Educational Review, Hill and Grossman (2013) critically revealed inadequacies of the current teacher evaluation transformation. They focused on the design aspect of the teacher evaluation process. Hill and Grossman argued that new observation systems must be subject-specific, include content experts, and provide accurate and useful information for teachers. They concluded by suggesting that states needed systems of evaluation that complemented existing systems.

Unfortunately, with little regard to empirical research, policy was completely changing how teachers were evaluated and how the information was used. Teacher evaluation and teacher education was directly impacted. As policy changes, so does school climate, teacher retention, and student achievement (for better or worse). This study builds on this body of literature by contributing more empirical research for policymakers to be adequately informed through the legislation process.

Research from Private Interest Groups

In a thorough search for literature about teacher evaluation, the search results were dominated by research from private interest groups instead of empirical studies in peer-reviewed journals. Groups such as Bellwether, the Center for American Progress,

and MetLife, have taken notice to policy reform about teacher evaluation by sponsoring their own research on the topic. With privately funded research, research quality and rigor was not clear, because of the lack of peer reviewing, but results and recommendations may still be relevant to understanding teacher evaluation.

Markow, Macia, and Lee (2013) conducted a survey for MetLife. Within their survey they interviewed teachers and principals, as well as other school stakeholders, by telephone. This large-scale survey reached 1,000 U.S. K-12 public school teachers. Their survey revealed that the primary concern for teachers and principals was the declining budget and of much lesser concern was the challenges of evaluating teacher effectiveness. This study showed that teacher evaluation was an issue for teachers and principals; however, it did not suggest reforming evaluation to increase student achievement or improve teacher quality. Within the study, educational issues were identified, but recommendations for improvement were not made. My study addressed this gap by looking more closely at this reported “challenge” of evaluating teachers, through in-depth interviews and data analysis, and aimed to empower participants to be agents of change.

The Center for American Progress has multiple reports about teacher evaluation. Of these reports, Peter Youngs (2013) conducted a study that found teacher evaluation to be a key component for the successful implementation of the Common Core standards and assessments. Youngs (2013) stated that, “past attempts to enact standards-based reform have been impeded by limitations in teacher evaluation” (p. 1). Youngs (2013) further explained ways that teacher evaluation can positively support the enactment of Common Core standards and assessments. This research appeared to support a CCSS

agenda, rather than critically looking at teacher evaluation as an effective process.

Bellwether, another private research group, also sponsored a comprehensive study that looked at the aggressive teacher evaluation legislation (Mead, 2012). Mead (2012) provided a straightforward analysis of teacher evaluation policy changes in each state. Each state was analyzed through a criterion-based assessment about what they did to assess their teachers. This study did a thorough analysis of what was happening with teacher evaluation, state-by-state, but it did not look at the ‘why’ and ‘how’ or ‘impact of’ policy changes. Mead (2012) concluded that states and schools must continue to evolve their evaluation practices as they learn about successes and mistakes. However, to do this, more peer-reviewed empirical research needed to look deeper at the ‘how’ and ‘why’ questions about new teacher evaluation trends in education reform, to better support evolving evaluation practices. This study builds on this body of research by addressing the topic at a deeper level, through its methodology and framework, and builds upon the existing literature to further look at potential mistakes in implementing evaluation systems.

Summary Discussion: Shift in Accountability

An underlying theme among the literature and evolution of evaluation through federal policy pointed towards a shift in accountability. Policy was pushing toward a focus on teacher accountability, rather than school accountability for student achievement (Darling-Hammond, Wise, & Pease, 1986; NCLB, 2002; USDOE, 2013). Teacher evaluation appeared to be the federal government’s primary approach to address low student achievement and teacher evaluation methods and uses were changing rapidly (USDOE, 2013). Unfortunately, as history has shown, “theory and practice of teacher

evaluation diverge” (Darling-Hammond et al., 1983, p. 285). In this same article by Darling-Hammond et al. (1983), *Teacher Evaluation in the Organizational Context: A Review of the Literature*, they wrote about the decade prior to 1986 as a time of teacher evaluation increasing in importance. In their article they “examine how external demands for accountability are at odds with internal organizational needs for stability and trust...and how teacher evaluation may affect organizational operations and teaching work” (p. 285). Based on what Darling-Hammond et al.’s (1986) report, and on conclusions from other leading teacher education and multicultural education scholars about NCLB and the policy impact on teacher evaluation, little has changed in the divergence between theory and practice. This study was important in elucidating a different perspective that could potentially reveal more information about that divergence.

The drastic and competitive nature of teacher evaluation policy reform, surrounding student achievement, had made a very clear shift from school to teacher accountability. Under the 2001 NCLB amendment, “schools are held accountable for results” (USDOE, 2002, p. 9). Schools and districts were directly responsible for the achievement of their students. Under NCLB, districts were required to only employ highly qualified teachers (HQT). By definition, this meant that a HQT have: a bachelor’s degree, a teaching certificate, and proven content knowledge. NCLB was designed to keep schools and districts accountable for teacher quality. Districts were given flexibility to choose where to allocate funds to ensure teacher quality was sufficient. This included “hiring new teachers, increasing teacher pay, and improving teacher training and professional development” (USDOE, 2002, p. 10). This flexibility was designed to help

districts meet their unique needs to raise student achievement.

Under the waivers to NCLB, schools were no longer charged with the responsibility of ensuring teachers were highly qualified and the waiver system has actually changed the NCLB definition of HQT. Districts began implementing teacher evaluation systems that used multiple sources, including value-added sources, rather than focusing on teachers having the attributes previously defined as HQT. With a state waiver, teachers are held more accountable than schools for student achievement. One of the major changes from NCLB to the waivers was to mandate using teacher evaluation to inform personnel decisions. Markley (2004) noted that, “99.8% of public schools use principals’ classroom observation as the primary source of data for teacher evaluation” (p. 4). Previous studies have indicated that principal reviews were often inflated and inaccurately positive and sometimes unfairly negative. Various reasons have been noted, including not disrupting the work environment, lack of content knowledge, and minimal observations (Markley, 2004). With a shift in evaluation purpose and student achievement accountability, I suspect that with the changing conditions, less schools will use the principal-based evaluation system.

Focusing teacher evaluation on accountability could be detrimental to the overall school climate, teacher retention, and administrator-teacher relations. Hill and Grossman (2013) reviewed the waivers to NCLB and found major issue with the reform movement. In their argument against the current design of the waivers in regard to teacher evaluation, they stated that, “these practices have proven unsuitable for generating and sustaining instructional improvement in the past and will fail to do so now if implemented as currently planned” (Hill & Grossman, 2013, p. 372). Their conclusion

was based on the history of policy reform failure. Hill and Grossman (2013) found all the conditions of past reform failure present in the new teacher evaluation reform, as outlined in the waivers to NCLB. Hill and Grossman (2013) suggested a different approach to teacher evaluation that complemented current practices and was based on research on teaching and teacher development. They identified three features that must be in a new teacher evaluation system: 1) evaluators with subject-specific knowledge, 2) subject-specific observation instruments, and 3) feedback must be useful and accurate (Hill & Grossman, 2013).

Hill and Grossman (2013) concluded that, “such an observation system would prioritize teacher learning over accountability and also require numerous elements not contained in existing policy blueprints” (p. 373). It could be useful to have discussions with teachers who had negative experiences throughout the evaluation process to find out if their experiences matched what Hill and Grossman (2013) contended in their article as faults of teacher evaluation.

Summary Implications for the Study: Marginalized Group of Teachers

This review of literature has not only revealed limited empirical research about teacher evaluation, it has also revealed a missing perspective. Research has focused on the positive experiences of teacher evaluation (Goldstein, 2004). Among the research giving voice to teachers about their experiences with teacher evaluation, none have disaggregated the data about positive and negative teacher experiences. Consequently, the negative experience with evaluation was not easily understood because it had not been the focus of prior research. Through a critical multiculturalist framework, this issue was directly addressed in my research by giving voice to this marginalized group of teachers.

The research and participants symbiotically worked together to increase the body of knowledge about teacher evaluation through empirical research and empower participants to be agents of change. This supported the overall sequential transformative research design and the critical multiculturalist theoretical framework. This study builds on current sequential transformative research by continuing to promote social change and empowerment of marginalized groups.

The current reform movement that focuses on teacher evaluation had not yet been connected to teacher preparation. Prominent teacher education scholars claimed this as a major fault of the United States education system and addressed the drastic differences in systems with countries like Finland that outperformed the United States (Ravitch, 2012). Diane Ravitch (2012) noted that Finland did not have strict teacher evaluation and further explained that in the 1960's educators in Finland were mandated to journal everything they taught, every hour of the day. At that time, Finnish reformist pushed for what many called an elitist system, to infuse professionalism in to teaching. This meant that teacher preparation programs made changes in programmatic issues, such as strict pre-service teacher qualifications, rather than firing the worst teachers (Ravitch, 2012). Ravitch (2012) explained that this shift raised the level of teacher quality and student achievement across the board. Finland moved away from teacher evaluation completely. The United States, however, was moving rapidly toward increased teacher evaluation. This was the time that teacher educators and programs in the United States needed to take a critical look at policy and impacts of teacher evaluation on teacher education.

In conclusion, current empirical research has not suggested that teacher evaluation reform results in improved student achievement, teacher instruction or continued

employment, yet it was deemed the key to improvement by the federal government (Donaldson, 2012). This indicated that research in teacher education, focused on teacher evaluation, needed to be strengthened to ensure that student achievement remained central to education reform. With the federal government continuing to make decisions that directly impacted students and teachers, teacher education research needed to focus on the impacts of federal policy decisions, and make suggestions for further reform based on empirical research. From a critical multiculturalist perspective, I aimed to elucidate the negative experiences teachers had with teacher evaluation.

This chapter has explained critical reports and policies that have led to the current state of teacher evaluation and what prominent education scholars thought about the reform and policies. No Child Left Behind (2001) served as the catalyst to aggressive reform addressing student achievement (Darling-Hammond et al., 1983). Reports that instigated NCLB included The National Commission on Teaching and America's Future (1996), The Goals 2000: Educate America Act (1994), and A Nation at Risk (1983). After NCLB, further legislation addressing low student achievement included waivers to NCLB (USDOE, 2012, 2013), Race to the Top (McGuinn, 2012; USDOE, 2009), and the Teacher Incentive Fund (USDOE, 2012). Prominent education scholars adamantly disagreed with such aggressive pressure toward increased teacher evaluation to address student achievement. Prime examples, such as the high student achievement and nearly non-existent teacher evaluation protocol in Finland, and other effective means of improving student achievement, were cited as more logical approaches to education reform (Cochran-Smith & Lytle, 2006; Darling-Hammond, 2013; Hill & Grossman, 2013; Ladson-Billings, 2005; Nieto, 2006; Noguera, 2003; Ravitch, 2010).

The chapter has also reviewed related empirical research and research from private interest groups about teacher evaluation. Four themes emerged from the review of literature: evaluation as positive, evaluation as negative, focus on teacher quality, and research from private interest groups. From this review, it was apparent that very limited body of knowledge about current teacher evaluation trends and pressure for reform exist within the context of the United States. The chapter concluded with an explanation of the impacts on teacher education. The next chapter will explain the specific methodology of this study. Chapter four will present the findings and chapter five will explain the interpretations and implications of the findings.

Chapter 3: Methodology

The previous chapter provided a review of literature about teacher evaluation. Chapter three builds on the literature foundation by providing details about the methodology of the study. National reports touting an education crisis in the United States prompted education reform movements aimed to improve student achievement. Federal policy was focused on teacher evaluation as the primary means to improve student achievement. States were adhering and revising practices of evaluation with aggressive teacher evaluation policy reform.

However, as chapters one and two revealed, limited empirical research existed that supported such a primary focus on teacher evaluation as a means of improving student achievement. Also, among the existing research, the perspective was limited, leaving a group of marginalized teachers. Teachers who had negative experiences within teacher evaluation remained voiceless in the existing research. Numerous teacher education scholars have expressed contention regarding current trends in education reform that link teacher evaluation to student achievement (Adamson & Darling-Hammond, 2012; Cochran-Smith & Lytle, 2006; Nieto, 2006; Noguera, 2011; Ravitch, 2010; Rowley & Wright, 2011). However, policy continued to push toward reforming teacher evaluation systems, such as PAR (NCTQ, 2012).

Restatement of Problem

To date, minimal empirical research exists to inform and support policy decisions about teacher evaluation reform (Darling-Hammond, 2013). Current reform decisions had a major impact on schools, teachers, and students. Policies and grants, such as waivers to NCLB and Race to the Top, were implicating teacher evaluation as the primary means to

improve student achievement and the working condition was changing for teachers (NCTQ, 2012). It is necessary to increase the body of research representing multiple perspectives involved in teacher evaluation policy so policy makers are informed and teacher preparation programs could adequately prepare preservice teachers for their future working conditions.

Restatement of Purpose

The purpose of this study was two-fold. The first was to elucidate the negative experiences that teachers had with evaluation. These experiences were missing from the research and could reveal a misuse of teacher evaluation that ultimately impacts student learning. I aimed to advocate and empower a marginalized group of teachers who had negative experiences with teacher evaluation. Through a sequential transformative design, the study was grounded in a critical multiculturalist framework that sought to advocate for and empower a group of teachers who had negative experiences through the PAR teacher evaluation process in their school district.

The second purpose was to juxtapose the intended use of current popular teacher evaluation reform movements to the evident implementation. I aimed to reveal dissonance and/or congruence between how the Peer Assistance and Review evaluation policy was designed and how it was being used in school districts, specifically focused on the negative experiences. My study's focus on juxtaposing the intended use of evaluation policy to evident use was based on a clear need to represent this voiceless perspective, as it has not been included in research or policy reform to this point.

Review of Research

This study utilized the theoretical framework and methodology to empower

participants through their negative teacher evaluation experiences. The three aspects of the study, theoretical framework, methodology, and the topic of teacher evaluation, came together to work toward empowering a marginalized group of teachers. Figure one shows the interrelated nature of these aspects of the study and illustrates the use of the theoretical framework and methodology to support the need of empowerment within the topic of teacher evaluation.

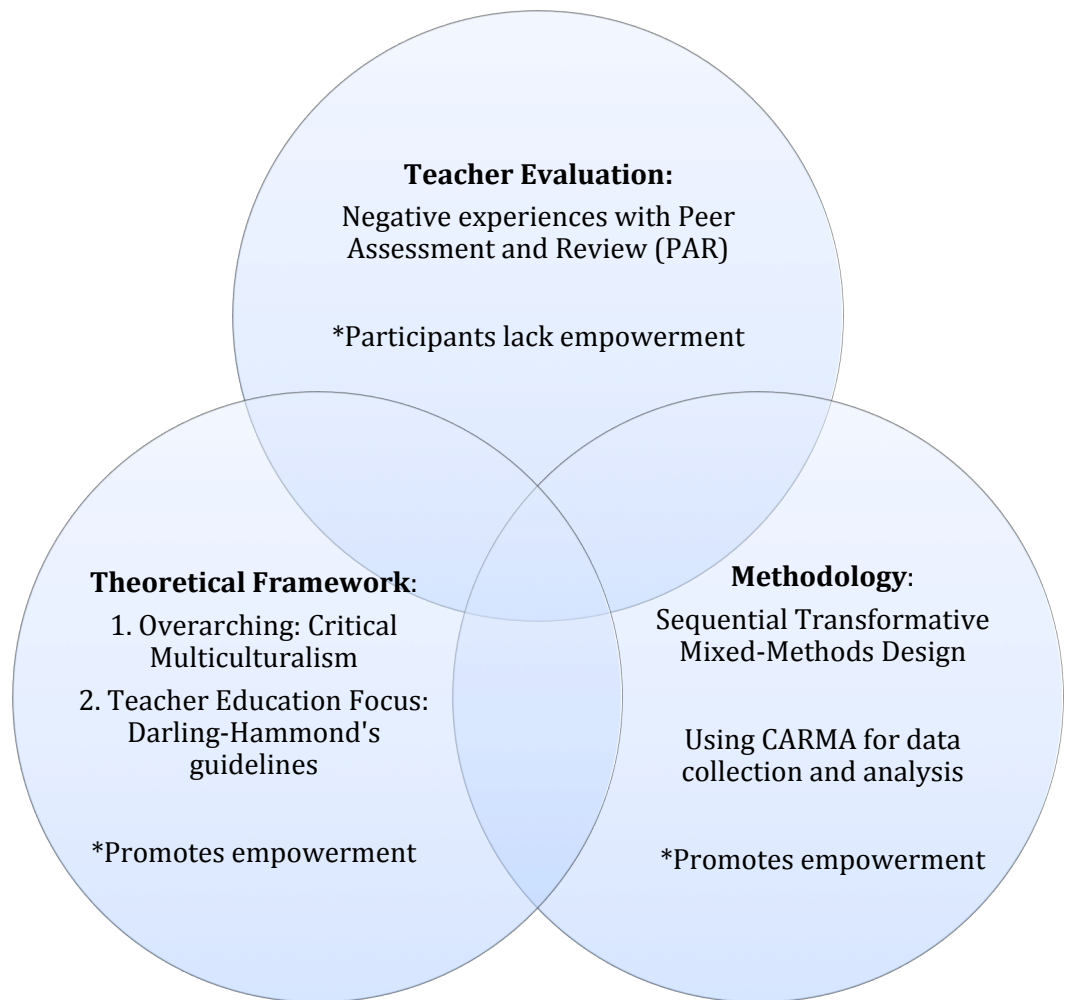


Figure 1. *Theoretical framework, methodology, and the topic of teacher evaluation.*

The theoretical framework and methodology were both designed to promote empowerment, while the topic of teacher evaluation was focused on a subsection of

teachers who had negative experiences in teacher evaluation and lacked empowerment.

The following table (Table 1) provides an overview of the research process in terms of how the research questions align with the data sources, data analysis, and timeline of this study.

Table 1

Review of the Research Process

| Research Questions | Data Sources | Data Analysis | Timeline |
|---|---|--|--|
| How does the PAR teacher evaluation process negatively impact teachers? | Interviews Teacher survey District data | Content Analysis CARMA Darling-Hammond | District data & survey analysis: March 1, 2014 to April 1, 2014 Interviews: April 1, 2014 to June 30, 2014 Interpretation: May 1, 2014 to March 1, 2015 |
| What, if any, parallel traits exist among those teachers who had negative experiences with the PAR evaluation system? | Interviews Teacher survey District data | Content Analysis CARMA Darling-Hammond | District data & survey analysis: March 1, 2014 to April 1, 2014 Interviews: April 1, 2014 to June 30, 2014 Interpretation: May 1, 2014 to March 1, 2015 |
| How does the intended use of the PAR teacher evaluation process compare to the evident use of PAR? | PAR policy Interviews Teacher survey District data | Content Analysis CARMA Darling-Hammond | District data, PAR policy, & survey analysis: March 1, 2014 to April 1, 2014 Interviews: April 1, 2014 to June 30, 2014 Interpretation: May 1, 2014 to March 1, 2015 |

Table one outlines the research process and will be further explained in this chapter. The three research questions were directly connected to the five different data sources. The

process of data analysis connected to each data source, as well. Content analysis was conducted with existing data, including: district data, teacher survey, and PAR policy. CARMA and Darling-Hammond's (2013) guidelines were used with all data sources, including pre-existing data and interviews. Table one was provided as an overview of the research process and will be described in detail throughout this chapter.

Research Questions

The study elucidated marginalized perspectives to empower participants, to better inform policymakers and teacher preparation programs, as well as increasing the body of empirical research about teacher evaluation. There were three guiding questions to the study:

1. How does the PAR teacher evaluation process negatively impact teachers?
2. What, if any, parallel traits exist among those teachers who had negative experiences with the PAR evaluation system?
3. How does the intended use of the PAR teacher evaluation process compare to the evident use of PAR?

Sequential Transformative Design

This study utilized a mixed methods sequential transformative design that aligned with the theoretical framework with the aim to provoke change. The study consisted of two sequential data collection phases; quantitative analysis of existing survey, PAR policy, and district data that led in to phase two with qualitative interviews (Figure 2). As Creswell (2009) stated, there are many options to choose from with an advanced mixed methods design. Sequential transformative design was a specific combination of elements that was closest in alignment to this study. Using the two phases of the sequential

transformative design for this study made analysis of multiple data sources possible and supported the use of quantitative and qualitative analysis.

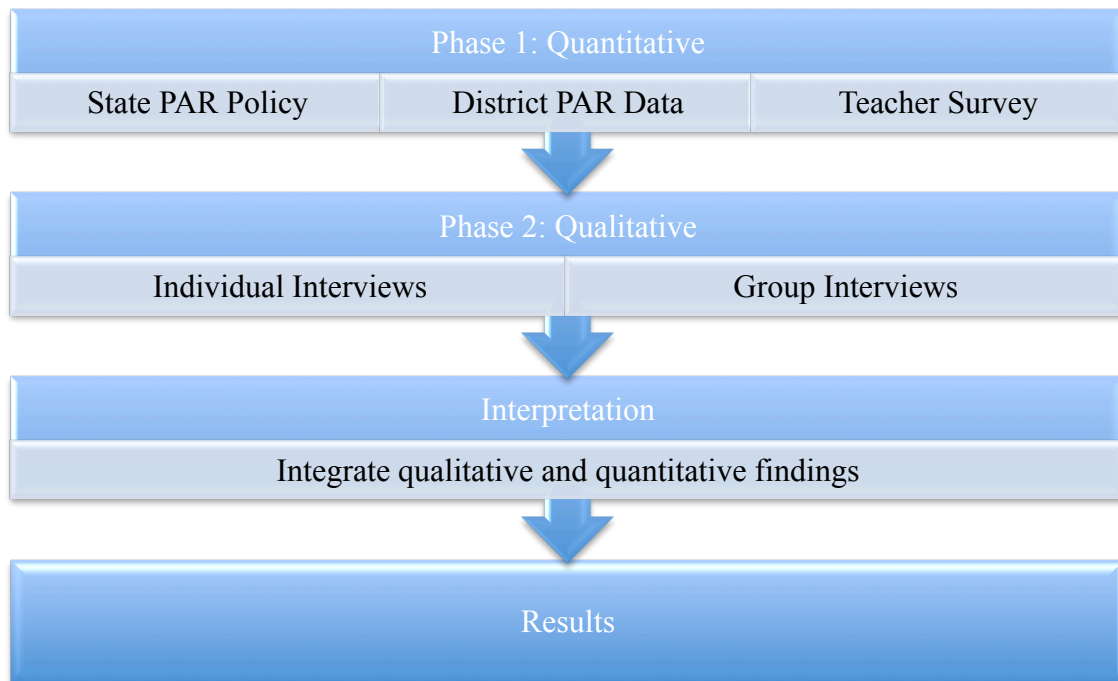


Figure 2. Phases of a sequential transformative design.

I chose a sequential design because I wanted to utilize the data from the district and survey to inform the interviews. Themes revealed themselves through phase one and were addressed at a deeper level in the interviews. With limited existing empirical research about teacher evaluation, this was an appropriate way to approach the data.

I chose a transformative study design because of the nature of the research and how closely it aligned with the theoretical framework. Transformative designs are rooted in the theoretical lens of the study (Creswell, 2009). For this study, I used a critical multicultural framework that aimed to give voice to a marginalized group of teachers. I also have personal experience with the misuse of evaluation systems. When using a transformative design, “no matter what the domain of inquiry, the ultimate goal of the

study is to advocate for change” (Creswell, 2009, p. 222). Issues of inequity, oppression, domination, and alienation were also core elements of a transformative design and were core elements of this study, as well as my personal experiences with teacher evaluation. Therefore, even though other methodologies would have worked for this study, a sequential transformative design was the most appropriate methodology for this study about the negative experiences of teacher evaluation. Sequential transformative design, rather than other methodological designs, not only facilitated data collection and analysis that was in line with the research questions, but the aim of this methodology was to give voice to marginalized groups. A sequential transformative design worked to advocate for and empower participants. In both phases of the study, a collaborative nature of data collection and analysis was present. Sequential transformative design more freely allows for this type of collection and analysis. Participants saw the data and interpretations. Their feedback and thoughts were utilized to advocate for change on a personal level through empowerment, as well as on a larger scope of state and federal policy design about teacher evaluation.

Phase One

The first phase of the study was a quantitative analysis of a previously published survey that was created by the teachers in the district where this study took place, during the 2012-2013 school year. Phase one also included analysis of district data about who had been referred to Peer Assistance and Review. Data included multiple demographic indicators for each person referred to PAR over the past ten years. PAR was the secondary evaluation step after a teacher had been rated as unsatisfactory on his/her evaluation by a school administrator. After the teacher had participated in PAR for one

school year, the panel of ‘peers’ determined if the employee would be retained or dismissed. The district where the participants were from in the study was one of the first districts in the nation to implement PAR as the evaluation system for teachers and had been using PAR for over ten years. Federal mandates encouraged more schools to adopt a PAR system for their teacher evaluations.

Phase Two

The second phase in the sequence of the study was a qualitative design that included individual interviews of five teachers who self-reported that they had negative experiences with the teacher evaluation process, specifically Peer Assistance and Review. Following the individual interviews, one group interview was conducted to elucidate themes that emerged in the individual interviews. The group interview focused more on policy design and implementation similarities and differences. After the two phases, data were integrated for collective interpretation.

Sampling and Participants

Initially purposeful unique sampling targeted teachers who had negative experiences with teacher evaluation. Due to a lack of existing research from the negative perspective, their experience was considered unique for this study. After a few key participants were selected, snowball sampling was utilized within the same school district in Northern California where PAR had been used for over ten years (Merriam, 2009). The participants referred other teachers they knew who had negative experiences with teacher evaluation. Participants from this school district who had negative evaluation experiences were chosen because they had previously been recognized for their outstanding teaching, through positive evaluations, awards, and honors, and they taught within the school

district that had an established use of PAR for evaluating teachers. Other studies focused on school districts that had newly adopted PAR and where PAR was reported as a successful reform to teacher evaluation. The school district in Northern California, having a long standing use of PAR, and a group of teachers who had negative experiences, was an ideal choice for this study. The unique participant perspective had yet to be represented in previous empirical research.

The ethical considerations for this study were centered on privacy of the participants and avoiding additional negative repercussion because of self-advocacy and expressing negative experiences. To address these issues, the school district in Northern California remained anonymous and the teachers who participated were also kept anonymous. Additionally, the district name was omitted from all data-sets used for the study. All participants were adults and signed an informed consent to voluntarily participate in the study. Participants were made aware of their right to stop participation at any point during the research.

Data Collection

Through a sequential transformative design, five data sources were utilized to elucidate the negative experience the participating teachers had with teacher evaluation. Data sources were specifically focused on the continued use of the Peer Assessment and Review evaluation system. All participants were from a school district in Northern California that had been using PAR for over ten years.

A sequential transformative study design separated data collection in two distinct phases. One phase focused on qualitative data and the other on quantitative. The order and weight of the qualitative and quantitative methodology is flexible (Creswell, 2009).

For this study, phase one was a quantitative analysis of district PAR data, district PAR policy, and an anonymous teacher survey from the 2012-2013 school year. Phase two was qualitative, with individual and group interviews.

Following the distinct phases of the sequential transformative design, data from each phase was integrated for data analysis as a collective interpretation with participants. This collective interpretation was done through email correspondence. After interpretation was complete, results were formulated. The Complementary Analysis Research Matrix Application (CARMA) tool was used for data collection and analysis (Putney et al., 2006).

Data Collection using CARMA

CARMA is a four-stage matrix used as a tool to systematically collect, organize, and analyze data by juxtaposing the intended policy use and the evident policy use (Appendix A). For this study, teacher evaluation was the focus. Data collection took place during the first two stages and data analysis took place in the last two stages.

The first stage of CARMA consisted of note taking about the evaluation program policy, who was being served, who was involved, how teachers were to be evaluated, and what would be the intended results of the evaluation system. Detailed data were collected about the intended use and outcomes of teacher evaluation, specifically the Peer Assistance and Review evaluation system.

The second stage of CARMA consisted of note taking about the participants' experiences with the teacher evaluation system. Detailed data were collected about the participants and who they were, how the participants had experienced teacher evaluation, and the end results of their experiences with teacher evaluation. This stage focused on

hearing from the participants directly and the sharing of negative experiences.

The CARMA tool was used throughout the sequential transformative design. Phase one of the sequential transformative design aligned with stage one of CARMA. Phase two aligned with stage two of CARMA. The interpretation and results components aligned with phase three and four of CARMA. The results phase of CARMA was a straightforward report of congruencies and divergence between expected and evident use of teacher evaluation. This was the integration of data from phase one and two of the sequential transformative design. The final stage of CARMA was the interpretation component, where participants and I co-construct meaning and implications of the data and findings from phases one, two, and three of CARMA. This aligned with both the interpretation and results portions of the sequential transformative design.

Data Sources

This study examined five data sources (see Figure 3). These sources included: a teacher survey, district data, California Peer Assistance and Review (PAR) Policy, individual interviews, and one group interview. Data collection happened sequentially, from quantitative to qualitative, and then interpreted together before compiling results.

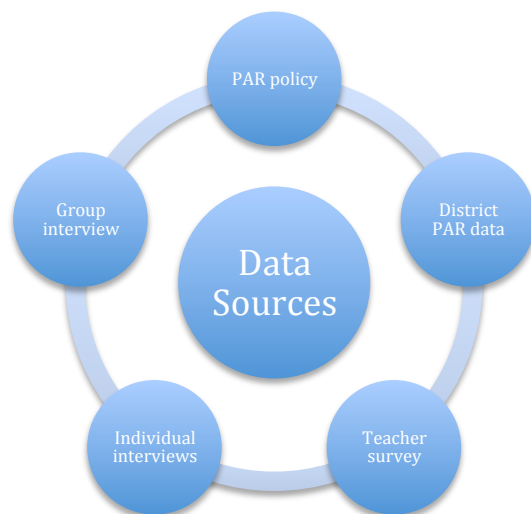


Figure 3. Data sources.

Teacher survey. During the first phase of this study, an existing anonymous survey was analyzed. A teacher in the school district in Northern California designed a survey about the evaluations and working relationships at a high school within the district during the 2012-2013 school year. There were a total of nine questions on the survey and 94 respondents (Appendix B). Expected survey results revealed themes of feelings and experiences that teachers at the school district had with their evaluations. Basic statistical measures, such as mean, average, and mode, were utilized to understand the data. Phase one took place from March 1, 2014 to April 1, 2014.

District data. In addition to the survey, district data about the PAR evaluation system and overall district demographics were analyzed in the quantitative phase one of this study. District data included the age, ethnicity, gender, and salary placement of the 41 teachers that had been referred to PAR in the school district in Northern California from 2002 to 20013 (Appendix C). A binomial distribution analysis was conducted.

California Peer Assistance and Review (PAR) policy. Additionally, during phase one, the California Department of Education PAR policy that the school district must adhere to, was analyzed using the CARMA data analysis tool. The CARMA tool was a matrix to help systematically collect, organize, and analyze data. The specific PAR policy outlined by the California Department of Education was juxtaposed to how PAR manifested itself in the school district. The manifestation was more closely analyzed during interpretations and results. The PAR policy, as written by the state of California, was analyzed in phase one, outlining the intent and design of the policy. The California PAR policy also transcended to phase two through CARMA because of the design versus implementation juxtaposition that CARMA addressed. Interviews in phase two referred to policy and illuminated how participants had experienced the policy. Phase one took place from March 1, 2014 to April 1, 2014 and phase two followed, from April 1, 2014 to June 30, 2014.

Interviews. Transformative interviews were conducted during the qualitative phase two between April 1, 2014 and June 30, 2014. By design, transformative interviews, which align with critical multiculturalism and sequential transformative study design, aimed to elucidate issues of power, privilege, and oppression (Merriam, 2009). Transformative interviews also aimed to empower participants (Merriam, 2009). Using this interview design aligned with the theoretical framework and the methodology. For this study, four teachers in the school district in Northern California were interviewed based on their negative experience with evaluation and history of excelling as teachers. The four teachers were interviewed individually and then interviewed as a group after all individual interviews were completed. The individual interviews were semi-structured, to

ensure that the participants defined their negative experiences as authentically as possible. A flexible question guide was used to ensure issues were addressed, but emerging views and ideas from the respondent ultimately guided the interview (Merriam, 2009). Refer to Appendix D for individual interview questions. Probing questions were asked as needed to elucidate more detail about the topic. Some interviews were conducted through the internet and others were conducted in person, depending on the participant's availability. I conducted the interviews in person in Northern California during the month of June, 2014.

One focus-group interview was conducted after all four individual interviews were conducted. The focus group interview was designed to empower the participants toward change. Merriam (2009) suggested at least six participants in focus group interviews; however participation was difficult to sustain. A group interview intended to help participants express their experiences in more robust ways than individually to possibly reduce feelings of isolation (Merriam, 2009). This aligned with the methodology and theoretical framework, as both aimed to give voice to marginalized groups and empower participants. The whole-group interview was intended to be conducted in person but defaulted to an online interview because of scheduling. The online interview consisted of group emails.

Data Analysis

After phase one and two, collaborative interpretation took place, where researcher and participants co-construct meaning of the data. Data were coded by emerging themes and by pre-determined themes based on the guidelines of Darling-Hammond (2013). This thematic coding was composed in to initial findings that were emailed to participants.

Each participant had the opportunity to individually or collectively respond to interpretation with affirmations or corrections. Data from phase one and two was integrated and collectively interpreted to formulate the results through triangulation of data. Triangulation assisted in increasing consistency and credibility (Merriam, 2009). Data analysis was done with both theoretical frameworks as guides through the Complementary Analysis Research Matrix Application (CARMA) tool.

Data Analysis using CARMA

As previously mentioned in this chapter, the first two stages of the CARMA matrix were data collection. The data analysis stages were stages three and four (Appendix A). In stage three of CARMA, results were compiled through what Putney et al. (2006) call note making (rather than note taking). The data collected from stage one and two was juxtaposed and systematically analyzed. For this study, the results were co-constructed with participants. This phase focused on comparing and contrasting the expected with the evident data, and looking at the different aspects of congruency or divergence of the data collected in phase one and two.

The final phase of CARMA was the conclusions and recommendations. This was what Putney et al. (2006) call note remaking. This is where the participants and I interpreted the data to address implications and modifications that were necessary to work toward change their negative experience with evaluation to positive experiences.

Category Construction

Following phase one and two of the sequential transformative design, open coding took place. Within CARMA, this took place in the note making and note remaking stages previously described. During open coding, analysis was very flexible and open to any

emerging themes to inductively form categories. After open coding, I grouped open codes through analytic coding by interpreting and reflection on the meaning of the data sets. Analytic coding, rather than axial coding, was chosen because it best fits the design of this research to co-construct meaning with participants and work toward empowerment. Axial coding simply focused on description (Merriam, 2009). Each data set was analyzed individually and then coding was merged to capture recurring patterns across all quantitative and qualitative data (Merriam, 2009).

Timeline

The study took place in phases to uphold the sequential transformative design. The quantitative phase one took place between March 1, 2014 and April 1, 2014. The qualitative phase two took place between April 1, 2014 and June 30, 2014. Interpretation and results of the data from both phases took place between May 1, 2014 and March 1, 2015.

Role of the Researcher

In this study, I, as the sole researcher, found a gap in the research and designed this study to elucidate the negative experiences of teacher evaluation. I implemented all aspects of the research. However, my role in this study was more closely aligned as a facilitator, as I collaboratively co-constructed the meaning of the data with the participants. This role of co-creating meaning was specifically chosen with the aim to empower participants and ensure that their voice was appropriately expressed. This role aligned closely with the methodology and theoretical framework,

As a high quality teacher who had negative evaluation experiences, I have an understanding about the feelings that go along with the misuse of teacher evaluation.

With this first-hand experience, I was able to conduct authentic interviews with participants who knew that I appreciated their perspective and the situation, without judgment. The participants may have been more open with me because of my experiences and the facilitator role that I had during data collection and analysis.

Trustworthiness

The mixed methods sequential transformative design was conducted in two phases: quantitative leading to qualitative (Creswell, 2009). In this design, I first collected and analyzed the preexisting teacher survey from the 2012-2013 school year and the district PAR data. The qualitative data were collected in the second phase of the study through in-depth individual interviews, followed by one group interview. The qualitative phase built on the quantitative phase and then the two phases were connected in the interpretation stage of the study. The rationale for this approach was that it was a methodology that not only facilitated data collection and analysis that was aligned with the research questions, but the aim of this methodology was to give voice to marginalized groups to better advocate for and empower participants. This also provided triangulation of data, which was a key component of trustworthiness.

Triangulation increased credibility, also known as validity and reliability. This study utilized multiple facets of triangulation. This study utilized multiple methods (quantitative and qualitative), multiple data sources (district data, teacher survey, policy, individual interviews, and a group interview), and multiple theories as the guiding framework (critical multiculturalism and Darling-Hammond's 2013 guidelines). Merriam (2009) expressed the inability to capture an objective truth or reality, but through strategies like robust triangulation, researchers can increase credibility.

Additionally, to support internal validity or credibility, also known as respondent validation, this study used ‘member checks’. Member checks aimed to rule out misinterpretation and identify any research bias or misunderstanding (Merriam, 2009). In both phases of the study, a collaborative nature of data collection and analysis was present to form continual member checks (Merriam, 2009).

To support consistency, also known as reliability, I kept a detailed audit trail of all data collection and analysis. This consisted of a record of how data were collected, the process of conducting the research, and a running record of interactions with the data. During analysis, I kept account of how themes were derived. The detailed journal also included reflections, questions, problems, and decisions.

The study took place in a large school district in Northern California. This district was chosen because of its history of using the PAR evaluation system for the past ten years and the ease of imploring participants in the study. Previous research has focused on school districts where PAR has been successful and results have been discussed holistically, not accounting for the unique knowledge and understanding of those teachers who had a negative experience with teacher evaluation. This study provided thick, rich description about the participants’ experiences, which supported transferability. However, the district remained anonymous and was referred to as a school district in Northern California.

Additionally, all data were protected through sole storage on my personal computer and backed up through cloud storage. I am the sole researcher and I am the only person that had access to the data files. All participants and district information was kept anonymous. This level of protection was deemed as appropriate and adequate

(Merriam, 2009).

Chapter Summary and Transition

To conclude, it was evident that more research needed to be conducted about new trends in teacher evaluation and specifically, research that represented multiple perspectives. Through a critical multiculturalist theoretical framework and by using Darling-Hammond's (2013) guidelines to improving teacher evaluation, this sequential transformative study elucidated the negative evaluation experiences of teachers.

The study utilized a mixed methods design that analyzed five data sources: PAR policy, district PAR data, a teacher survey, individual interviews, and a group interview. In phase one, quantitative data collection was done on the PAR policy, district PAR data, and a teacher survey. In phase two, qualitative data collection was done with four individual interviews and one group interview. Following, a co-construction of analysis and results of the data from phase one and two was done with participants to advocate for change and empowerment.

CARMA, a data collection and analysis tool, was used throughout the study to organize data collection and interpret the findings during data analysis. CARMA juxtaposed the intended use of policies or programs with the evident usage. By comparing and contrasting the two, we better understand issues surrounding negative experiences of teacher evaluation to enable policymakers and teacher preparation programs to make informed programmatic decisions.

CHAPTER 4: Findings of the Study

Past research and policy implementation make it evident that a discrepancy between education policy and education research exists in the area of teacher evaluation. Research that represents multiple perspectives, specifically from teachers who had negative experiences is missing or inadequately addressed (Goldstein, 2004; SRI, 2012). Through a critical multiculturalist theoretical framework, and by using Darling-Hammond's (2013) guidelines for improving teacher evaluation, this sequential transformative study elucidates the negative evaluation experiences of teachers. With a sequential transformative design, "no matter what the domain of inquiry, the ultimate goal of the study is to advocate for change" (Creswell, 2009, p. 222).

This study's methodology was a mixed methods design that analyzed five data sources: Peer Assistance and Review (PAR) policy, district PAR data, a teacher survey, individual interviews, and a group interview. A data collection and analysis tool called CARMA was used to guide the research (Putney et al., 2006). CARMA analysis allowed for juxtaposition of intended policy use of current teacher evaluation trends with evident use of policy in the schools. Data collection and analysis were conducted in two sequential phases, quantitative to qualitative, and then the transformative component was a co-construction of data analysis and results to advocate for change and empowerment. The purpose of this study was to better understand issues surrounding negative experiences of teacher evaluation to enable policymakers and teacher preparation programs to make informed programmatic decisions about teacher evaluation.

Chapter one introduced the reform movement focused on teacher evaluation and provided the rationale of this study. Chapter one explained the significance of exploring a

broader perspective of teacher evaluation reform policy and the implications it has on teaching and learning. Chapter two reviewed the existing literature and found a paucity of research focused on the negative impacts of PAR. Chapter three focused on methodology by explaining the sequential transformative design, critical multicultural theoretical framework, and the data collection and analysis tool called CARMA.

This chapter reports the findings from phase one and two of the sequential transformative method. Phase one consisted of quantitative data analysis and phase two was qualitative data analysis. However, the qualitative data were given greater weight than the quantitative data for this study. This chapter introduces the emergent themes, presents each data source separately, and then further applies the four major themes to the datasets holistically. The main sections of the chapter include: introduction of emergent themes, alignment to research questions, phase one, phase two, major themes, discrimination suggested through CARMA, making connections, collaboration, and conclusion. Part one, two, and three of the CARMA assessment matrix, which guided research and analysis, are discussed in this chapter as well. The interpretation of results, implications for teacher education, and further research will be discussed in chapter five.

Introduction of Emergent Themes

As the data were collected and analyzed, four major themes emerged (see Figure 4). The first major theme was ‘The Scarlet P’. Within this thematic organization of the data, the use of the letter ‘P’ referred to PAR. This theme was in reference to the classic American novel by Nathaniel Hawthorne, *The Scarlet Letter* (Hawthorne, 1850). Within this novel the main character, Hester Prynne, was charged with the sin of adultery and made to wear the letter ‘A’. The letter ‘A’ was the physical manifestation of her sin, as

defined by the societal norms of that time. Within teacher evaluation trends, such as PAR, conceptions of teaching and learning emphasized testing and standardization. Teachers were branded as a result of this peer evaluation process if their methods were not aligned with the norms, somewhat like Hester Prynne was in the novel. Issues such as sin, isolation, shame and nonconformance were thematic in the novel, as they were in the findings of this research. Within education, these issues included the bad-teacher narrative and education movements such as Common Core State Standards that required teacher compliance. The novel ends, as does this research, with some form of redemption. Redemption emerged in this research when some participants filed lawsuits against their school district and others were empowered to speak out about the negative impacts of teacher evaluation after hearing about this study. These unexpected results are discussed in chapter five.

The second theme that emerged was the Kangaroo Court. Merriam-Webster (2014) defines the term Kangaroo Court as, “a court that uses unfair methods or is not a proper court of law” (para. 1). Multiple data sources from this study suggested that PAR was a mock court operating irresponsibly and placing judgments outside of the legal procedure. Some data suggested that decisions surrounding who went to PAR and who got out of PAR were predetermined regardless of work done.

The third theme that emerged was The Salem Witch Trials, which historically were characterized by false accusations, lapses in due process, mass hysteria, and isolationism. Multiple data sources suggested these same tendencies within teacher evaluation processes. Additionally, current court case decisions, such as *Vergara v. The State of California* (Education Counsel, 2014), indicated lapses in due process through

litigation funded by the wealthiest people and organizations. Mass hysteria was a theme illustrated by the media, reports, federal policies, and competitive grants that were designed without the empirical research to support the shift in accountability from school to teachers. Interview data suggested that false accusations and isolationism existed among those who were referred to PAR.

The final theme that emerged was Nonconfirming Data. These data did not support the existence of negative impacts of evaluation. Although I specifically looked at the negative impacts of evaluation, because this perspective was missing from previous research, some data revealed a positive impact of teacher evaluation and specifically a positive impact of PAR. Each theme is discussed in detail later in this chapter, after the findings from the five data sources are explained.

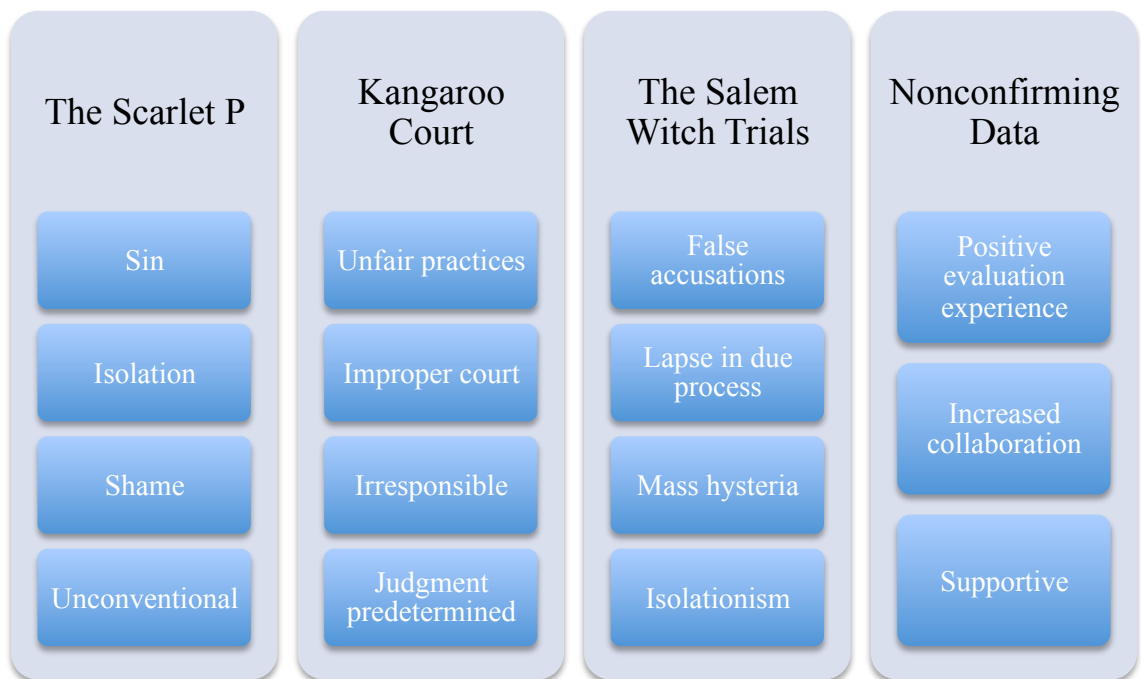


Figure 4. Emergent themes and their major attributes.

These themes were closely aligned with and embedded within the critical

multiculturalist theoretical framework. The concepts of The Scarlet P, The Salem Witch Trials, and Kangaroo Court refer to incidents and situations where critical multiculturalism takes issue. These historical references, chosen as themes, represent major components of the theoretical framework, such as unjust systems, unequal and unfair treatment of women and minorities, and empowering marginalized groups. The unique structure of the education system, historically dominated by women and characterized by minimal respect and pay, insidiously appears to have led in to the negative impacts of teacher evaluation.

Alignment to Research Questions

This study was guided by three research questions. Each question was addressed in phase one, phase two, and the reflection and results stages. Embedding these questions kept the research focused on understanding the negative experiences of teacher evaluation and working toward social justice. The CARMA analysis tool organized the data collection and analysis. The guiding research questions were as follows:

- How does the PAR teacher evaluation process negatively impact teachers?
- What, if any, parallel traits exist among those teachers who had negative experiences with the PAR evaluation system?
- How does the intended use of the PAR teacher evaluation process compare to the evident use of PAR?

Phase One

Phase one of this study was a quantitative analysis of existing data and policy. Three data sources were used for this phase: a teacher survey, district data, and the California Peer Assistance Review (PAR) policy. The data sources were analyzed using

the CARMA matrix through note taking. Analytic coding was used to co-construct meaning with participants and work toward empowerment. Each data set was analyzed individually and then coded concurrently to capture recurring patterns across all quantitative and qualitative data (Merriam, 2009).

CARMA Matrix for Phase One

The CARMA matrix was divided into four section or columns (Putney et al., 2006). The first three columns of the matrix are discussed in this chapter. The fourth column, which makes recommendations to maintain or modify programs, is discussed in chapter five. A CARMA worksheet was used for each column of the CARMA matrix (see Appendix G for completed CARMA worksheets). The first column provided orienting information that explained what was expected from a program or policy. In this case, the expected outcomes or intentions of PAR were noted by examining the California education code pertaining to the PAR evaluation system. The second column provided information about evident implementation of the program. This was where the evident use of the PAR policy was noted. The existing district data about who has been referred to PAR over the past ten years and the existing teacher survey were evaluated to explore the evident use of the PAR policy. Evident use was also explored in the second phase of the study when participants were interviewed individually and as a group. The third column of CARMA was, “the point where NoteTaking changes to NoteMaking, and the interpretation of the data is made” (Putney et al., 2006, p. 2). This column juxtaposed the intended use of PAR, noted in column one, with the evident use of PAR, noted in column two. Congruence or divergence between notes in the first two columns was noted in the third column.

Table 2 shows how the data organization, using the CARMA matrix and sequential transformative design, was integrated in this study. This chapter analyzes the program expectations, evident implementation, and results columns of the CARMA matrix. Those three columns aligned with phase one, phase two, and interpretations of the sequential transformative design. The conclusions/recommendations column of the CARMA matrix will be discussed in chapter five, which also aligns with the results phase of the methodology. Both the methodology and data organizer uphold the overarching critical multiculturalism theoretical framework by working toward improving teacher evaluation through the inclusion of multiple perspectives, especially those of marginalized groups, in order to empower participants. This is important to note because transformative designs are rooted in the theoretical lens of the study (Creswell, 2009). With a critical multiculturalist theoretical lens, the data organization and analysis was conducted with the aim to improve the negative experiences of the participants with the use of PAR. The table shows the integral interrelated components of this study: data collection and analysis, methodology, and theoretical framework.

Table 2

Integration of CARMA Matrix and Sequential Transformative Design.

| Overarching Theoretical Framework: Critical Multiculturalism | | | | |
|--|----------------------------------|---|---|--|
| Data Organizer: CARMA Matrix | Column 1 Program Expectations | Column 2 Evident Implementation | Column 3 Results | Column 4 Conclusions/ Recommendations |
| Methodology: Sequential Transformative | Phase 1 | Phase 1 & 2 | Interpretation | Results |
| Data Sources | PAR Policy | District Data Teacher Survey Individual Interviews Group Interview | Juxtapose notes from Column 1 & 2 Co-construct results with participants | Co-construct the recommendations to modify or maintain program with participants |

California Peer Assistance and Review (PAR) Policy

Peer Assistance and Review (PAR) was a program for teachers who received ‘unsatisfactory’ or ‘needs improvement’ marks on their performance evaluation. At that point, the teacher was referred to PAR, as an intervention, and the intended outcome was to improve teacher performance. This was done through a peer evaluation model. For this study, the PAR policy was located on the California department of education website and sections 44500-44508 of the California Education Code were analyzed (California Department of Education [CDE], n.d.). The six main sections of the PAR policy consisted of: 1) the seven principles, 2) the consulting teachers, 3) the peer-review panel, 4) general program expectations, 5) electing to participate, and 6) receiving funds.

The California PAR policy placed a focus on local control to meet local conditions. The policy indicated seven principles with which the locally constructed PAR program must conform. The first principle was that teacher participants must be permanent employees or, in smaller schools, a probationary employee. Participants were to be referred to PAR or volunteer to participate in the program as a result of an evaluation. Also, the performance goals had to be, “in writing, clearly stated, aligned with pupil learning, and consistent with Section 44662.” Section 44662 indicated evaluation expectations. The program was required to include multiple observations of a teacher during classroom instruction. The program had to, “strongly encourage a cooperative relationship between the consulting teacher and the principal with respect to the process of peer assistance and review.” Additionally, the school districts were expected to

provide sufficient professional development to help teachers improve pedagogy and knowledge. A monitoring component was included as well. Monitoring must be included in the PAR program with a written record. The final evaluation must be placed in the participating teacher's personnel file.

The policy explained the qualification criteria for a consulting teacher who acted as the 'mentor' or 'coach' of the participating teacher referred to PAR. The consulting teacher must be credentialed, have completed at least three consecutive years of employment at the school district, have substantial recent classroom instruction experience, effective communication skills, subject matter knowledge, and teaching strategy skills. Defining qualities about each qualification were not indicated.

The next section of the policy outlined expectations for the peer-review panel. The panel had to be composed of teachers and administrators who jointly selected the consulting teachers, review their reports about the participating teacher's progress through PAR, and made recommendations for personnel decisions to the school board. The panel must tell the school board which teachers did not "demonstrate satisfactory improvement." The majority of the panel must be certificated classroom teachers and the remaining members are school administrators chosen by the school district. The panel, by a majority vote, chooses which teachers will serve as consulting teachers. The panel must also conduct an annual impact evaluation of the program. The policy states, "This evaluation may include, but is not limited to, interviews or surveys of the program participants."

When a school board accepts state funds for a PAR program in their district, they agree to work with the union to negotiate, develop, and implement the program. Further,

certificated employees cannot be in management or in supervisory functions. The cooperating teachers are protected from liability and have access to appropriate defense, just like any other employee.

Funding is further explained; not more than five percent of the money received for PAR can be allocated to administrative expenses. Administrative expenses were defined as personnel costs, the cost of consulting teachers, and any indirect costs associated with the PAR program. School districts received one thousand dollars for each cooperating teacher (also referred to as mentor teacher); the state calculated how many mentor teachers the district should have for the school year. The funding for PAR can also be used for Beginning Teacher Support and Assessment (BTSA), a district intern program, professional development or other educational activities, or a program to support training and development of new teachers.

Unclear expectations. Overall, the policy was written with vague language. Not surprisingly, the individual and group interviews, presented later in this chapter, indicated that participants felt they were held to unclear expectations. These experiences appeared to be the result of the vague language outlined in the education code. Descriptions of PAR were overwhelmingly non-specific in the policy (see Table 3), leaving many areas open for misuse, as exemplified in the participants' accounts of their negative experiences. Past research has also made similar conclusions about the challenges of implementing PAR. One major challenge reported was program ambiguity (Goldstein, 2004). Policy language is important and previous reform policy had dramatic impact on teachers and students. The policy language for NCLB assumed and implied the role of

teachers and teaching. NCLB language masked many negative consequences of the reform policy (Cochran-Smith & Lytle, 2006).

Table 3

Education Code Analysis

| Education Code Section | Vague Language | Issues and Implications |
|--|--|--|
| 44500 (4) The program shall expect and strongly encourage a cooperative relationship between the consulting teacher and the principal with respect to the process of peer assistance and review. | strongly encourage cooperation | Cooperation is not expected, it is encouraged. What does a cooperative relationship mean? |
| 44500 (5) The school district shall provide sufficient staff development activities to assist a teacher to improve his or her teaching skills and knowledge. | provide sufficient staff development | How is sufficient defined? |
| 44501 (b) The consulting teacher shall have substantial recent experience in classroom instruction. | substantial recent experiences in classroom instruction | How much is substantial experience? Why is this not quantified? |
| 44501 (c) The consulting teacher shall have demonstrated exemplary teaching ability, as indicated by, among other things, effective communication skills, subject matter knowledge, and mastery of a range of teaching strategies necessary to meet the needs of pupils in different contexts. | demonstrated exemplary teaching ability; effective communication skills; subject matter knowledge; mastery of a range of teaching strategies | How is this specifically demonstrated? How is this defined? How is this assessed? How is this assessed? |
| 44502 (a) The governance structure of a program designed pursuant to this article shall include a joint teacher administrator peer review panel that shall select consulting teachers, review peer review reports prepared by consulting teachers, and make recommendations to the governing board of a school district regarding participants in the program, including forwarding to the governing board the names of individuals who, after sustained assistance, are not able to demonstrate satisfactory improvement. | not able to demonstrate satisfactory improvement; sustained assistance | What is satisfactory improvement? What constitutes “sustained assistance”? |

| Education Code Section | Vague Language | Issues and Implications |
|--|---|---|
| 44502 (d) The panel shall also annually evaluate the impact of the district's peer assistance and review program in order to improve the program. This evaluation may include, but is not limited to, interviews or surveys of the program participants. The panel may submit recommendations for improvement of the program to the governing board of the school district and to the exclusive representative of the certificated employees in the school district, if the certificated employees in the district are represented by an exclusive representative. | evaluation may include, but not limited to, interviews or surveys of the program participants | Why is the program evaluation component not explicit and statewide? |

District Data

A critical multiculturalist framework guided the analysis of district data. Four variable categories were considered: African American teachers, female teachers, teachers over the age of 55, and pay scale classification of teachers. These variables were chosen because of their historically marginalized status and their potentially disproportionate representation in PAR. Each variable was reported by percentage of teachers within each variable referred to PAR, as compared to total number of teachers referred to PAR. A binomial distribution analysis was calculated for African Americans and women over the age of 55 to find the probability that the exact number of each variable group was referred to PAR (see Table 4). This showed the degree of randomness that teachers within each variable were referred to PAR. This data analysis addressed research question one by revealing how the PAR evaluation process negatively impacted teachers. Data would reveal a disproportionate number of teachers of each variable in PAR, thus negatively impacting specific groups of teachers. The data also addressed question two by revealing parallel traits of those in PAR. The traits of concern were the variables that were considered in analysis (race, sex, age, and pay scale). The data analysis also partially addressed research question three by contributing to the evident use

data that was juxtaposed with the intended use data. Binomial distribution analysis was chosen because of its previous use in Equal Employment Opportunity (EEO) law cases. In the EEO field, binomial distribution statistics have been used to prove discriminatory treatment and discriminatory impact (Biddle, 1995).

The district data set included various demographic variables of the district as a whole and for those placed in the PAR evaluation program in the district. The demographic data included race, sex, age and pay scale factors. The data revealed that of the teachers referred to PAR, from 2002-2012, 24% were African American. Within the entire district, African Americans accounted for 6.8% of the total number of teachers (CDE, 2013). These data were further analyzed by using a binomial probability analysis to compute the probability that 24% of teachers in PAR would be African American. This analysis indicated .027% probability that 10 out of 41 teachers referred to PAR would be African American if chosen randomly from a pool of 6.8% African American teachers.

The next demographic indicator in the dataset was pay scale, known as “step and column” within the school district. Step referred to the number of years a teacher taught in the district and column was how many units of education the teacher had completed. The higher the step and/or column, the higher a teacher was on the pay scale. The highest column on the pay scale was seven, for those teachers who had a Bachelor of Arts degree with 84 additional units or a Masters degree with 36 additional units, or a doctorate degree. The demographic data showed that 35 of the 41 teachers placed in PAR were in column six or seven on the pay scale. The average step placement of all teachers in PAR was 15 years experience. The most highly educated and most experienced teachers made up 80% of all teachers referred to PAR.

The next demographic indicator was sex and age of the teachers referred to PAR. Of the 21 women placed in PAR, 19 women were over the age of 55. The women teachers had more years of experience than their male counterparts who were placed in PAR. Another binomial distribution analysis was conducted on the probability that having 19 out of 21 female teachers be over the age of 55 was a random incident. The analysis indicated that the probability that this happened at random was 0.000000013%. The district-level age data were not made available, despite multiple requests under the Freedom of Information Act (USDOJ, 1966). State-level data were used in-lieu of hidden/missing data sets about age. The state-level demographics on age indicate that teachers older than 55 represented 21.5% of the state teacher population (CDE, 2013). In this district, 90% of women placed in PAR are over the age of 55.

Table 4

Binomial Probability

| Variable | Probability of Success *Percentage of variable within the school district | Number of Trials *Total number of teachers referred to PAR | Number of Successes *Teachers within the variable referred to PAR | Binomial Probability *Probability that the number of successes happened at random or by chance |
|------------------|--|---|--|---|
| African American | 6.8% | 41 | 10 | 0.027% |
| Women Age 55+ | 21.5%* | 41 | 19 | 0.000000013% |

Note. California state data for age used in lieu of unavailable district data.

Teacher Survey

In April 2013 a teacher created and disseminated a survey through email to all teachers (approximately 200) at the high school in the district where this study took place. The survey was focused on the health and wellbeing of teachers in regard to teacher evaluation. The survey results were made public and then later analyzed in this study. While a total of 94 responded, many did not answer all of the questions. A total of nine questions were written to elicit the negative experiences of teachers at the high school. Some responses appeared to be from participants who did not recognize or consider the teacher evaluations to be a negative experience. Those participants appeared to have positive experiences with evaluation. However, most respondents reported at least one negative aspect of the teacher evaluation system.

The survey revealed that 33% of respondents (n=26) did not feel secure about their job. Some did not worry about themselves, but saw their highly esteemed colleagues treated unfairly. Some reported fear for self and others. Of those participants, 80% reported a negative impact upon their physical or mental health, sleep, or with other relationships in their life. As reported in the survey, these negative impacts on their health included high levels of stress, depression, weight gain, despondence, insomnia, hostility toward family, deteriorating relationships, high blood pressure, and overall distraction. One participant reported,

My doctor who I have a long relationship with was disturbed by my level of stress. During my appointment he prescribed anti-depressants and time off from work. I lost at least thirty pounds due to stress. My family was very concerned about me because I was despondent. I take anti-depressants everyday now. I have to unfortunately. (personal communication, April 28, 2013)

When participants were asked about feelings of bias in their evaluation, 23% (n=19) reported feeling as though their administrator was biased against them, their style, or their pedagogy. One teacher reported in the survey that evaluation was, "decidedly biased against my style and pedagogy and possibly against me personally." This aligns with findings from previous research that indicated participants felt the teacher evaluation system was not fair or objective (Donaldson, 2012). Several respondents explained feelings about their administrator evaluator with words such as, "criticized", "uninformed", and "unqualified." Hill and Grossman (2013) described the major attributes of past policy reform failure, and evaluators being uninformed and unqualified was noted as an epic failing of the current teacher evaluation reform.

When asked for further explanation of their evaluation experience, only 55% of participants (n=52) responded. However, of the 55% who responded, 40% (n=21) reported having a negative experience. Their responses centered mostly on evaluations lacking impact or meaning, and being adversarial, harassing, and intended to find fault. Hill and Grossman (2013) also noted this finding in their research about teacher evaluation. They suggested a different approach to teacher evaluation where feedback was useful and accurate (Hill & Grossman, 2013). Also, 52% (n=42) of respondents reported that they had negative feelings about administrators coming into their class unannounced. Some respondents explained feeling anxious, nervous, apprehensive, criticized, stressed, and scrutinized.

Alternatively, many respondents were satisfied and happy with their evaluation process. They liked their administrator coming into their classroom and giving them feedback. Some respondents reported discontent with the survey because they did not see

the connection between performance evaluation and health. Some faulted the question construction and claimed inherent bias in the survey as a whole.

Are you really trying to link people's personal health issues to being evaluated as a professional? What about if I answered positively to the previous question and am experiencing excellent health, sleeping great, and am newly in love? Is that somehow connected to my evaluation or feelings of job security? Clearly you are just trying to collect quotes to support your agenda and this 'survey' is intellectually dishonest. I expect that if you report the results of this 'survey' you will include all results. Right? (personal communication, April 28, 2013)

It appeared that a completely different experience of teacher evaluation occurred between two groups of teachers, with the majority of respondents reporting that their evaluation experience was positive. The dichotomy of the positive and negative evaluation experience reported in the teacher survey revealed that indeed contrasting experiences for teachers occurred (personal communication, April 28, 2013).

Phase Two

Phase two continued with the CARMA matrix as the organizational tool for data collection and analysis (Table 2). Four participants were interviewed individually and four participants were interviewed in an online group interview. Two participants did not continue in the research for the group interview that followed their individual interviews. Two participants did not participate in the individual interview but did participate in the group interview. Their lack of continued participation is discussed in chapter five. Pseudonyms were used for all participants throughout this study and the district where they worked was only referred to as a 'the district' to keep the research anonymous. The

six participants were: Reggie, Tina, Ana, Alex, Dawn, and Judith. Reggie and Dawn did not participate in the group interview, Tina partially participated in the group interview, and Alex and Judith did not participate in the individual interview.

Individual Interviews

The individual interviews took place at a mutually agreed upon public space in the summer of 2014. Four of the six participants were interviewed individually. The two participants who did not participate in the individual interview could not schedule a time to interview and did not respond to emails with the interview questions. Those participants did, however, participate in the group interview.

Reggie's interview. Reggie participated in the individual interview and later dropped out of the study for the group interview. He sent an email explaining the difficulty he was having with administration and it was too much for him to continue for the group interview. However, the individual interview with Reggie was very insightful.

Reggie was a white male teacher who had been teaching for 17 years and been teaching at his current school for 14 years. He was 56 years old and taught mathematics in grades 9-12. Reggie had a strong background in music and studied music through a math perspective. He often taught through connecting math to music and was able to do this because of his formal education in music theory and his continuing work as a professional bass player. This year Reggie was appointed to mentor three teachers in the state mandated Beginning Teachers Support and Assessment (BTSA) program that all new teachers must complete to receive their full teaching credential. BTSA is a program to transition newly credentialed teachers into the teaching profession through the guidance and mentorship of experienced teachers.

Reggie explained that his general disposition was as a team player. He did not prefer leadership roles, but enjoyed working as a team to support the cause. He taught through a student-centered approach where he connected with students and got to know them on a human level, beyond simply teaching mathematic concepts. Many of the classes he taught were remedial credit recovery courses, as well as advanced algebra courses.

Fourteen years ago, when Reggie started teaching at the school, he received ratings of 'satisfactory' on his evaluations. Every year after that he improved and got more and more ratings of 'distinguished' on his evaluation reports. This peaked five years ago when he received 'distinguished' in all categories of his evaluation. However, this was when everything changed. The following year, after receiving a perfect evaluation, the administration changed, and with that so did his evaluations. He no longer received ratings of 'distinguished' and received many marks of 'unsatisfactory' that made him fear the infamous referral to PAR. He stated that everyone's evaluation marks went down with this new administration and explained that there was an aura of fear around the entire school.

Other teachers look around and they go 'I know that teacher. That's a good teacher and they are being harassed by the administration who doesn't really know what they are doing. So that means that could happen to me.' And so everyone has gotten into a state of panic looking around at these negative evaluations that lead to potential termination or harassment to the point where teachers don't want to work there anymore. (personal communication, June 4, 2014)

Reggie further explained that this fear came from confusion about the evaluation process. He was told for ten years that he was a great teacher and then with a new administration he was given much lower evaluation marks. It was not clear to Reggie why he was not distinguished anymore. Reggie gave an example of a time he was observed and told he had poor time management. He later explained to the evaluator why he did not have the closure activity that the evaluator had wished to see. Immediately the evaluator switched his mark to state Reggie had good time management. At that point, Reggie knew with such huge flip-flopping in evaluation, that the administrators really had no clue. “They have no idea how to make us better teachers. They are clueless. They are in over their head and they are being pushed by the hill people who want it to be a private school. They are being pushed to be harder on the teachers.” Reggie explained that “hill people” were the affluent families who lived on the hill and used their money to influence process and procedure in the schools. He explained that the ‘hill people’: “want to give some money, not as much as they would give to a private school, and give some direction to the school to influence the school.”

Reggie was suggesting that money had an influence on how policy and procedure played out in the school. He continued to describe a weak administration at the school that did not work with teachers to support student success and speak with a unified front to parents. Reggie referred to times when the administration met with parents behind the teachers’ backs, which was against the contract. He also felt that the public school was being turned into a private school, which inherently created a segregated school.

When asked to discuss his most recent experience with evaluation Reggie explained, in a frustrated manner, that his evaluator had no content knowledge and

merely focused on procedural aspects of teaching. Reggie also felt that the administrator was not experienced in an administrative role or with proper evaluation procedures.

Debrief meetings, which were required by contract, consisted of a presentation of horrible evaluation marks and an approximately six minute verbal debrief between classes. Reggie recalled sitting in the meeting, looking at the horrible evaluation, overwhelmed with the unsatisfactory marks, and thinking how scared he was that he would be put in PAR.

Reggie explained that, “PAR would be humiliating.” He was so scared that he offered to do anything so he did not have to go to PAR the following year. He placated administration out of fear of PAR. He felt that the “PAR process is there to scare people. It’s there to try to terminate people. It’s there to get the union to work with the administration to manage the teachers.” Reggie tried everything to answer any concerns from administration but always felt that there was a predetermined bias against him.

They have been more harping on test scores and listening to parent complaints than they have been looking and understanding the teaching situation. So they are coming in to the teaching situation already having made up their mind. I don’t like this teacher; I’m going to look for anything bad I can find. (personal communication, June 4, 2014)

It was as though no matter what Reggie did, they would still find fault in his work as a teacher.

The result of the evaluation was that Reggie needed to change procedural items, like closure activities and classroom management techniques of off-task students. His evaluating administrator told him to address off-task students through humiliation. The administrator told Reggie he should compare student achievement and participation

publicly. He also insisted that Reggie use the Gradual Release of Responsibility (GRR) model. GRR is one of many teaching models that can be used to improve student outcomes (Lin & Cheng, 2010). The evaluating standard that the administrator used appeared to be aligned with the GRR model, however, Reggie explained that this was not previously communicated to Reggie prior to the observation, nor was it part of the school-wide policies.

After this evaluation Reggie wrote a rebuttal, which is a teacher's right within performance evaluation. However, Reggie stated that the evaluator refused to consider the points that he had made about the evaluation. He was told the evaluation was done and it would not be updated to reflect anything stated in the rebuttal.

In another example of Reggie's negative evaluation experience, he described an incident when he was teaching an advisory course that the school mandated all teachers teach in addition to their assigned courses. Reggie explained that advisory was implemented in this school to connect with students on a personal level and he wanted to make his advisory course the best in the school. One day, when the students had completed all of the advisory assignments, the students requested that Reggie play his guitar. Reggie felt this was a good way to connect with students and inspire their participation and attendance in the advisory course. At the time he started playing, the principal and other administrators walked in the classroom and Reggie could sense they were not pleased with what they saw. Later that day he received a message to go to the principal's office. At this meeting he was told he did not follow advisory protocol when he played music to engage students. When he asked for direct feedback about how to improve, the principal would not give him directives. At that point, Reggie felt that, "he

[principal] just didn't like me.”

Other teachers were assigned to help Reggie improve during this “Pre-PAR” year where he was marked as unsatisfactory for procedural and classroom management issues. Reggie expressed feeling that those mentor teachers were not better teachers than him. He described himself as a teacher who got to know his students and has a “less-is-more” presence. He also felt the teachers who mentored him had a stronger presence in the classroom, but Reggie did not feel that it made them better teachers.

Overall, Reggie felt that the evaluation process made him a worse teacher. When he received the evaluation as being an unsatisfactory teacher, he feared the following year he would be in PAR. To protect himself, he placated the administration and advocated for himself to various other leaders in the school to speak up and help him. After all, he was a veteran teacher who was also chosen as a mentor teacher for the BTSA mentorship program for three teachers that year. These were conflicting messages about performance. How could he be an unsatisfactory teacher and chosen to be a mentor teacher for three new teachers at the school? Ultimately he was released from the PAR trajectory and given a ‘2-year pass’ until his next evaluation. He explained feelings of relief and calm, as a result, that enabled him to be a better teacher.

Reggie further discussed the concept of ‘bad teacher.’ “The term ‘bad teacher’ has become the national mantra and I think most people have no idea what that means.” He explained that sometimes a teacher is liked merely because they are “bubbly and kids like them” and they get a good evaluation for that reason. However, if a teacher knows the subject matter but is not the likable type among students and parents, then they may get a poor evaluation. It essentially came down to a popularity contest. Reggie explained that,

“the media has also heightened the few stories of bad teachers in the nation as a description of all teachers.”

Reggie believed “they [administrators] were coming after me because I’m an older teacher who gets paid more.” He explained he knew he had to play the game and give the evaluator all the procedural steps of teaching very explicitly when he was being evaluated. It was as if he was putting on a show that addressed all of the things that Reggie guessed would be important in the evaluation, in an overly expressed manner. For Reggie, he knew administrators “come in with a negative attitude toward older teachers, teachers of color, toward anybody who’s gotten negative complaints from parents.” The only time Reggie remembered receiving useful evaluation feedback was when other math teachers evaluated him. He recalled that when evaluation changed from departments to administration, he felt that was the beginning of biased, targeted attacks to push out the older teachers. For Reggie, evaluation became less helpful to the teachers and more to fulfill an agenda.

But Reggie did not completely denounce administration. “I think they [administrators] are trying, but they don’t know how [to evaluate] because none of them have really taught long enough.” He also explained that, “the administration is just trying to keep their jobs.” He felt that the administration had been, “hijacked by special interest groups making demands.” Reggie explained that the administration was misusing the evaluation process but, “they are following what they are being told from the state - to toughen it up on the teachers.” He also briefly mentioned that the union was on the decline because he did not feel their union supported the teachers. Instead, Reggie felt that the union and administration were working together (personal communication, June

4, 2014).

Tina's interview. Tina was an experienced educator with Ivy League education. She taught in the district for 15 years and a total of 19 years in her career. Tina was an African American woman who was 61 years old. She completed her undergraduate degree at the University of California in Berkeley and received a Master's degree in Education on a full scholarship at Harvard, where she received perfect marks. As a history and anthropology instructor at the high school level, she continued to take history courses throughout her career from university extension programs and local community colleges to continue increasing her content knowledge. Tina also received recognition from parent groups as teacher of the year and other awards of excellence in teaching.

Additionally, Tina had been an integral part of the Advancement Via Individual Determination (AVID) program in multiple school districts. AVID was an untracking program in grades 7 through 12 designed to support previously underachieving students, typically from low-income and minority backgrounds, prepare for entrance to universities (USDOE, 1998). The district in which this study took place disbanded AVID and Tina, along with many parents and students, fought a losing battle to keep the program alive. Tina also participated in other programs that supported college readiness for students of color.

Tina explained the power shift she noticed within the district. She explained that previously there was a strong union and veteran teacher presence in the district. When those teachers retired there was a vacuum that occurred and "there was no longer a partnership between teachers and administrators." She also explained that the union was dismantled and working with the administration against the interests of the teachers. The

school was segregated into small schools. This segregation took away power and created small schools of color within a larger school. Tina felt the district chose to, “divide and conquer the power of teachers” in this shift.

Through the evaluation process Tina received marks of proficient or distinguished until a change in administration. At that point her evaluations became a very negative experience for Tina. “So I understood that she [the administrator] was coming in not to help me, not to present, um, constructive criticism, but to find areas in which she could present a negative evaluation in my pedagogy. She, um, rarely had positive comments to make. They were very subjective.” Tina recalled an evaluation where she was marked as unsatisfactory for not having a warm-up in a two-minute segment of the class where a student presenter was having technology difficulties with their presentation. That was when Tina was referred to PAR for the following year.

The year leading up to PAR, Tina’s evaluation was conducted but not debriefed with her and filed with the district until after the evaluation deadline, in violation of the teachers’ contract. She spoke with the principal about the missed deadline and instructed the union to stay out of the issue because she had handled it herself. Unbeknownst to Tina, the union had gone to the administration to help them file the negative evaluation and place her in PAR. Tina was faced with a principal who had gone back on his word and a union that had operated behind her back to support the administration rather than a teacher. She was placed in PAR and given a consulting mentor teacher who was a former administrator in the district. She was told the reasons she was placed in PAR were because she did not have a two-minute warm-up, objectives and goals were not written on the board, and she did not have an exit ticket activity. Tina explained she had never

heard the terms that landed her in PAR. She did the components of those items, but she was not using them in the explicit fashion that newer teachers directly out of credentialing programs used them.

Tina explained the evaluation process while in PAR. The evaluator came to her class three times that year. The consulting teacher with PAR observed her class 11 times and stayed for the entire class period. The consulting teacher had data that showed that Tina had improved through the year in PAR. However, during the PAR panel meeting, following the consulting teacher's report that Tina had improved, the consulting teacher was asked to leave while the administrator presented her findings from the three observations. "It felt like the consulting teacher was being grilled when reporting back to the panel to make recommendation for further employment or termination because she wasn't making any negative comments." Despite the improvement the consulting teacher had reported, the panel recommended Tina be placed back in PAR for a second year. "The consulting teacher was shocked when she heard that I was put back in PAR." Tina explained that the consulting teacher told her:

This has nothing to do with your pedagogy: this is all political. You met the standards. I watched you meet the standards. I was there 11 times. I am angry and I am in shock. This has nothing to do with your teaching. (personal communication, June 5, 2014)

During the second year of PAR, Tina's negative experiences continued. She had a different evaluator and the critique was similar to the first year in PAR. Tina explained the procedural teaching that she was written up for: "Nobody else had to do this, only me. I'm the only teacher at the school that has to do this." Tina explained that the consulting

teacher did not know why Tina was in PAR because her teaching was so good. Soon the students found out that the district was targeting her and contacted the news to speak out about the injustice. Tina attributes her continued employment with the district to the students' activism.

However, her negative experiences did not stop there. The second year in PAR ended with a recommendation to terminate her employment. The consulting teacher changed her recommendation to the PAR panel and recommended Tina be placed in PAR for a third year. The panel explained their recommendation for Tina's termination was because she showed the movie *Gone with the Wind*. She showed this movie in the history course she taught while they were studying the Civil War and Reconstruction. The evaluator observed this lesson and initially wrote that it was a good use of film to show historical bias in Hollywood films. Then the PAR panel used the film to justify termination. However, when the PAR panel presented their recommendation for termination to the school board, they denied the recommendation because the observation data showed she passed all of the standards.

Tina explained that "PAR is a program that is meant to come in and support teachers. The way that we use PAR is not to support teachers but to target teachers for termination or to harass teachers to the point where they quit." Tina explained the plight of one of her fellow teachers who taught Latin in the district. Tina said, "She is wonderful, oh, she is awesome. She is quitting this year because she got a bad evaluation." Tina accounted for yet another teacher in the district who had taught history for over 30 years and had been a contributor to high school history textbooks. He quit because of the pressure from teacher evaluation.

Health was another major issue that arose from the negative experiences with evaluation. During the years she was in PAR she gained 40 pounds and was diagnosed as pre-diabetic. Even now that she is out of PAR Tina explained, “It’s still looming in the back of my head, what’s going to happen?”

New expectations and a new educational climate seemed to plague Tina.

The one thing that really attracted me to teaching was your ability to be creative and to use different strategies to teach different students at different levels. They have now moved to what we refer to as scripted teaching. There's a script that you're going to follow and this also ties in to Common Core, where teachers have very little autonomy, you have to follow a script. (personal communication, June 5, 2014)

Tina also explained she was not made aware of or trained in the new expectations. For her, an ideal evaluation system would include constructive criticism that is directly tied to content and not a script. “It should be a partnership where it’s give and take in a constructive manner.” However, currently “It’s anything but constructive and you feel like they have already made up their mind before they even come in.” Tina claimed her pedagogy was a different approach, but it was successful.

Policy construction was also a major concern for Tina. She felt the PAR panel was out of compliance because the administrators outnumbered the teachers. The policy stated the majority of panel members were to be teachers. Additionally, Tina did not support the switch in accountability because “there’s a systemic problem and issues need to be addressed on the pre-school level, not high school. Students have been socially promoted all the way to high school and now high school teachers are expected to make

all the changes.” (personal communication, June 5, 2014)

Ana’s interview. Ana was an extremely experienced educator with 38 years of teaching experience. She taught at the school district where this study took place for 22 years and retired one year prior to this interview, at age 65, to stop the negativity from PAR. Ana was a white female teacher who taught at the arts magnet elementary school within the district. Prior to her work there, she taught many years in nearby school districts, at children’s centers, and was the director of a nursery.

Ana considered herself a good teacher. She explained that students did well in her class, were happy to come to school, and their scores on standardized tests were very high. “I always want children to feel that they’re in the right place and doing the right thing and feeling good about themselves. I also want them to feel challenged and I think that I met those goals regularly.” The year Ana retired, to end the stress she was under from PAR, her students scored higher than the other classes in mathematics. However, she never received any recognition for her achievements as a teacher the entire duration of employment with the district.

The negative experience Ana had leading up to and while in PAR caused her to retire. I looked things over and I just said, you know this process is, I had a wonderful class last year, and I said I can’t take this stuff anymore. I can’t take the negativity anymore. And so I just said I don’t have to do this anymore.

(personal communication, June 5, 2014)

When Ana was asked about her negative experience with evaluation, she went directly into a discussion about policy design and the use of standards for evaluation. Ana attributed the subjectivity of evaluation, under the current system, to the abstract policy

language.

I think that with this evaluation, which would have been in the 2011-2012 school year, that this was with a shift in standards and I would say the standards are written in an abstract language that can shape an evaluation when an administrator, an evaluating administrator, has come in and is essentially doing snap-shots of what goes on in your class. And I think the concept of snap-shot is important because I don't think they see the overall picture. You are to demonstrate certain things and you are to demonstrate them in ways that they think, um, create the evidence that they want specifically. (personal communication, June 5, 2014)

Ana continued her discussion, expressing her doubts about evaluators coming into the classroom and checking off competencies and procedures. She felt that she was being observed through a general evaluation with very specific criteria. For her it felt like the evaluator was just looking for proof to uphold what they had already decided.

Ana had multiple examples of negative evaluation experiences. As an educator who got overwhelming praise from the reading specialists who were in her classroom daily, the principal's evaluations of her did not reflect their feedback. The principal specifically had issue with Ana's timing. Ana explained that the principal "had a specific formula of timing that apparently we were not supposed to deviate from. I was aware of it but I wanted to run a lesson that was right for the children and felt right for me. I was never criticized and in fact was lauded by the reading teachers. They thought I was doing a fine job."

Classroom management was another issue that came up in Ana's evaluations. The

year that her evaluations led her to a referral to PAR was a year with a cohort of difficult students. Many students had behavioral and special education needs.

We knew, the three teachers, three second grade teachers knew at the very beginning of the year that we needed help and support from the administration on this and we were told, ‘well you are the classroom teacher. You have to take care of this. This is your job’. Well it was each classroom teacher's job but sometimes there are situations that need more than the support of the classroom teacher and we all had a difficult year that year. Another teacher had a negative evaluation at that time and that teacher decided to retire that year and I was not ready to do that, so I didn't. (personal communication, June 5, 2014)

That following year, Ana was placed in PAR because of her timing and classroom management techniques. Ana explained that the PAR coach was very supportive and nice to work with for the year. They had good conversations and the coach provided great feedback. The PAR coach that worked with Ana also coached teachers in other districts, but Ana was the only one that she worked with in this school district. The coach was not quite sure why Ana was in PAR, “she would just shrug her shoulders because in other districts the people who were getting coaching needed it extremely and I don't think I was that needed in the extreme [*sic*].” Ana mentioned that the mentor teacher was let go by the district.

Ana emphasized that, “There’s no teacher that is infallible. Every time anyone comes in to your room for any reason you see that there is some little thing that you could have done better.” With continual room for improvement in teaching, Ana was open to suggestions for improvement. However, she expressed a negative experience with the

way the principal corrected teaching practices. “I felt like the principal would sometimes scold in a very very grave airless way” and “that [the scolding] didn’t show up in my evaluation but you can tell how something like that can kind of shadow you.”

Additionally, Ana discussed natural differences in teaching. She talked about people naturally having different approaches to talking with children. The district did not have an adopted way of delivering curriculum, such as the scripted Open Court model, so individual differences should have been acceptable. Ana indicated that the district did not require a script be followed but clearly had a format of expectations that were not made apparent to teachers until after evaluations were completed.

Ana’s advice to new teachers was composed of what things should look like. The appearance was of utmost importance, rather than a focus on quality teaching. “Your class needs to look like it's managed well, where children look like they are paying attention. Children need to look like they are engaged in their work. Needs to look like you have a structure during individual work times that accommodates any special needs. Class needs to look like it's a comfortable place for them to learn.”

As the interview continued, Ana moved the discussion to link Common Core State Standards and teacher evaluation. She surmised that she was pushed out of the district for financial reasons communicated to her directly and indirectly through policy implementation. Ana explained her systematic understanding:

It makes sense monetarily for the district to expedite getting teachers out who are close to retirement because of the training aspect of Common Core. They would have to train another teacher in delivering CCSS in the district very soon after training the older teacher. (personal communication, June 5, 2014)

Ana also explained a meeting with the principal of her school: "My principal even said at one point, 'Do you know how much this district has spent on in-servicing you?' I don't think she meant just me, but all teachers." These experiences contributed to her negative experience with teacher evaluation: "That definitely kind of lowers the ceiling on you or puts a cloud over your head."

As she approached her decision to retire while going through the year in PAR, she explained that, "I found myself withdrawn to my classroom more and more and just staying with my own cohort in recent years. It was much safer. I tried avoiding the office. It did feel negative." Ana felt that her principal just had something against her. No matter what she tried to do, the principal would continue to target her. She explained that the union was there to help but they did not always provide recourse when it involved personal attitudes and bias toward teachers: "There's some union politics that I'm not understanding these days." She also mentioned a change in district administration left her unknowing about the future.

Ana suggested an ideal evaluation include more dialogue and support from staff. She also suggested that teachers evaluate their administrators because teacher evaluation is one-sided the way it currently operates, "it isn't a process we ever dignified completely." Ana felt that further development of the evaluation process was necessary. Overall, her experience with PAR and evaluation leading up to PAR was punitive, negative, and ruled by fear. "It was like a cop giving tickets. You did this and I'm putting it in your file." It was a police state where power was exercised arbitrarily (personal communication, June 5, 2014).

Dawn's interview. Dawn was an experienced, highly involved teacher with an administrator's credential. She currently taught third grade, previously taught sixth grade, and prior to that taught a combination class with third and fourth grade students. Dawn had a Master's degree from the University of California in Berkeley and was formerly a professional musician. She was a 67-year-old, white, Jewish female teacher who had been teaching in the district for 14 years with 17 total years of teaching.

Dawn's involvement in teaching stretched far beyond the classroom. She was on every art committee within the district and had been on committees at the county and state levels. She held leadership roles in the arts foundations for public schools and was on the site council. Additionally, Dawn was a site representative for the union. Outside of education, she was a professional musician with recorded albums.

Dawn had multiple years of negative experiences with evaluation that left her in therapy battling Post Traumatic Stress Disorder from the PAR evaluation process. Although she had received high accolades from parents and teachers, she felt completely overlooked with awards from the district.

I feel invisible in terms of praise and I don't know why. But it makes the job harder because you work so hard and you just get criticized and you don't get praised on top of it. You know, it's kind of lonely. (personal communication, June 4, 2014)

The parents had even equated a year in her class to being in the best private school. She attributed this to all the extra things she provides her students. Dawn incorporates her experience with music and the arts into her curriculum. She also wrote grants annually to bring folk dancing to her school. Her grant covered the entire grade level and

incorporated elements of folk dancing, cooperation, and community.

The year that Dawn was referred to PAR she was written up for misuse of academic time. The ‘misuse’ of time was during the folk dance grant activities. The other teachers were not written up despite their involvement, as it was a grade-level grant. The grant was from an education foundation and folk dancing counted as physical education minutes. “I was so traumatized. I felt like I had been ambushed.” She felt alone with no support. As the union site representative, she had nobody to go to for assistance. Once she was referred to PAR the union stripped her of the position and the president of the union would not engage in conversation. “I was completely shocked and stressed and humiliated and embarrassed and I didn’t, like, really have a trusting relationship, on a confidential level, with people there.”

The school referred Dawn to PAR for having slow transitions, not intervening when a student with special needs was doodling while Dawn was teaching, and the timing of instruction. The observers did not have conversations with Dawn about the changes needed. Instead, she was written up and referred to PAR. Dawn explained, “I tried to get help from the union. The only thing the union will help with on *[sic]* is if it’s a timeline thing in the contract.” By contract, however, Dawn utilized her right to submit a rebuttal to the evaluation. “I spent hours writing a finely word-smithed, very revised, very factual rebuttal. Nobody read them. I spent hours writing very factual clear rebuttals.” Dawn explained that the process of rebutting evaluation was her right, but not actually utilized. Dawn felt that the PAR evaluation system was to “get rid of dead wood and start a paper trail in the evaluation process.”

Another negative evaluation experience for Dawn was when she was teaching

grammar. The school's literacy coach saw her teaching in a different way and reported her to the principal. Dawn explained that the principal came to her angrily and expressed her disapproval of teaching cursive and sentence diagraming. She was told she had "old-school" approaches to teaching and she was too traditional. Dawn felt the evaluation system was "abused and misused. I think the evaluations are useless because they are so subjective and because there is nobody to defend a teacher who feels that they have been unfairly treated." Dawn was happy to receive feedback and simply wanted to have conversations with the evaluator and work together toward improvement, rather than being placed in PAR and written up. She felt nervous about the process and equated it to The Gulag.

Dawn further explained her feelings as she faced the PAR panel meeting, where the participating teacher goes in front of the PAR panel at the end of the year to receive the recommendation for further employment or termination. "It was like an execution thing. It's like 'Prisoner A' can come in now." At that point the mentor teacher explained to the panel how Dawn, the participating teacher, improved (or did not improve) throughout the year in PAR. Dawn's PAR coach was very supportive and reported that Dawn had improved through the year in PAR, as documented in the data she collected and presented to the PAR panel. However, the PAR panel asked the mentor teacher probing questions about Dawn's use of a technique called equity sticks. Dawn used many techniques, but had not used this specific technique using names on popsicle sticks to ensure randomly selecting students. This 'failure' led her to a second year in PAR.

I thought I was going to have a heart attack when I found out that I wasn't exiting and I had to go through this again. I had no idea why because I had jumped

through all the hoops. I had done everything. (personal communication, June 4, 2014)

Dawn attributed her confusion about remaining in PAR to the PAR panel's private conversations. She was not sure why there were so many conversations about her performance that did not include her input or allowed her to even hear what was said about her performance.

The second year of PAR proved to be equally as stressful as the first for Dawn. "So year two the stress was killing me. I would sit at home and have my heart racing and I would go through my head, there's no way." She further explained the depression that ensued. "I was really anxious. It was on my mind all the time." Dawn was not allowed to have the same mentor coach the second year. She was comfortable with the previous mentor, so this inconsistency between years added stress to the situation. Dawn explained that her new mentor was not able to address the pedagogical issues. Her students scored high on standardized tests and she was fully engaged in educating her students. However, as Ravitch (2010) suggested, an emphasis on teacher evaluation has created a punitive atmosphere.

"You're guilty before you go in" was the recurring feeling that Dawn had while going through PAR. Dawn explained:

I don't believe in the evaluation process. I believe it's a sham. I believe it's a total worthless sham because, um, because of what happened to me and because I can see how they can write anything they want and there's nobody going to listen to you. Nobody is going to protect you. Maybe if you get a lawyer, but otherwise they can do anything they want and guess what, if PAR doesn't kill you, you

know, and get you out of there, the stress will. I'm just hanging in because it's like, I have to. But the stress is terrible and the fear of what if, it's always on my mind. What if I lose my job and why aren't I appreciated? (personal communication, June 4, 2014)

Dawn suggested a better evaluation system would include the teachers, create a safe space for them to express issues they may have, and work toward improvement without the immanent fear of losing their job.

Missing data. Judith and Alex did not participate in the individual interview portion of the research. They did, however, participate in the group interview. Judith did not participate because of scheduling and comfort level. Judith finally decided to participate when I emailed a notice about the group interview. She responded to the group with all questions answered. Similarly, Alex could not participate in the individual interview because of scheduling difficulties. He offered to answer the questions through email; however, he never responded to multiple messages asking for his participation. Alex did participate in the group interview (personal communication, June 4, 2014).

Group Interview

The group interview was initially planned for an in-person meeting with the participants of the individual interviews. Unfortunately, scheduling a time for the teacher-participants was very difficult because of their work obligations and out-of-school family/extra school duties they had in the evenings. Skype or Google Hangout was attempted, but many of the participants hesitated about a meeting enabled through technology. Finally, I was able to gain partial participation through group emails. I indicated to all participants that they needed to "reply-all" when responding to the four

questions. Three out of the four participants that responded (two participants did not continue in the study) did not select “reply-all.” I had to follow-up with everyone to select “reply-all” because this was a group interview conducted via emails. Other follow-ups were done with participants, as well, to merely gain further participation in the study. Reggie indicated that he could not answer the questions because he was so overwhelmed and upset by further harassment from an administrator. Tina answered the questions partially because she indicated that answering the remaining questions was emotionally upsetting her. Dawn, who had previously indicated that she had Post Traumatic Stress Disorder from the harassment she received through PAR fully dropped out after an emotional individual interview. Dawn completed the individual interview but was clearly distressed throughout the 40-minute interview. Dawn did not respond to the group interview questions but stated that she hoped that the individual interview would help change current policy and practice surrounding the PAR evaluation system (personal communication, June 27, 2014).

The group interview was focused on the PAR policy, unlike the individual interviews that were centered on the participants’ negative experiences of evaluation. The group interview was still about the negative evaluation experiences of the participants but the focus was on policy. Participants were instructed to read the California education code sections 44500-44508 and then respond to four questions. The questions were these:

- 1) [After reading the PAR policy as stated from the California Department of Education] What has been your experience with evaluation in the school district?
- 2) How does your experience with evaluation align with the state PAR policy?
- 3) Tell us about a time when you had a negative evaluation experience.
- 4) What similarities or differences do you see

in all of our negative experiences with PAR? (answer this question as part of your response to at least two others).

Three themes emerged from the group interview. Participant responses significantly overlapped in the following themes: Emotional overload, privatization and standardization of education, and inconsistent and unknown expectations. This section presents the data from the group interview grouped by these three main themes, as well as the data that did not fit within them.

Emotional overload. A total of six participants were contacted for the individual interviews that were conducted prior to the group interview. Two individual interviews did not take place because of scheduling and technology avoidance amongst the participants. When the group interview took place, three of the six participants seemed to be emotionally overloaded with the evaluation process. This overload was so serious, that it seemed to be the cause for Reggie, Dawn, and Tina to drop out of the study. Tina answered the first question of the group interview and wrote at the end, “I will finish this today. I have to take a break because this brings up too many unpleasant memories. I hope that you understand.” Tina did not send this to the entire group, as I explained in the procedural email, and she did not finish the interview response.

Reggie did not respond to the questions, but emailed me directly explaining, “I am very busy and tired of fighting the ‘unnamed’ administrators that keep changing every time the water gets hot. Then we have to jump through a new set of inconsistencies all over again.” This participant did not respond to any further emails and dropped out of the study (personal communication, June 27, 2014).

Privatization and standardization of education. The next theme that emerged from the group interview was the privatization and standardization of education. These findings supported the larger narrative in education, with specific mention of Common Core State Standards and the dismantling of unions. In Tina's response, she explained how she was not supported by the union when she told the union president that the principal had decided to "throw out the evaluation." She explained:

I asked them to stay out of it and they went behind my back and helped my evaluator so that when I returned to work I was placed in PAR, and told by the union VP that it would be good for me. (personal communication, June 5, 2014)

Instead of the union following the wishes of the union member, they pushed the PAR agenda. Tina explained, "I felt totally betrayed."

Alex also expressed experiences that indicated a larger narrative surrounding teacher evaluation. Alex stated, "I do think that it [PAR policy] conflicted with our contract which states that the purpose of evaluations is to improve a teacher's practice." Alex explained that his experience was not supportive: "There was nothing helpful offered in the evaluation, and it caused me to feel anxious." Like Tina, Alex felt that the union was not supporting its membership. Alex stated, "I believe that PAR came about as an instance of collaboration between unions and districts." This cooperation between unions and districts was a mystery to Alex. Alex furthered this statement by stating, "Teachers' unions are overdue for an investigation as to why on a statewide if not nationwide basis, they began promoting PAR and healthcare caps to locals at around the same time." His overall feelings were that the PAR policy, healthcare caps, and other "anti-teacher policies" were disastrous and divisive of teachers. One of the major

foundational issues that Alex emphasized in the group interview was that he was present when the district adopted the PAR program for evaluation, nearly twelve years ago. He stated, “PAR was not presented by union leaders as a program for improving teaching, as many now claim it to be.” He felt that PAR was actually, “a superior pathway to dismissing bad teachers.”

Ana’s discussion about factors that support a larger narrative, beyond the scope of simply an evaluation program, included targeting older teachers for financial reasons. Ana felt that she was negatively evaluated simply because she was a year away from retirement age. Additionally, Ana stated, “I wondered also whether the principle had been given directives by the district offices.” Ana surmised the reason she was targeted was for financial investment reasons, especially with the mass implementation of the Common Core State Standards (CCSS). She explained that since she was in her mid-sixties, and if she did not retire, the district would have to pay for her to be trained on the CCSS’s. She would only use this training for a couple years because she was at retirement age and the district would need to train her replacement. Ana’s responses brought forth a focus on a master narrative that suggested the timing of teacher evaluation reform and the introduction of CCSS’s was not by chance. Instead, teacher evaluation reform appeared to support the implementation of the CCSS’s (personal communication, June 27, 2014).

Inconsistent and unclear expectations. The final theme that emerged was inconsistent and unclear expectations. All respondents reported inconsistencies and unclear expectations in their evaluations. Reggie completely dropped out after the individual interview and stated being tired of “administrators that keep changing,” and that he has to “jump through a new set of inconsistencies all over again.”

Tina concurred with Reggie when she described her negative evaluation experience. She stated, “My evaluator came in with the intent of not providing constructive criticism, but rather subjectively citing petty scripted parts of the lesson from the administration that I was unaware of.” She further explained a specific instance of these unclear expectations:

She cited me for not having a warm-up while my AP U.S. History class was preparing for group presentations. The first group needed help with my laptop and I assisted them. It took all of three minutes assisting my students but when I met with my evaluator she cited that I didn’t have a warm-up on the board for my students to answer as I assisted the group presenting. (personal communication, June 5, 2014)

Tina explained she was unaware of the expectations that all warm-ups needed to be written on the board. She did have warm-ups in her lesson plans, but did not write them on the board. She was written up for this, which led to her referral to PAR.

Alex also cited unclear expectations with his evaluations. The experience that Alex shared was focused on his evaluation rating. He explained that he was “rated as ‘improvement needed’ on Standard II.” Standard II is from the California Standards for the Teaching Profession (CSTP) and pertains to creating and maintaining effective environments for student learning (Commission on Teacher Credentialing [CTC], 2009). After receiving this rating, he explained that he asked his evaluator on which standard he was rated ‘distinguished’. California uses a four-level rating system: distinguished, proficient, improvement needed, and unsatisfactory. Alex stated “My evaluator told me that having given me an ‘improvement needed’ he would not consider giving me a

distinguished rating in any other category.” There are six categories in the CSTP’s: 1) engaging and supporting all students in learning, 2) creating and maintaining effective environments for student learning, 3) understanding and organizing subject matter for student learning, 4) planning instruction and designing learning experiences for all students, 5) assessing students for learning, and 6) developing as a professional educator (CTC, 2009). Alex further explained that his evaluator told him that was his usual practice because he felt that all the standards were interconnected. Alex stated, “I felt that he was paving the way for multiple ‘improvement needed’ ratings at the next evaluation.” This was foreshadowing an eventual referral to PAR, although, at the time of this interview, Alex had not yet been referred to PAR.

Ana shared Alex’s concern about all the standards being interconnected and rated on each one together rather than individually. Ana stated, “I felt this bias, too;” furthermore, she “felt strangled by this perception which left little space to clear individual hurdles.” Ana also discussed the way standards were written abstractly, which led to various interpretations by local evaluators. Ana explained that she “wondered how subjectively they were being interpreted by my principal.” Ana gave an example of one year when she was assigned a class with tough students. Ana felt that she had been very successful with her class, from the beginning and throughout the year. She explained that parents and reading specialists gave her accolades for her success. When parents and reading specialists found out she had negative evaluations, they were surprised. Her experience was an inconsistency between perceptions of performance by different parties. Ana did feel that the PAR policy was consistent, but wondered about the local interpretation aspect (personal communication, June 27, 2014).

Outliers. In the group interview, Tina explained that her evaluation experience was positive before the implementation of PAR. Tina did not have a negative experience with evaluation in general: it was specifically the PAR program. She also expressed a cooperative relationship between the principal and herself. It was the union's involvement that caused issue for Tina. Cooperation did seem to be present at some level for Tina and the administration. Additionally, Tina explained, "The PAR process was positive because of my coach. She observed me eleven times compared with the three times that my evaluator observed me." This shows a positive aspect of the PAR program through an overall negative experience. Additionally, in Ana's interview, she stated that "It [PAR] seemed consistent with state guidelines" (personal communication, June 27, 2014).

Major Themes

Four major themes emerged from the data analysis through the CARMA matrix. These themes included, the scarlet P, kangaroo court, the Salem witch trials, and discrepant data. The themes were previously introduced in this chapter. This section further explains the implications in the final column of the CARMA matrix.

The Scarlet P (PAR)

The branding of the Scarlet P (P as in PAR) was the physical manifestation of the sin of a poor evaluation and referral to PAR. This, like in *The Scarlet Letter* (Hawthorne, 1850), was a burden to bear for teachers. In *The Scarlet Letter* (Hawthorne, 1850), society judged Hester Prynne's behavior as a sin regardless of the situation. Similarly, teachers in this study felt judged and shamed because of their ostensible lack of adherence to administrators' definition of a good teacher. The bad teacher narrative has

been hard at work through education reform since the shift in accountability from school to teacher that followed the failure of NCLB. Accountability has been hastily switched, despite the lack of evidence to show that teacher evaluation actually improves the quality of classroom instruction, improves student achievement, or increases the dismissal of underperforming teachers (Donaldson, 2012). Teachers were targeted to improve student achievement because specific interest groups and political figures made that decision (McGuinn, 2012; USDOE, 2009).

The teacher survey and interviews revealed that participants felt isolated and ashamed about their work in the school. Teacher self-doubt and blame, exacerbated by the evaluation process and based on societal focus on the bad teacher narrative, has left some teachers burdened with the mark of the P from PAR. “I was completely shocked and stressed and humiliated and embarrassed and I didn’t, like, really have a trusting relationship, on a confidential level with people there” (personal communication, June 4, 2014). Participants reported administrators working behind the teachers’ backs to support alternate agendas following district orders. Teachers were suspicious about the coincident timing of teacher evaluation reform and the implementation of the Common Core State Standards (CCSS). Participants’ suspicions were inline with what education scholars had been reporting (Ravitch, 2013, 2014; Youngs, 2013). Diane Ravitch clearly stated her position that teacher evaluation reform has ulterior motives of dismantling unions, promoting the implementation of CCSS, and the overall privatization of public schools (Ravitch, 2013, 2014). Youngs (2013) stated that, “Past attempts to enact standards-based reform have been impeded by limitations in teacher evaluation” (p. 1).

Data from multiple sources revealed a propensity toward resegregation and the systematic pushing out of teachers of color and segregating students of color. Qualified teachers had already been migrating out of urban schools (Birman et al., 2009). The district data showed a disproportionate number of teachers of color referred to PAR, branded with the Scarlet P. This further exacerbated the already evident pushing-out of teachers of colors. Ladson-Billings (2005) had warned that teacher evaluation systems could negatively impact students and teachers of color. Within this research, the number of teachers referred to PAR was not representative of the number of teachers of color within the entire district. When analyzing the policy, it did not appear to have the intent to disproportionately impact teachers of color. However, analysis through the CARMA matrix revealed that evident implementation of the policy, as evidenced in the district data and interviews, disproportionately affected teachers of color more than their white counterparts. Additionally, women and teachers over the age of 55 were also disproportionately referred to PAR. Within the scope of this research, disparate impact discrimination or disparate treatment could have been considered the result, but that discrimination classification was not the focus of the research questions. Further research and data analysis would assist in classifying the policy implications as either intentional or unintentional in the evident discrimination.

Kangaroo Court

Some data suggested that decisions about termination or referral into and out of PAR had been predetermined and not based on performance. This embodies the essence of a Kangaroo Court, to appear to be upholding justice while actually disregarding its

legal and ethical obligations. This theme was most pronounced in the individual interviews, the group interview, and the teacher survey.

Throughout the individual interviews, participants explained their negative experiences and noted their feeling that ulterior motives landed them in PAR. None of the participants believed they were referred to PAR for legitimate reasons or to improve their teaching practices, which is the stated purpose of teacher evaluation in the contract and the state education code. As an example, Tina reported that the PAR panel suggested she be terminated because she showed the film *Gone with the Wind*. The reasoning hardly seems adequate to substantiate termination. Tina and other participants felt like no matter what they did, a referral to PAR was imminent. Participants reported feeling as though evaluators were coming in to find something wrong. When participants explained their negative evaluation experiences they all spoke of very specific procedural components of teaching and learning that were listed by evaluators as reasons they should be in PAR. Examples of these include the use of equity sticks, warm-up activity written on the board, objectives written on the board, classroom management, and timing. When the participants felt they addressed the low evaluation, through changing pedagogy or writing a rebuttal statement, they still found themselves referred to PAR. Some teachers stayed in PAR for multiple years. If the intent of evaluation was to help underperforming teachers, it would seem that the procedural items at issue could have been addressed through professional development rather than a punitive structure potentially resulting in termination.

Additionally, data about PAR was difficult to obtain and sometimes denied by school districts. Demographic data about teachers referred to PAR is public data under

the Freedom of Information Act (USDOJ, 1966). Hiding data and forcing data-seekers to go to court to obtain it suggests a hidden agenda. Statistical analysis for this study required the use of state demographic data because the school district withheld their data.

Previous research supports the Kangaroo Court theme that emerged through this study. Education scholars suggested that teacher evaluation reform had ulterior motives that are aimed at dismantling the unions, promoting the Common Core State Standards (CCSS), and the overall privatization of public education (Ravitch, 2013, 2014; Youngs, 2013). Youngs (2013) explicitly stated that one of the key components for CCSS's to successfully be implemented was teacher evaluation reform. "Past attempts to enact standards-based reform have been impeded by limitations in teacher evaluation" (Youngs, 2013, p. 1). Youngs (2013) concluded that teacher evaluation reform supported a CCSS agenda rather than being used to improve teaching practices and student achievement.

The Salem Witch Trials

The Salem Witch Trial theme was rooted in the notions of false accusation, lapses in due-process, mass hysteria, and isolationism. The current issue of teacher evaluation in schools seemed to be part of larger agendas to support the bad teacher narrative and the overall privatization of education functioning in a crony capitalism model. This notion appeared to transcend themes that emerged in this study.

Following the failure of NCLB, a salvo of federal policy and grants made education a competition rather than holding high expectations for all. Policies, decisions, and research about teacher evaluation have all revolved around a small group of people and private interest groups. This was seen in various data from this research, as well as

policy and court cases. For example, *Vergara v. California* (2014) was a clear case of a money-motivated decision. The case was contrived by one private interest group with large capital. The case seemed more appropriate for a legislative review rather than a court case. It also seemed to be very narrowly focused on teachers being the cause of the children in the lawsuit not getting equal access to education. The research used to support the plaintiff was from a small group of private interest groups. This also aligned with the results of the literature review for this research. Private interest groups dominated all research about teacher evaluation and seemingly guided education reform efforts. The *Vergara v. California* (2014) case further amplified the blatant witch-hunt going after teachers.

The media also created elements of mass hysteria about teacher performance and accountability. Media coverage of education issues focused narrowly on ‘bad teachers’ as the reason student achievement was low throughout the United States (Himes, 2015; Porter, 2015). This narrow focus has also reached teacher education programs, through proposed federal regulations on teacher preparation programs (American Association of Colleges for Teacher Education [AACTE], 2014). The federal regulations were proposed to hold teacher preparation programs accountable for the effectiveness of their students once they enter the teaching field. However, many teacher education scholars and universities critique the regulations along lines similar to those of scholars and participants in this study critiquing teacher evaluation reform. These critiques included federal overreach, unfunded mandates, no empirical evidence to support the reforms, disproportionately impacting teachers and students of color, and failure to support equity (AACTE, 2015).

Further supporting the Salem Witch Trial theme, data in this study suggested that teachers in the PAR evaluation system felt isolated. One participant began to avoid interactions and stayed in her classroom where she felt safe. Another participant reported feeling like she had no support and nobody to talk to about her negative experiences. She, and other participants, felt humiliated by their placement in PAR (personal communication, June 4, 2014).

Discrepant Data

The most evident divergence of perspectives appeared in the teacher survey. The survey revealed that the majority of teachers were satisfied with the PAR evaluation system. Many reported how supportive and necessary a PAR evaluation system was to supporting student achievement. Some teachers even denounced the survey all together because, in their opinion, there was no possible correlation between teacher evaluation systems and teacher health or performance. This data confirmed previous research that found that, across stakeholders, there was a positive view of a peer evaluation system (Goldstein, 2004). Additionally, another study reported that “PAR programs offer a rigorous and comprehensive way to evaluate teachers” (SRI, 2012).

Although the majority of survey respondents supported the PAR evaluation system, thirty-three percent of respondents openly stated how negative their experiences with evaluation were. Thirty-three percent is still a large portion of teachers negatively impacted by teacher evaluation systems and should be appropriately acknowledged, despite their minority status.

Another data set that provided different perspectives was the interviews. As the researcher, I anticipated some participants to be subpar teachers who were not involved in

school activities or committees and not highly motivated to improve student achievement. However, this was far from what the data revealed. All participants were highly involved, highly educated, and passionate about the teaching profession.

Discrimination Suggested Through CARMA

This study was guided using the CARMA matrix. The third column of CARMA juxtaposed data of the policy intent, column one, and the evident use of the policy, column two. The CARMA matrix elucidated an appearance of disparate impact discrimination or treatment. Disparate impact discrimination and disparate treatment are two different forms of discrimination. It is beyond the scope of this research to identify which type of discrimination is most applicable with the data from this study. However, it is worth noting these types of discrimination and why there may be larger implications about the teacher evaluation policy.

Disparate impact discrimination is a unique form of discrimination because it is not considered deliberate discrimination, which is usually used to describe unequal or unfair practices and policies. In this situation, with new teacher evaluation practices, disparate impact discrimination appears to have operated insidiously. The intended use of evaluation policy did not appear, through this study, to be discriminatory in nature: however, the myopic policy language and implementation (at the federal and state levels) suggested that teacher performance evaluation trends qualify as disparate impact discrimination.

Disparate impact theory explains that, “When an action has a disproportionate effect on some group (racial, ethnic, gender, whatever), it can be challenged as illegal discrimination--even if there was no discriminatory intent” (Clegg, 2000, p. 79).

Analyzing the data of this study through the CARMA matrix helped elucidate the issue of teacher evaluation reform, including PAR, having major discrepancies between intent and evident use. The use of CARMA to collect and analyze data was to explore negative experiences with the peer teacher evaluation system, PAR. However, the results suggested more than negative experiences were evident here; policy issues that disproportionately, and possibly unintentionally, impacted teachers of color, veteran teachers (over the age of 55), and women. PAR evaluation of teachers, intended perhaps to operate in a fair and harmless way, actually had quite grave effects for many, as this study revealed in its close look at the negative impacts of teacher evaluation.

However, when analyzing the data of this study through the CARMA matrix and looking at the binomial probability of the district data, an argument could be made that the discrimination was far from unintentional. Disparate treatment is similar to disparate impact discrimination, but is more overt and intentional. The binomial statistics, reported early in this chapter, can be used to support a classification as disparate treatment (Biddle, 1995). Binomial probability has been used in numerous Equal Employment Opportunity (EEO) court cases to support arguments that qualify policies and practices as discriminating through disparate impact or disparate treatment. In line with critical multiculturalism, to support equity and inclusion, a further statistical analysis of PAR data from various school districts in future research is necessary to clearly define the type of potential discrimination evident in teacher performance evaluation data. A larger sample size set would increase significance in the findings and delineate discrimination classification.

Making Connections

Data collection and analysis of the five data sources was aimed to elucidate negative evaluation experiences. The research questions and theoretical framework guided the research. Additionally, a focus on the impacts of the data on teacher education was ubiquitous. The CARMA matrix, methodology, and theoretical framework worked synchronistically to support the purpose of this study. This section looks specifically at the research questions, theoretical framework, teacher education, and missing data to explicate the interconnectedness of the elements of the study.

Research Questions

This study was guided by three research questions. The questions were focused on the PAR evaluation system that was highly supported by federal mandates and grant competitions. The questions were:

- How does the PAR teacher evaluation process negatively impact teachers?
- What, if any, parallel traits exist among those teachers who had negative experiences with the PAR evaluation system?
- How does the intended use of the PAR teacher evaluation process compare to the evident use of PAR?

Phase one consisted of data from the PAR policy, a teacher survey, and PAR data. Question one, “How does the PAR teacher evaluation process negatively impact teachers?” was addressed with the PAR data, interviews, and the teacher survey. The data from PAR for this district suggested that the PAR policy disproportionately refers teachers to the PAR program based on race, age, and salary. The teacher survey explained the negative impacts in more detail from the teacher perspective. A subsection of the teachers surveyed experienced negative health impacts and a scorn for teaching

from their teacher evaluation experiences. Overall, those experiences were negative. While the majority of teachers did not report negative experiences from teacher evaluation, 20% of teachers who responded related that they had negative experiences. These experiences ranged from anxiety to major health issues that impacted them professionally and personally. The interviews also revealed how the PAR process negatively impacted teachers. One teacher's PAR experience resulted in a diagnosis from her doctor of Post Traumatic Stress Disorder. Another teacher reported that she retired early because of the stresses from the PAR evaluation process. Other teachers interviewed reported an overall feeling of scorn toward teaching and felt that their pedagogy worsened while in PAR.

Question two, "What, if any, parallel traits exist among those teachers who had negative experiences with the PAR evaluation system?" was addressed with the PAR data and interviews. The teacher survey did not reveal any parallel traits among those who had negative evaluation experiences. The survey was not designed for this research, so there were some questions that were not asked of participants that could have been useful in addressing this question. The PAR data suggested that demographic traits are similar among those referred to PAR. Similarities include age, race, gender, and salary level. Those interviewed also had similar traits.

Question three, "How does the intended use of the PAR teacher evaluation process compare to the evident use of PAR?" was addressed with the PAR policy, the PAR data, and the teacher survey. Within the CARMA matrix this was explained in the third column where the intended and evident use were juxtaposed to reveal dissonance or congruence. Question three was further addressed in the interviews in phase two of the

study. Phase one was more focused on the intended use of PAR and then deeper understanding of the negative experiences was elucidated in phase two.

Theoretical Framework

This study was approached through a critical multiculturalist framework, with the goal to empower marginalized groups (McDowell & Fang, 2007). The marginalized group in this study was the teachers who had negative evaluation experiences and remained voiceless through their plight. This approach to the data was very explicit, as the intent was to uncover the negative experiences of teacher evaluation, discover any parallel traits of those who had negative experiences, and reveal any dissonance between the intended and evident use of the evaluation policy. The perspective of this research, throughout the data collection and analysis, was to empower participants and support equity and inclusion of multiple perspectives in teacher evaluation reform.

As a balance in perspectives, Darling-Hammond's (2013) guidelines to improve teacher evaluation also framed the research. Darling-Hammond (2013) suggested that a PAR evaluation system, that utilized a peer format of evaluation, was part of an ideal evaluation system. She stated that a peer evaluation system addressed many issues in the current evaluation methods: a lack of clear consistent standards for good practices, failure to focus on improving practice, inadequate time to evaluate, minimal concern about student outcomes, rigid systems that do not consider the needs of the individual teacher, and inadequate connection between evaluation and professional development. Logically, it seemed that a peer evaluation system would address these issues. This grounded the collection and analysis of data to explore the issues of PAR.

However, as the study progressed, two major changes in the theoretical

framework occurred. First, the theoretical framework expanded. Through data collection and analysis, the critical multiculturalist framework encompassed more than just the negative experiences of teacher evaluation, as originally stated. Data suggested that PAR policy design disproportionately targeted African American teachers, female teachers, and older veteran teachers. This discovery made it evident that a critical multiculturalist perspective, along with the methodological choices of the study, was an appropriate fit for the study.

The other major change in the theoretical framework was a divergence from Darling-Hammond's (2013) suggestions that were initially used to frame and balance the research. Darling-Hammond's (2013) statement that a peer evaluation system would address the weaknesses she described in the current evaluation system was not supported by this research. Data revealed quite the opposite. In theory, if looking at the policy in isolation, one could come to the conclusion that peer evaluation would be a better way to evaluate teachers and address current weaknesses. These types of assumptions, though, may lead to flawed systems. Empirical research, like this study, showed the negative impacts of evaluation in one school district and on particular participants. When participants described their experiences in the PAR system, they noted all of the issues that Darling-Hammond (2013) noted about the current (non-peer) evaluation systems seen across the nation. Darling-Hammond (2013) noted PAR as a "bright spot" that was a "labor-management breakthrough" (p. 2). However, this study suggested that PAR was far from a breakthrough or bright spot, as Darling-Hammond (2013) claimed. Participants in this study reported that the PAR system in their school district created a negative evaluation experience because of the same issues that plagued earlier evaluation systems:

a lack of clear and consistent standards of good practices as a benchmark for evaluation, focus was not on improving practice, inadequate time for evaluation and feedback, more concern on teaching procedures rather than student outcomes, the PAR system was subjective and did not consider the needs of the individual teacher and class, and a detachment between evaluation and professional development.

Although the intent was to use Darling-Hammond's (2013) teacher evaluation suggestions as part of the theoretical framework for this study, the research took a different direction that did not support Darling-Hammond's (2013) conclusions about teacher evaluation, at least in this circumstance. With data suggesting that specific groups of teachers were targeted, critical multiculturalism was used in a way to empower participants and work toward a policy awareness to reduce potential discriminatory actions through teacher evaluation. This divergence from Darling-Hammond's (2013) suggestions led to this study's primary focus on the overarching critical multiculturalism theoretical framework.

Teacher Education

This teacher evaluation study was extremely timely in the education political climate and reform movements. Teacher evaluation and teacher preparation programs were starting to be the focus of improving student achievement. A shift in accountability occurred from schools to teachers, leading to accountability of teacher education programs. At the time of this study, teachers were designated as the most important factor in student achievement and it was decided that increased performance evaluation was the means to improve teachers and thus improve student achievement. Teacher preparation courses typically did not consider teacher evaluation as a component of the credentialing

program.

However, this myopic separation of teacher evaluation and teacher education has come to the forefront with a “Notice of Proposed Rule Making” from the federal government to increase accountability of teacher preparation programs (Kumashiro, 2015). Just as PK-12 evaluation has increased in ways that were not grounded in research, so too are university teacher preparation programs being held accountable in ways not empirically supported. This hasty movement can be connected to the overall privatization and standardization of public education. If teacher preparation programs are to be evaluated on outcomes such as placement and retention rates of graduates, and teacher performance evaluation that determine termination or continued employment, these issues should be high priority topics for the field of teacher education.

Missing Data

I have identified four major areas of missing data for this study. The first incomplete data set was the district data. The school district withheld data about district-wide demographics. This information should be made public; however the process to have the district release the information would have required court involvement. For this research, state-wide demographic data was used in lieu of district data.

Additionally, data were limited for the interviews because participants of this study could not be in any form of litigation. The study continued with alternate participants, however, the perspective of those teachers in litigation about the misuse of performance evaluation could have provided rich data for this study. One participant had to drop out between the individual and group interview because he was empowered to file a lawsuit against the school district and union because of the way he was placed in PAR.

The survey data also had limitations because the survey was not designed for this research. Many questions focused on linking teacher performance evaluation and teacher health. However, additional questions and different question structures could have provided more robust data for the study.

Lastly, three participants dropped out of the research for various reasons. Studying people's negative experiences in their current situation proved to be a difficult task because they were in high-stress situations. Had they remained in the study, their information could have provided a more complete picture of negative evaluation experiences.

Collaboration

By design, collaboration was an integral part of the study. The methodology, theoretical framework, and the CARMA matrix were all specifically chosen for this study because of the synchronistic focus on empowerment and change. The co-construction of data was a critical part of the sequential transformative design. After the initial data were recorded, it was shared with participants for feedback. Participants reviewed the data for phase one and two. Four of the six participants emailed their feedback about the data set. All of the responses reported that the data from all the data sets fit together and accurately portrayed their experiences with teacher evaluation. Participants were also pleased to see that others shared their experience.

Participants also used the collaboration time to inform me of their advocacy toward changing teacher evaluation policy. Participants had requested PAR demographic data from multiple school districts throughout the state and found similar trends about race, pay scale, and gender as was seen at their school district. Participants also referred

my findings about policy inconsistency to others they knew in litigation about teacher evaluation. Of specific interest was my finding about the role of the teachers on the PAR panel and the mentor teacher. The California education code stated that teachers could not serve in supervisory roles over other teachers; however, the PAR sections of the education code specifically stated that the mentor teacher and the PAR panel, made up of a majority of teachers, were directed to give a recommendation for further employment or termination. Based on that point, their recommendation was about personnel matters; they were de facto operating in a supervisory role. This finding was added to the cases of the teachers who were not allowed to participate in this research because they were in litigation at the time of the study.

Conclusion

The purpose of this research was to better understand current teacher evaluation reform and specifically elucidate negative evaluation experiences. Previous research had not focused on the negative experiences teachers had with teacher evaluation reform, such as PAR, that utilized a peer evaluation structure. Rather, research was focused on positive impacts of PAR and was conducted through private interest groups. There was a paucity of empirical research that accounted for multiple perspectives such as the negative experience. The sequential transformative research design considered both qualitative and quantitative data from five data sources: district data, PAR policy, a teacher survey, individual interviews, and a group interview.

The data sets collectively elucidated the negative evaluation experiences for some teachers. Although not the experience of all teachers, a subset of teachers experienced evaluation in ways contrary to supporting improved pedagogy and student achievement.

The district data revealed a disproportionate number of African Americans, women, and teachers over the age of 55 referred to PAR. Through a binomial distribution analysis, data indicated that either disparate impact discrimination or disparate treatment existed for these populations.

The analysis of the PAR policy revealed the state policy language in the education code was written in vague language that supported potential misuse or mis-implementation of the PAR evaluation system at the local level. Interviews corroborated this finding when participants explained the many ways that PAR did not support them by appropriately evaluating their pedagogy or helping them become better teachers. Their experiences revealed a negative impact on their health, personal relationships, efficacy as a teacher, and trust in the education system. Themes of misuse of power, isolation, demoralization, demonizing, and unclear expectations emerged across all data sets. These were coded as the following themes: The Scarlet P, The Kangaroo Court, The Salem Witch Trials, and Nonconfirming Data.

The group interview supported the overall negative experience reported in the individual interviews. This data was reported thematically to integrate participant responses adequately. Themes included emotional overload, privatization and standardization of education, inconsistent and unclear expectations, and outliers.

The teacher survey added support and depth to the negative experiences of teachers in the school district, while also revealing that the majority of teachers did not have negative experiences with teacher evaluation. However, when survey responses indicated that 20% of the teachers in the district reported negative health and wellness impacts from the performance evaluations and interviews elucidated extremely negative

experiences, concerns arise. Additionally, when those experiences are combined with findings from the education code policy, showing vague language that could lead to the misuse of policy, and the demographic data from those referred to PAR, showing a disproportionate number of African American, women, and older teachers targeted, the results are alarming. An imperfect and potentially harmful evaluation system is currently in use in many school districts and has been touted by the federal government as the key to improving student achievement. Chapter four has explained these findings and the emergent themes. Chapter five will continue with interpretations of these findings and relate them to the larger context of teacher education.

Chapter 5: Implications

The previous chapter reported the findings of this study about teacher evaluation reform. Chapters one, two, and three established the rationale for the study, reviewed the prior literature, and explained the methodology. As explained in chapters one and two, teacher evaluation systems, such as Peer Assistance and Review (PAR), have recently been favored in federal policy reform aimed to improve student achievement (USDOE, 2009, 2012, 2013). However, evidence was lacking that showed teacher evaluation improved the quality of classroom instruction or assisted in the dismissal of underperforming teachers (Donaldson, 2012). Additionally, existing research about teacher evaluation did not analyze the negative experiences of teachers through the evaluation process. Rather, studies were narrowly focused on schools that reported positive implementation of PAR. With this limited view of teacher evaluation reform, especially reform focused on peer evaluation, multiple perspectives had not been considered. However, despite the lack of research establishing positive results, and/or

examining negative impact on teachers, there has been an unprecedented wave of state teacher evaluation reform across the country (McGuinn, 2012). President Obama gave a call to action to improve student achievement and make education America's national mission (USDOE, 2009). The primary means to accomplish this was to hastily change teacher evaluation.

With a paucity of existing empirical research, this study aimed to better understand the potentially negative impacts of teacher evaluation reform focused on peer evaluation. Cochran-Smith and Lytle (2006) stated that teachers play a large part in student achievement and all students deserve to have effective teachers. However, "theory and practice of teacher evaluation diverge" (Darling-Hammond et al., 1983, p. 285). Teachers who had negative experiences with teacher evaluation remained marginalized through the research. This study looked specifically at negative experiences of evaluation through a sequential transformative design that was framed by critical multiculturalism. This methodology and theoretical framework, as well as the data collection and analysis tool called CARMA (Putney et al., 2006), synchronistically aimed to empower participants and inform policy makers.

Five data sources were analyzed in two sequential phases, followed by interpretation and suggestions for change. The PAR policy was analyzed, followed by an analysis of an existing teacher survey and a district data set that included demographic information about teachers referred to PAR. This led in to the second phase of the study that consisted of interviews. Individual interviews with teachers from the same school district were conducted. Reggie, Tina, Ana, and Dawn discussed their negative experiences of evaluation. Next, a group interview was conducted online with Tina, Alex,

and Ana. Due to the stressful nature of focusing on negative experiences, some participants dropped out of the study.

However, this mixed methods design, utilizing five data sources, supported a deeper understanding about how teacher evaluation negatively impacted teachers, the common traits of those who had negative evaluation experiences, and the intended versus evident use of the evaluation policy. Beyond gaining a deeper understanding, this study also aimed to give voice to a marginalized group and empower underrepresented people. Additionally, unexpected results strengthened the significance of the study.

Review of Research Questions

There were three research questions that guided this study. Each question was addressed throughout the research phases. The three questions were: a) how does the PAR teacher evaluation process negatively impact teachers?, b) what, if any, parallel traits exist among those teachers who had negative experiences with the PAR evaluation system?, and c) how does the intended use of the PAR teacher evaluation process compare to the evident use of PAR?

Interpretations of Findings

The findings of this study were extremely timely in the context of education. Teacher evaluation reform was not operating in isolation. Within the larger context of education, federal mandates and grants emphasized teacher evaluation reform at the state level. States were pressured to make hasty changes in their current teacher performance evaluation systems. Peer evaluation, such as PAR, was proclaimed to be an optimal choice of reform (USDOE, 2009, 2012, 2013). However, research was limited about PAR

and the negative experiences of teacher evaluation were not addressed in previous research. This study revealed three major findings. Figure 5 shows the major findings and their alignment to the research questions.

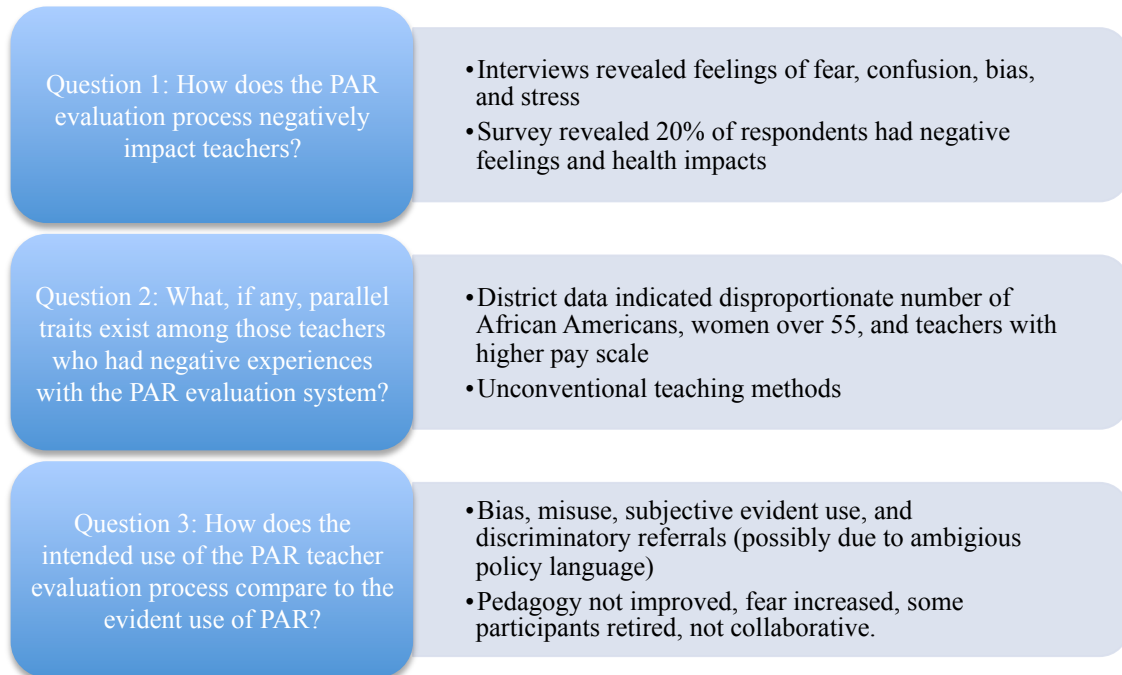


Figure 5. Major findings and their alignment to the research questions.

The first major finding, aligned with research question one, was that significant numbers of teachers identified teacher evaluation as a negative experience. Interviews gave a detailed account of participants and demonstrated how the PAR policy negatively impacted them. Reggie, Tina, Ana, and Dawn were dedicated professionals with a passion for education. Unfortunately, teacher evaluation practices plagued them with feelings of fear, confusion, bias, and stress. Teacher survey data concurred with interview data, finding that 20% of respondents had negative feelings about teacher evaluation and it had an overall negative impact on their health and wellbeing.

The second major finding, aligned with research question two, indicated similar traits of those referred to PAR. The district data revealed that teachers at a higher pay scale, more well educated, African American, and women over 55 were disproportionately referred to PAR. The interviews revealed a commonality of those referred to PAR as well. The teachers who participated in this study all seemed to have unconventional teaching practices, as defined by the norms of the current education context. Reggie discussed an instance where he incorporated music in to his lesson and was reprimanded. He also spoke about the ways his formal education in music informed his teaching of math. Tina also indicated she was attracted to teaching because she could be creative and use different strategies to teach different students at different levels. She saw her positive creative intention in conflict with what education has become for her; scripted teaching to support the Common Core State Standards. Tina said directly that her pedagogy was a different approach but successful. Ana had similar experiences with her successful, though unconventional timing of lessons and way of talking with students. Dawn also expressed unconventional methods of teaching. She incorporated music and the arts, such as folk dance, into her curriculum. Dawn also used what her principal called, ‘old school methods’, such as teaching grammar and cursive.

During participant interviews, teachers reported the rigid pedagogical expectations of teaching that were put upon them through their PAR experience and evaluation leading up to PAR. Participants felt restricted and forced to teach in ways that were not most appropriate for them or their students. They felt targeted for their alternate, though successful approaches to education. In this study, participants either conformed to the demands, to save their jobs, or they chose to retire. Ana and Judith chose to retire,

while the other participants who conformed, continued to feel targeted and expressed anxiety about future evaluations. Dawn stated, “the stress is terrible and the fear of what if, it’s always on my mind. What if I lose my job and why aren’t I appreciated?” (personal communication, June 4, 2014).

Ladson-Billings and Jackson (2007) also explained that state and federal reform, like NCLB, undermined the goals and aspirations that have historically drawn African American’s to teaching. Tina’s interview confirmed this data when she described how the opportunities to inventively teach that originally attracted her to the profession, as an African American woman, were taken away and degraded to scripted teaching. The data also revealed that 25% of teachers referred to PAR were African American while African Americans only account for 6.8% of the teaching staff in the school district.

Another common trait among interview participants, aligning with research question two, was a lack of feeling appreciated or recognized. Participants expressed receiving recognition from students and parents, but not from the school administrators. Dawn felt invisible and wondered why the school did not appreciate or recognize her dedication. Tina discussed parent and student appreciation for her but never any formal recognition from the school leadership. Reggie and Ana had similar experiences.

The third major finding, aligned with research question three, was dissonance between intended and evident use of the PAR policy, as illustrated through the CARMA matrix during data analysis (see Appendix G for CARMA analysis worksheets). The interviews, district data, and the survey revealed that a seemingly neutral PAR policy actually had negative impacts on a significant subgroup of teachers. Policy did not explicate discriminatory acts in its intended use, but the evident use, as illustrated in the

district data, targeted African Americans, women, and teachers over the age of 55. PAR also resulted in a negative impact on teacher's health and wellbeing. Tina explained that during the years she was in PAR she gained 40 pounds because of the stress and was diagnosed as pre-diabetic.

Another point of divergence between intended and evident use was the amount of ambiguous language that was perpetrated throughout the policy. Examples of vague language include, "strongly encourage cooperation", "sufficient", and "substantial". These terms were neither directive nor measurable and supports Goldstein's (2004) finding that program ambiguity was a challenge to implementing PAR. This could have resulted in the amount of confusion and frustration that participants expressed throughout their interviews. After a year in PAR and fulfilling all requirements, Dawn expressed that, "I thought I was going to have a heart attack when I found out that I wasn't exiting [PAR] and I had to go through this again. I had no idea why because I had jumped through all the hoops. I had done everything" (personal communication, June 4, 2014). The evident use ambiguity during evaluation in this district was described in various extremely negative terms, such as: the Gulag, prison, and a kangaroo court.

Participants did not understand expectations. All participants in the individual interviews and group interview indicated confusion about why they were evaluated poorly and what they could do to improve their rating. The PAR policy was also written vaguely, leaving space for local interpretation and implementation. Concurrently analyzing these findings shows a potential interrelatedness. PAR policy was written in vague language to allow it to meet local needs. However, vague expectations left room for misunderstanding among the population the policy was designed to help.

Major Themes

The aforementioned major findings of this study addressed the three research questions and were framed by the theoretical framework to give voice to a marginalized group and advocate for change. Through that lens, data collection and analysis revealed four main themes (see Appendix F for thematic analysis tables). The themes will be discussed in this section, followed by unexpected results, implications for teacher education, limitations, researcher reflections, and conclusion.

Theme One: The Scarlet P

The Scarlet P theme addressed research question one and encompassed reports of shame, fear, isolation, a negative view of unconventional teaching methods, sin, poor health, and negative impacts on relationships. This theme was mostly focused on the individual impact and results. The teacher survey, although not specifically designed for this study, revealed teachers did not feel comfortable with evaluation. Responses included negative health impacts such as, high levels of stress, depression, weight gain, despondence, insomnia, hostility toward family, deteriorating relationships, high blood pressure, and overall distraction. The majority of those who did not feel secure in their job reported a negative impact on their physical or mental health, sleep, or with other relationships in their life. Likewise, the individual interviews revealed feelings of fear, shame, and isolation. Reggie, Tina, Ana, and Dawn mentioned the shame and fear about the PAR system. Interestingly, they all reported utilizing unconventional methods as well. Reggie and Dawn incorporated their expertise in music to their lessons. Ana and Tina used student-centered approaches to teaching to meet the unique needs of their students.

Their unconventional methods were often used as the reason for their referral to PAR.

This can be equated to the adulterous branding of Hester Prynne.

Theme Two: The Kangaroo Court

The kangaroo court theme addressed research question three and was illustrated in the research through notions of unfair practice, ambiguity, subjectivity, predetermined judgments, and ulterior motives. This theme encompassed the evident use of PAR and revealed negative experiences with the union, privatization, and standardization of education. PAR policy was written ambiguously with significant grey area that could be attributed to the bias and subjectivity that participants reported. The teacher survey also revealed the perception of bias against teachers individually, their pedagogy, or their style. All participants in the individual interviews discussed changes in the administration at their school and district preceding judgment against them. Participants noted administration change as the point of negative change in their evaluation experience. This further supported the subjectivity of the PAR evaluation system.

Theme Three: The Salem Witch Trials

The Salem Witch Trials theme addressed all three research questions and was defined by experiences of false accusations, lapses in due-process, mass hysteria, and isolationism. This theme was applied more to the group level rather than individually. Bias revealed in the district data illustrated a disproportionate referral to PAR from specific demographics (race, age, gender, and pay scale). In the teacher survey data, respondents reported feelings of division among teachers and viewing each other through an “us and them” lens. They saw their highly esteemed colleagues treated unfairly, resulting in a punitive fearful atmosphere among all. Individual interviews corroborated

this finding as well. Teachers reported having to “play the game” and other survival techniques. Interviews also revealed that teachers felt their teacher evaluations were predetermined, regardless of their performance or improvements. Participants speculated teacher evaluations were insidiously connected to a larger movement toward privatization, standardization, implementation of Common Core State Standards, and the dismantling of unions. This finding was in stark contrast to Papay and Johnson’s (2012) finding about PAR’s connection to the union, “PAR encouraged a culture of collaboration not only among teachers but also between labor and management at both the school and district levels” (Papay & Johnson, 2012, p. 723).

Theme Four: Nonconfirming Data

Nonconfirming data consisted of positive findings about evaluation and specifically PAR. Although these data were approached to elucidate the negative experience of evaluation, it is important to recognize the data to support successes of teacher evaluation reform. The teacher survey had the largest sample size and indicated that most teachers appreciated the peer evaluation process. Interview participants also reported positive experiences through evaluation. Ana, Dawn, and Tina discussed the good conversations they had with their mentor teachers. They also liked the in-depth observations that happened regularly, as opposed to the two to three basic observations from the traditional principal-centered observation system. These data supported previous research that found school stakeholders had a positive view of the new collaborative process of PAR (Goldstein, 2004) and that, “PAR programs offer a rigorous and comprehensive way to evaluate teachers” (SRI, 2012).

Unexpected Results

The collaborative structure of this study seemed to create a movement toward change, which was the purpose of the study, but happened in unexpected ways. Data sources emerged for future research and teachers began to speak out about the PAR evaluation system. These unexpected results have inspired future research and advocacy for participant and have been integrated in to court cases already.

While examining the PAR policy within the California Education Code, I discovered incongruent sections that suggest that the PAR proceedings may be against education code and illegal. This finding was included as evidence in a discrimination case with one school district. The finding centered on the role of management and questioned who was allowed to manage teachers. The education code clearly stated that teachers cannot serve in managerial positions over other teachers (CDE, n.d.). However, PAR is designed to have a panel of teachers and administrators who make personnel recommendations to the school board about the continued employment of the teacher referred to PAR. In addition, the mentor teacher, who works with the participating teacher for the year to improve teaching practices, makes a recommendation to the PAR panel. This recommendation is their opinion about the growth of the participating teacher over the year and opines if the teacher is up to proper performance for the school. The mentor teacher's opinion directly impacts the recommendation of the panel, which consists of other fellow teachers, to the school board. The school board bases their decision to terminate or retain the participating teacher based on the reports from the teachers. With this said, it is clear that the PAR evaluation system has teachers performing managerial or supervisory duties over other teachers. This is against the California Education Code 44503B (CDE, n.d.).

Additionally the Public Employment Relations Board (PERB) has defined what constitutes a manager or supervisor. This definition was indicated in two court cases addressing what was allowed by the education code (Berkeley Unified School District v Berkeley Federation of Teachers, 1979; San Mateo City School District v PERB, 1983). Applying those cases to the structure of PAR would suggest that the PAR policy is out of compliance, however, a deeper analysis is necessary and out of the scope of this research.

The cooperative nature of this study also brought forth media coverage and the sharing of documents. Participants, and non-participants who heard about the research, shared articles, radio broadcasts, and data to support a larger understanding of the negative impacts of performance evaluation on teachers. A support and research network was formed.

Another unexpected result was the lack of sustained participation for the interviews. I did not anticipate the difficulty in scheduling or the consequences of the outside stresses of their teaching situation that would limit participation. These results make sense, but were unexpected.

Additionally, the hesitance to use technology during the group interview was unexpected and worth noting. This finding, although seemingly insignificant, could attribute to negative evaluation experiences. This coupled with the unconventional methods of teaching that participants reported utilizing could have been a key factor in to their referral to PAR. This is not to say that the PAR system is infallible, but poor technology use could be a contributing factor. This is also something that could be addressed in a professional development course rather than a referral to PAR.

Implications in Teacher Education

Teacher education research and practice is at an interesting crossroads. Previous focus on teacher evaluation was minimal within the field of teacher education. This divergence has left the research foundation minimal. At the time of this study, teacher evaluation within k-12 schools was increasingly in punitive in nature. Closely following, has been a focus on increased accountability on teacher education programs. Federal mandates on teacher education programs have been proposed to hold programs accountable for teacher performance and retention. This is directly connected to teacher evaluation reform. Increased research and activism throughout the teacher education community is imperative.

Based on the findings, it appears that teacher education programs could improve teacher effectiveness by studying and developing lifelong learning courses for teachers at different stages of their career to support success and retention. This would necessitate a more cooperative relationship between university programs and school districts.

Another topic in this study was the number of teachers of color and has been of specific interest in previous research. Education scholars have reported concerns about the systematic pushing out of African American teachers (Ladson-Billings, 2005; Ladson-Billings & Jackson, 2007; Nieto, 2006). Nieto (2006), and have discussed the ill effects of the standardization movement, which is what has happened in teacher evaluation and within the larger context of education. Nieto (2006) stated that standardization in schools is, “limiting the kinds of pedagogical approaches that teachers use, as well as constricting the curriculum, especially in classrooms serving the most educationally disadvantaged students” (p. 460). Nieto’s sentiments were supported by the findings of this study. Analysis of the data set of teachers referred to PAR revealed that

24% of teachers referred to PAR were African American, while the district had a mere 6.8% African American teacher population. With the use of a binomial distribution analysis, these statistics suggested disparate impact discrimination or disparate treatment, as it has proven in past Equal Employment Opportunity law cases (Biddle, 1995).

Teacher education must take a more concerted effort to increase the number of teachers of color. Evaluation programs at the K-12 level, and now proposed accountability policy at the university level, have the potential to decrease the already limited diversity in education. This needs to be studied further to promote a diverse teacher workforce that mirrors its student population.

Limitations

The initial limitations to this study were the scope and sample size. The study was bound by one district, but was still able to address the research questions and provide a better understanding of the negative experiences of teacher evaluation. This does not appear to have been a limitation of the study. However, the study could have been strengthened by data from other school districts. This suggests necessary further research that incorporates more data.

The sample size of teachers interviewed was anticipated to be a limitation as well. This did prove to be a limitation of the study. Not only was the initial sample size small, but also the continued participation was difficult to sustain. A larger pool of participants would have helped with this issue and improved significance of the findings. However, participation drop-out based on stress and overload from the evaluation process actually supports the overall significance of the negative impacts. Also, expanding participant requirements to including participants in litigation could have provided richer data.

Recommendations for Further Study

This research used the CARMA matrix for analysis. Following CARMA, conclusions and recommendations are made in what Putney et al. (2006) call NoteReMaking. NoteReMaking explicates the evaluators' and/or stakeholders' interpretations after collecting and analyzing the intended use, evident use, and results. Implications and suggested modifications to existing programs are developed in relation to the users/participants in the program, who was served or involved, how participants were served, and what the program produced.

The PAR evaluation system appeared to be the model teacher evaluation system that the federal and state policies and grants supported in current education reform. However, education research did not support a peer system of teacher evaluation as a way to improve student achievement. Rather, the lead researchers in teacher education made several recommendations to improve student achievement, rooted in empirical research and experience, which do not include teacher evaluation reform. The existing research that did support PAR evaluation systems was conducted through private interest groups and coupled with the promotion of the CCSS's.

The data sets in this research indicated that while teacher evaluation may be a contributing factor to student achievement, as shown in previous research through private interest groups, the way teacher evaluation was being used was not optimal and had negative effects. Data revealed that in some instances a peer formatted teacher evaluation system targeted specific groups of teachers, peer evaluation policy was written vaguely and enabled bias at the local level, and the health of teachers was compromised.

Based on the findings of this research, prior research about teacher evaluation, and education code and policy, urgency toward peer evaluation systems must be stopped. Other systems, such as academic mentorship programs, and other means to increase student achievement should be considered. Shifting accountability to teachers, rather than schools, has major negative impacts on teachers and the education environment and has not been linked to increased student achievement. Teacher evaluation is a necessary and important component of education. However, choosing a peer evaluation system, such as PAR, has shown to negatively impact teachers and may not be the most effective way to increase student achievement.

Further study is necessary to adequately enact high-stakes teacher evaluation reform that is just and serves to improve student achievement. Current teacher evaluation reform lacks a foundation in sound research from experts in the field of education. This is a call to education researchers to take immediate focus on studying current trends in teacher evaluation, as well as the impact on student achievement. Mead (2012) emphasized the importance of states and schools continuing to evolve their evaluation practices as they are informed about successes and mistakes. This study informs practice, by increasing understanding about how teachers are negatively impacted by evaluation, and should be used to improve evaluation systems so they accurately address teacher quality and student achievement while considering multiple perspectives.

Hill and Grossman (2013) suggested that new evaluation systems be subject-specific, include content experts, provide accurate and useful information for teachers, and complement existing systems. This study indicated that negative experiences stemmed from a lack of what Hill and Grossman suggested for better evaluation systems.

Further planning and research that includes education scholars could create a dynamic evaluation system that includes Hill and Grossman's (2013) elements and supports teacher quality and improved student achievement.

This study used five data sources that proved to be very rich in data and could support future research. Looking more in-depth in to each data source individually would increase the working knowledge of the negative impacts of peer evaluation systems that reformists claim to be the answer to improving student achievement. More knowledge about these impacts could help reformers better understand how teacher evaluation policy translates in to the schools and reduce dissonance between intent and evident use of evaluation policy.

Expanding the district data to multiple school districts across the nation could provide a more complete statistical analysis of those referred to the PAR evaluation system. The teacher survey could also be expanded and designed by an educational researcher rather than a teacher not trained in survey methods. The questions could be more intentionally aligned with the research questions. The PAR policy should be further analyzed and could be compared to other state education codes. Interviews, both individual and focus group, could be expanded to include more participants and ask deeper questions about the teacher evaluation experience. This would provide more information about the parallel traits that exist among those referred to PAR and the ways PAR teacher evaluation negatively impacts teachers.

Further research within the larger context of teacher education is imperative. This study should be expanded to connect more robustly to university teacher preparation programs and teacher educators. The topic of this study foreshadows accountability shifts

from PK-12 moving up to the university level. Now, more than ever with proposed increases in evaluation at the university level, teacher preparation programs must consider evaluation procedures. Teacher preparation programs are currently under critical scrutiny, just as the teachers are in PK-12 schools. While teacher evaluation increases rigid expectations, with student outcomes calculated in to teacher performance evaluation and peer systems are implemented, university teacher licensure programs are likely going to be evaluated on teacher success once they are graduated from university credentialing programs. Looking even further, in this same reform pattern, teacher educators are the next shift in accountability. At the time of this research, federal regulations had been proposed to increase accountability for teacher education programs. Teacher education research must look at the connection and reveal ways that best support a diverse group of teachers and improve student achievement.

Currently major dissonance exists between research, policy, and student/teacher outcomes. This is not a new problem, as Darling-Hammond et al. (1983) explained over thirty years ago that, “theory and practice of teacher evaluation diverge” (p. 285). Universities and school districts could improve teaching and student achievement through collaboration. Currently, many school districts work in isolation, completely disconnected from university credentialing programs. Through this collaborative and transformative research, it has become apparent that joint collaboration could be the single most powerful approach to improving education. Without action, following trends in PK-12 schools, teacher educators will be the next to experience a push in evaluation reform.

Additionally, a critical study that looks at the emergence of what appears to be disparate impact discrimination surrounding teacher evaluation systems, such as PAR, is necessary to ensure teachers of color, veteran teachers, and female teachers are not adversely targeted through a seemingly neutral evaluation policy. Clearly, a subgroup of teachers has reported negative experiences through teacher evaluation. Their negative experiences should be respected through critical analysis and outlet.

Researcher Reflections

The topic of teacher evaluation was of great interest to me as a result of my negative experiences with the misuse of evaluation and witnessing others treated unfairly through the evaluation process. This perhaps could be viewed as research bias, however, the research design, theoretical framework, and data collection and analysis were intentionally chosen to reduce bias and assumption. Throughout the study, the focus was on the research question and the purpose of better understanding negative evaluation experiences. Additionally, the study focused on PAR, of which I had no personal experience and had minimal knowledge.

Before the study I assumed that some of the participants would reveal some valid deficiencies in their pedagogy that would justify a referral to PAR. However, all of the participants appeared to be dedicated professionals who may have practiced variations of anticipated pedagogy from the evaluator. Professional development and school-wide expectations could have been established to address the variation in pedagogy if the school truly took issue with what the teachers were doing in the classroom.

After conducting this research I recognize how widespread the misuse of evaluation is in schools and the interconnection of teacher evaluation reform in the larger

context of education. Initially I had no understanding of how Common Core State Standards and unions could be tied to teacher evaluation. I found it precarious that federal mandates and funding hinged on teacher evaluation reform. Conducting this research has served to expand my perspective and motivated me to continue with this line of research.

Conclusion

Prominent education scholars discussed ways to increase student achievement but did not attribute increased teacher evaluation as the means to do so. Education scholars actually warned about the challenges of implementing PAR and the focus on teacher evaluation to improve student achievement. In countries with the highest student achievement rates, such as Finland, strict teacher evaluation systems did not exist. They took alternate approaches to improving student achievement that focused on strict preservice teacher qualifications and not firing teachers (Ravitch, 2012).

This advanced mixed methods study explored the negative experiences of teacher evaluation. Chapter one introduced how education reform was focused on teacher evaluation to address low student achievement and provided a rationale for exploring a broader perspective to better understand the implications it had on teaching and learning. Chapter two reviewed the existing research about teacher evaluation and identified themes that emerged throughout the literature. Chapter three provided specific details about the methodology design of the study and the use of the CARMA data collection and analysis tool that supported a critical multiculturalist theoretical framework. Chapter four explained the results of each data source and the themes that emerged through a

holistic data analysis. Finally, this chapter has interpreted the results from chapter four and explained the implications. Throughout each chapter, the critical multiculturalist theoretical framework guided data collection, methods, and thematic analysis.

Critical multiculturalism, with a commitment toward social justice, framed the study to promote transformation of the intended and evident use of PAR, in terms of participants and policy. Critical multiculturalism was closely tied to the emergent themes. The scarlet P, the Salem witch trials, and the kangaroo court were chosen as overarching themes based upon issues central to critical multiculturalist issues, such as unjust systems, unequal and unfair treatment of women and minority groups, and empowering marginalized groups.

The literature implied that existing research about teacher evaluation was limited and research about the negative impacts of evaluation was non-existent. Additionally, research supporting a peer evaluation system, as promoted by federal policies and grants, was developed from private interest groups. According to the five data sets included in this study, negative impacts of teacher evaluation did occur. These impacts include deteriorating health, pushing out of teachers of color and veteran teachers, negative feelings toward teaching, and reduced teacher effectiveness.

Myopic policy design and implementation, without a foundation in empirical research, leaves room for policy misuse and abuse, as well as a divergence from a focus on improving student achievement. This study has elucidated the negative impacts of a peer evaluation system in one district. A cross examination of multiple data sources improved understanding of the negative evaluation experience. Although one might assume that a teacher with a negative evaluation experience was simply a bad teacher,

this study has shown dedicated and experienced educators negatively impacted by the PAR evaluation system. PAR policy and demographics of referred teachers also shows vague and discriminatory use of evaluation. Policy makers and school district officials must take heed of multiple perspectives and consider the negative impacts of teacher evaluation reform. Evaluation systems that prioritize teacher learning over accountability are integral to successfully improving student achievement.

References

- American Association of Colleges for Teacher Education (2015). *AACTE — on behalf of 800 teacher prep programs — submits comments on proposed federal regulations*. Retrieved from <http://aacte.org/news-room/>
- Adamson, F. & Darling-Hammond, L. (2012). Funding disparities and the inequitable distribution of teachers: Evaluating sources and solutions. *Education Policy Analysis Archives*, 20(37). Retrieved from <http://epaa.asu.edu/ojs/article/view/1053>
- Aughinbaugh, A. (2001). Does Head Start yield long-term benefits? *Journal of Human Resources*, 36(4), 641-665.
- Banchero, S. (2012, December 11). Competitors still beat U.S. in tests. *The Wall Street Journal*. Retrieved from <http://online.wsj.com>
- Berkeley Unified School District v. Berkeley Federation of Teachers, 101 PERB (1979).
- Biddle, R. E. (1995). The role of two statistical approaches in EEO cases. *Biddle & Associates*. Retrieved from <http://www.biddle.com/documents/twostats.htm>
- Birman, B. F., Boyle, A., Le Floch, K. C., Elledge, A., Holtzman, D., Song, M., ... Yoon,

- K. S. (2009). *State and local implementation of the No Child Left Behind Act: Volume VIII-Teacher quality under NCLB final report*. Jessup, MD: U.S. Department of Education.
- Bode, P., & Nieto, S. (2003). Review of intergroup dialogue: deliberative democracy in school, college, community and workplace. *Teachers College Record*, 105(7), 1298-1301.
- California Department of Education. (n.d.). *California Laws & Codes*. Retrieved from <http://www.cde.ca.gov/re/lr/cl/>
- California Department of Education. (n.d.). *DataQuest*. Retrieved from <http://www.cde.ca.gov/ds/>
- Clegg, R. (2000). The bad law of “disparate impact.” *Public Interest*, 138, 79-90.
- Creswell, J. W. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches* (3rd ed.). Thousand Oaks, CA: Sage Publications, Inc.
- Cochran-Smith, M., & Lytle, S. L. (2006). Troubling images of teaching in No Child Left Behind. *Harvard Educational Review*, 76(4), 668-726.
- Commission on Teacher Credentialing. (2009). *California Standards for the Teaching Profession (CSTP)*. Retrieved from <http://www.ctc.ca.gov/educator-prep/standards/CSTP-2009.pdf>
- Darling-Hammond, L. (2012). *Creating a comprehensive system for evaluating and supporting effective teaching*. Stanford, CA: Stanford Center for Opportunity Policy in Education.

- Darling-Hammond, L. (2013). *Getting teacher evaluation right: What really matters for effectiveness and improvement*. New York City: Teachers College Press.
- 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.
- Donaldson, M. L. (2012). Teachers' perspectives on evaluation reform. *Center for American Progress*. Retrieved from www.americanprogress.org.
- Education Counsel. (2014). *TLE Legal Update: Vergara v. California Court Invalidates State Tenure, Dismissal, and Last-in-First-Out Statutes*. Retrieved from <http://www.educationcounsel.com>.
- Goldstein, J. (2004). Making sense of distributed leadership: The case of peer assistance and review. *Educational Evaluation and Policy Analysis*, 26(2), 173-197.
- Hawthorne, N. (1850). *The Scarlet Letter*. Boston: Ticknor, Reed & Fields.
- Hill, H. C., & Grossman, P. (2013). Learning from teacher observations: Challenges and opportunities posed by new teacher evaluation systems. *Harvard Educational Review*, 83(2), 371-384.
- Himes, T. (2015). Tentative ruling: LAUSD must repeal teacher evaluation system. *Los Angeles Daily News*. Retrieved from www.dailynews.com
- Kumashiro, K. (2015). Review of proposed 2015 federal teacher preparation regulations. *National Education Policy Center*. Retrieved from <http://nepc.colorado.edu/thinktank/review-proposed-teacher-preparation>
- Ladson-Billings, G. J. (2005). Is the team all right?: Diversity and teacher education. *Journal of Teacher Education*, 56(2), 229-234.

- Ladson-Billings, G. J. & Jackson, J. F. L. (2007). *The fragmented world of the social: Essays in social and political philosophy*. New York City: SUNY Press.
- Lin, N. C. & Cheng, H. (2010). Effects of gradual release of responsibility model on language learning. *Procedia Social and Behavioral Sciences*, 2, 1868-1872.
- Markley, T. (2004). Defining the effective teacher: Current arguments in education. *Essays in Education*, 11(3), 1-14.
- Markow, D., Macia, L., & Lee, H. (2013). *Challenges for school leadership: The MetLife survey of the American teacher*. Retrieved from <http://www.metlife.com/teachersurvey>
- McDowell, T. & Fang, S. S., (2007). Feminist-informed critical multiculturalism: Considerations for family research. *Journal of Family Issues*, 28(4), 549-566.
- McGuinn, P. (2012). The state of teacher evaluation reform. *Center for American Progress*. Retrieved from <http://www.americanprogress.org>
- Mead, S. (2012). Recent state action on teacher effectiveness: What's in state laws and regulations? *Bellwether Education Partners*. Retrieved from <http://bellwethereducation.org>
- Merriam, S. B. (2009). *Qualitative research: A guide to design and implementation*. San Francisco: Jossey-Bass.
- Merriam-Webster. (n.d.). retrieved from <http://www.merriam-webster.com/dictionary/evaluate>
- National Commission on Excellence in Education. (1983). *A Nation at Risk*. Retrieved from http://datacenter.spps.org/uploads/SOTW_A_Nation_at_Risk_1983.pdf

National Commission on Teaching and America's Future. (1996). *What Matters Most*.

Retrieved from <http://nctaf.org>

National Council on Teacher Quality. (2012). *NCTQ yearbook brief: State of the states 2012*. Washington, DC: Sandi Jacobs.

Nieto, S. (2006). Solidarity, courage and heart: What teacher educators can learn from a new generation of teachers. *Intercultural Education*, 17(5), 457-473.

Noguera, P. (2011, November 29). We must do more than merely avoid the NCLB train wreck. *Huffington Post*. Retrieved from <http://www.huffingtonpost.com>

Noguera, P. A. (2003). The trouble with Black boys: The role and influence of environmental and cultural factors on the academic performance of African American males. *Urban Education*, 38(4), p. 431-459.

Organization for Economic Co-operation and Development. (2012). *PISA 2012 results*. Retrieved from <http://www.oecd.org/pisa/keyfindings/pisa-2012-results.htm>

Papay, J. P. & Johnson, S. M. (2012). Is PAR a good investment? Understanding the costs and benefits of teacher peer assistance and review programs. *Educational Policy*, 26(5), 696-729.

Porter, C. (2015). Teacher evaluation plan draws new support: Coalition of teacher-preparation groups backs controversial federal plan, setting up showdown. *The Wall Street Journal*. Retrieved from www.wsj.com

Putney, L. G., Wink, J., & Perkins, P. (2006). Teachers as researchers: Using the Critical Action Research Matrix Application (CARMA) for reflexive classroom inquiry. *Florida Journal of Teacher Education*, IX, 23-36.

Ravitch, D. (2010, March 9). Why I changed my mind about school reform: Federal

- testing has narrowed education and charter schools have failed to live up to their promise. *The Wall Street Journal*. Retrieved from <http://online.wsj.com/>
- Ravitch, D. (2012). *How are teachers in Finland evaluated?* Retrieved from <http://dianeravitch.net>
- Ravitch, D. (2013). *Bombshell: Eli Broad secretly funded anti-public school, anti-union initiatives*. Retrieved from <http://dianeravitch.net>
- Ravitch, D. (2014). *Common core*. Retrieved from <http://dianeravitch.net>
- Rethinking Schools. (2006, Jan). *Getting to the heart of quality teaching*. Retrieved from http://www.rethinkingschools.org/archive/20_02/edit202.shtml
- Rowley, R. L. & Wright, D. W. (2011). No “White” child left behind: The academic achievement gap between black and white students. *The Journal of Negro Education*, 80(2), 93-107.
- San Mateo City School District v. Public Employee Relations Board, 33 Cal. 3d 850 (Cal. 1983).
- SRI International (2012). *Evaluating teachers: New study underscores effectiveness of rigorous peer assistance and review (PAR) programs*. Retrieved from www.sri.com
- Trends in International Mathematics and Science Study. (2011). *TIMSS 2011*. Retrieved from <http://timss.bc.edu>
- U.S. Department of Education. (1994). *Goals 2000: Educate America Act*. Retrieved from <http://www2.ed.gov/legislation/GOALS2000/TheAct/index.html>
- U.S. Department of Education. (1998). *Advancement Via Individual Determination (AVID)*. Retrieved from <http://www2.ed.gov/pubs/ToolsforSchools/avid.html>

U.S. Department of Education. (2002). *No Child Left Behind: A desktop reference 2002*.

Washington, DC: Office of the Under Secretary.

U.S. Department of Education. (2005, November 29). *New No Child Left Behind*

flexibility: Highly qualified teachers. Retrieved from

<http://www2.ed.gov/nclb/methods/teachers/hqtflexibility.html>

U.S. Department of Education. (2009). *Race to the Top program executive summary*.

Retrieved from <http://www2.ed.gov/programs/racetothetop/index.html>

U.S. Department of Education. (2012). *Teacher Incentive Fund*. Retrieved from

<http://www2.ed.gov/programs/teacherincentive/index.html>

U.S. Department of Education. (2013). *ESEA flexibility*. Retrieved from

<http://www2.ed.gov/policy/elsec/guid/esea-flexibility/index.html>

U.S. Department of Justice. (1966). *The Freedom of Information Act*. Retrieved from

<http://www.foia.gov/about.html>

Youngs, P. (2013). Using teacher evaluation reform and professional development to

support common core assessment. *Center for American Progress*. Retrieved from

<http://www.americanprogress.org>

Appendix A

Complementary Analysis Research Matrix Application (CARMA)

| 1 | 2 | 3 | 4 |
|---|--|--|--|
| Program/Community/ Classroom Expectations | Evident Implementation | Results | Conclusions/ Recommendations |
| NoteTaking | NoteTaking | NoteMaking | NoteReMaking |
| Management/program staff/administrator/teacher | Users/participants | Compare/contrast expected with evident | Evaluator and/or stakeholder Interpretations |
| Who is being served? Who is involved? | Who are evident participants? | Expected v. evident, congruent or divergent? | What are the implications? Modify or maintain program? |
| How are participants to be served? | How are participants using the service? | Expected v. evident, congruent or divergent? | What are the implications? Modify or maintain program? |
| What will be produced by participants in the program? | What was produced by participants in the program? | Expected v. evident, congruent or divergent? | What are the implications? Modify or maintain program? |

Dr. LeAnn G. Putney - Complementary Analysis Research Matrix Application adapted from Putney, L. G., Wink, J., & Perkins, P. (2006). Teachers as Researchers: Using the Critical Action Research Matrix Application (CARMA) for Reflexive Classroom Inquiry. Florida Journal of Teacher Education, IX, 23-36.

Appendix B

2012-2013 Teacher Survey Questions

1. Did you feel that your evaluator was biased against you and/or your style and/or your pedagogy? Enter a short response, which could simply be “no.”
2. Is there anything more that you’d like to add about your evaluation?
3. How do you feel when an administrator drops in, unannounced? Answer briefly:
4. Overall, how do you feel about the security of your job? Answer briefly:
5. If you answered at all negatively to the previous question: Has there been any impact upon your physical or mental health, sleep, or with other relationships in your life?
Answer briefly if any:
6. How is communication between you and district’s administrators and/or the school board? Positive? Negative? Explain. (Leave blank if you haven’t had a need to communicate with them.)
7. Enter any additional comments about your feelings about your relationship with administrators:
8. List and describe any other concerns you might have about [the high school in the school district in Northern California]. What are the most significant issues? (e.g. salary, healthcare, structure, governance, communication, testing, PD, needed programs, etc.)
9. Are there any other related concerns that you’d like comment upon that are broader than the [high school in the school district in Northern California].

Appendix C

District PAR Data

| Year | Age | Ethnicity | Gender | Salary Placement |
|---------|-----|------------------|--------|------------------|
| 2012-13 | 66 | White | F | 16.7.13 |
| 2012-13 | 64 | White | F | 16.6.21 |
| 2012-13 | 37 | African-American | M | 17.6.11 |
| 2011-12 | 63 | White | M | 16.7.18 |
| 2011-12 | 65 | White | F | 16.7.12 |
| 2011-12 | 33 | White | M | 16.4.06 |
| 2011-12 | 71 | African-American | F | 16.7.23 |
| 2011-12 | 59 | White | M | 17.5.23 |
| 2011-12 | 33 | White | M | 17.4.08 |
| 2010-11 | 62 | White | M | 16.7.18 |
| 2010-11 | 32 | White | M | 16.4.06 |
| 2010-11 | 35 | African-American | M | 17.6.11 |
| 2010-11 | 37 | White | M | 16.7.10 |
| 2010-11 | 58 | African-American | F | 17.6.15 |
| 2010-11 | 59 | White | F | 17.3.13 |
| 2009-10 | 40 | White | F | 16.7.15 |
| 2009-10 | 60 | Hispanic/Latino | M | 17.4.19 |
| 2009-10 | 59 | White | F | 16.6.19 |
| 2009-10 | 56 | White | M | 17.7.23 |

| | | | | |
|---------|----|------------------|---|---------|
| 2009-10 | 57 | African-American | F | 17.6.14 |
| 2008-09 | 39 | White | F | 16.7.14 |
| 2008-09 | 58 | White | F | 16.6.18 |
| 2008-09 | 55 | White | M | 17.7.23 |
| 2008-09 | 39 | White | M | 16.6.13 |
| 2007-08 | 59 | White | F | 17.7.20 |
| 2007-08 | 53 | White | M | 17.7.20 |
| 2007-08 | 62 | White | F | 17.5.13 |
| 2007-08 | 60 | African-American | F | 16.7.19 |
| 2006-07 | 58 | White | F | 17.7.19 |
| 2006-07 | 52 | White | M | 17.7.19 |
| 2006-07 | 65 | African-American | F | 17.7.23 |
| 2005-06 | 57 | White | F | 16.6.14 |
| 2004-05 | 55 | White | M | 16.6.12 |
| 2004-05 | 60 | White | M | 17.7.09 |
| 2004-05 | 37 | White | M | 16.7.05 |
| 2003-04 | 51 | Hispanic/Latino | M | 16.7.13 |
| 2003-04 | 56 | African-American | M | 17.7.23 |
| 2003-04 | 55 | White | F | 16.6.12 |
| 2002-03 | 54 | White | F | N/A |
| 2002-03 | 56 | African-American | F | N/A |
| 2002-03 | 59 | African-American | F | N/A |

Appendix D

Interview Questions

Individual Interview Questions

1. Tell me about your role at your current school district and the roles you have held through your career as a teacher.
2. What honors/awards/rewards have you received as a teacher in this district or any other district?
3. Tell me out your negative experiences when you have been evaluated in your current district, from the beginning of the year to the end of the year.
 - a. Give me an example of a negative experience.
 - b. What was it like for you when your evaluation experience turned negative?
4. Some people would say that teachers who have negative experiences with evaluation are bad teachers. What would you tell them?
5. Suppose I was a first-year teacher entering the profession and teaching in your school district. What advice would you give me about the evaluation process?
6. What would you describe as the ideal evaluation process?
7. Are you finding that evaluation at your current school is a different experience from what you expected or experienced in the past?
 - a. Tell me more about that experience.

Group Interview Questions

1. [After reading the PAR policy as stated from the California Department of Education]
What has been you experience with evaluation at in the school district?
2. How does your experience with evaluation align with the state PAR policy?

3. Tell us about a time when you had a negative evaluation experience.
4. What similarities or differences do you see in all of our negative experiences with PAR?

Appendix E

Informed Consent

Department of Teaching and Learning

TITLE OF STUDY: Negative Impacts of Teacher Evaluation Reform: The Role of Federal Policy in Student Achievement and Teacher Quality

INVESTIGATOR(S): Allison Smith

For questions or concerns about the study, you may contact Allison Smith at smith957@unlv.nevada.edu.

For questions regarding the rights of research subjects, any complaints or comments regarding the manner in which the study is being conducted, contact the UNLV Office of Research Integrity – Human Subjects at 702-895-2794, toll free at 877-895-2794 or via email at IRB@unlv.edu.

Purpose of the Study

You are invited to participate in a research study. The purpose of these study is to better understand the negative experiences you have had with teacher evaluation.

Participants

You are being asked to participate in the study because you fit this criteria: you are a teacher who has indicated that you have had negative evaluation experiences and/or someone has heard that you have had negative experiences with evaluation.

Procedures

If you volunteer to participate in this study, you will be asked to do the following: participate in a video-recorded individual interview that should take no more than one hour and one group interview with up to five other teachers who have also had negative evaluation experiences in your school district. The group interview should take no more than one hour.

Benefits of Participation

There will not be direct benefits to you as a participant in this study. However, we hope to learn more about the misuse of teacher evaluation to better inform policymakers about teacher evaluation reform.

Risks of Participation

There are risks involved in all research studies. This study may include only minimal risks. You may become uncomfortable answering some of the questions. At that point, you can always choose not to answer the question or end the interview immediately.

Cost/Compensation

There will not be financial cost to you to participate in this study. The study will take up to two hours of your time. You will not be compensated for your time.

Confidentiality

All information gathered in this study will be kept as confidential as possible. No reference will be made in written or oral materials that could link you to this study. All records will be stored in a locked facility at UNLV for 5 years after completion of the study. After the storage time the information gathered will be destroyed and deleted.

Voluntary Participation

Your participation in this study is voluntary. You may refuse to participate in this study or in any part of this study. You may withdraw at any time without prejudice to your relations with UNLV. You are encouraged to ask questions about this study at the beginning or any time during the research study.

Participant Consent:

I have read the above information and agree to participate in this study. I have been able to ask questions about the research study. I am at least 18 years of age. A copy of this form has been given to me.

Signature of Participant

Date

Participant Name (Please Print)

Audio/Video Taping:

I agree to be audio or video taped for the purpose of this research study.

Signature of Participant

Date

Participant Name (Please Print)

Appendix F

Thematic Analysis Tables

The Scarlet P

| Data | Data Source | Notes |
|--|--------------------|---|
| The survey revealed that 33% of respondents did not feel secure about their job. Some did not worry about themselves, but saw their highly esteemed colleagues treated unfairly. There was report of fear for self and others. Of those participants, 80% reported a negative impact upon their physical or mental health, sleep, or with other relationships in their life. As reported in the survey, these negative impacts on their health included high levels of stress, depression, weight gain, despondence, insomnia, hostility toward family, deteriorating relationships, high blood pressure, and overall distraction. | Teacher survey | Isolation Shame Fear Health Relationships |
| “My doctor who I have a long relationship with was disturbed by my level of stress. During my appointment he prescribed anti-depressants and time off from work. I lost at least thirty pounds due to stress. My family was very concerned about me because I was despondent. I take anti-depressants everyday now. I have to unfortunately” | Teacher survey | Shame Health Relationships |
| 52% of respondents reported that they had negative feelings about administrators coming in to their class unannounced. Some respondents explained feeling anxious, nervous, apprehensive, criticized, stressed, and scrutinized. | Teacher survey | Fear Shame Health |
| He stated that everyone’s evaluation marks went down with this new administration and explained that there was an aura of fear around the entire school. | Reggie’s interview | Fear Sin |
| He no longer received marks of ‘distinguished’ and received many marks of ‘unsatisfactory’ that made him fear the infamous referral to PAR. | Reggie’s interview | Fear Sin |
| “PAR would be humiliating.” He was so scared that he offered to do anything so he did not have to go to PAR the following year. He placated administration out of fear of PAR. | Reggie’s interview | Shame Fear |
| One day, when the students had completed all of the advisory assignments, the students requested that Reggie play his guitar. Reggie felt this was a good way to connect with students and inspire their participation and attendance in the advisory course. | Reggie’s interview | Unconventional Sin |
| Tina explained that the consulting teacher did not know why Tina was in PAR because her teaching was so good. Soon the students found out that the district was targeting her and contacted the news to speak out about the injustice. | Tina’s interview | Isolation |

| Data | Data Source | Notes |
|--|------------------|-----------------------------|
| Tina explained the plight of one of her fellow teachers who taught Latin in the district. Tina said, "She is wonderful, oh, she is awesome. She is quitting this year because she got a bad evaluation." Tina accounted for yet another teacher in the district who had taught history for over 30 years and had been a contributor to high school history textbooks. He quit because of the pressure from teacher evaluation. | Tina's interview | Sin Isolation |
| Health was another major issue that arose from the negative experiences with evaluation. During the years she was in PAR she gained 40 pounds and was diagnosed as pre-diabetic. Even now that she is out of PAR Tina explained, "it's still looming in the back of my head, what's going to happen?" | Tina's interview | Health Fear Shame |
| "The one thing that really attracted me to teaching was your ability to be creative and to use different strategies to teach different students at different levels. They have now moved to what we refer to as scripted teaching. There's a script that you're going to follow and this also ties in to Common Core, where teachers have very little autonomy, you have to follow a script" | Tina's interview | Unconventional Sin |
| Tina claimed her pedagogy was a different approach, but it was successful. | Tina's interview | Unconventional Sin |
| The principal specifically had issue with Ana's timing. Ana explained that the principal, "had a specific formula of timing that apparently we were not suppose to deviate from. I was aware of it but I wanted to run a lesson that was right for the children and felt right for me. I was never criticized and in fact was lauded by the reading teachers. They thought I was doing a fine job" | Ana's interview | Unconventional Sin |
| "that [the scolding] didn't show up in my evaluation but you can tell how something like that can kind of shadow you" | Ana's interview | Sin Shame |
| Additionally, Ana discussed natural differences in teaching. She talked about people naturally having different approaches to talking with children. The district did not have an adopted way of delivering curriculum, such as the scripted Open Court model, so individual differences should have been acceptable. | Ana's interview | Unconventional Ambiguous |
| As she approached her decision to retire while going through the year in PAR, she explained that, "I found myself withdrawn to my classroom more and more and just staying with my own cohort in recent years. It was much safer. I tried avoiding the office. It did feel negative" | Ana's interview | Isolation Shame Fear |
| Although she had received high accolades from parents and teachers, she felt completely overlooked with awards from the district. "I feel invisible in terms of praise and I don't know why. But it makes the job harder because you work so hard and you just get criticized and you don't get praised on top of it. You know, it's kind of lonely." | Dawn's interview | Isolation Shame |

| Data | Data Source | Notes |
|---|------------------|----------------------------|
| Dawn incorporates her experience with music and the arts in to her curriculum. | Dawn's interview | Unconventional Sin |
| "I was so traumatized. I felt like I had been ambushed." She felt alone with no support. | Dawn's interview | Isolation |
| "I was completely shocked and stressed and humiliated and embarrassed and I didn't, like, really have a trusting relationship, on a confidential level with people there." | Dawn's interview | Shame Isolation |
| The school's literacy coach saw her teaching in a different way and reported her to the principal. Dawn explained that the principal came to her angrily and expressed her disapproval of teaching cursive and sentence diagraming. She was told she had "old-school" approaches to teaching and she was too traditional. | Dawn's interview | Unconventional Sin |
| "So year two the stress was killing me. I would sit at home and have my heart racing and I would go through my head, there's no way" | Dawn's interview | Shame Fear Health |
| This overload was so much that it seems to be the cause for Reggie, Dawn, and Tina to drop out of the study. | Group interview | Shame Fear Health |
| Tina answered the first question of the group interview and wrote at the end, "I will finish this today. I have to take a break because this brings up too many unpleasant memories. I hope that you understand." | Group interview | Shame Fear Health |
| "I'm just hanging in because it's like, I have to. But the stress is terrible and the fear of what if, it's always on my mind. What if I lose my job and why aren't I appreciated?" | Dawn's interview | Survival Health Shame Fear |

Kangaroo Court

| Data | Data Source | Notes |
|--|----------------|--|
| The California PAR policy placed a focus on local control to meet local conditions. | PAR policy | Unfair practices Ambiguous Subjective |
| Overall, the policy was written with vague language. | PAR policy | Ambiguous Subjective |
| When participants were asked about feelings of bias in their evaluation, 23% reported feeling as though their administrator was biased against them, their style, or their pedagogy. | Teacher Survey | Unfair practices Judgment predetermined Subjective |

| Data | Data Source | Notes |
|---|--------------------|---|
| When participants further explained their evaluation experience, only 55% responded. However, of the 55% who responded, 40% reported having a negative experience. Their responses centered mostly on evaluations lacking impact and meaning and being adversarial, harassing, and intended to find fault. | Teacher survey | Unfair practices Judgment predetermined |
| Respondents explained their feelings about their administrator evaluator with words such as, “criticized”, “uninformed”, and “unqualified.” | Teacher survey | Subjective Unfair practices |
| "decidedly biased against my style and pedagogy and possibly against me personally." | Teacher survey | Subjective Judgment predetermined |
| The following year, after receiving a perfect evaluation, the administration changed. With that change, came a complete change in his evaluations. | Reggie's interview | Subjective Ambiguous |
| Reggie further explained that this fear came from confusion about the evaluation process. He was told for ten years that he was a great teacher and then with a new administration he was given much lower evaluation marks. It was not clear to Reggie why he was not distinguished anymore. | Reggie's interview | Ambiguous Subjective |
| “They have no idea how to make us better teachers. They are clueless. They are in over their head and they are being pushed by the hill people who want it to be a private school. They are being pushed to be harder on the teachers.” | Reggie's interview | Privatization Subjective |
| “PAR process is there to scare people. It's there to try to terminate people. It's there to get the union to work with the administration to manage the teachers.” Reggie tried everything to answer any concerns from administration but always felt that there was a predetermined bias against him. “They have been more harping on test scores and listening to parent complaints than they have been looking and understanding the teaching situation. So they are coming in to the teaching situation already having made up their mind. I don't like this teacher, I'm going to look for anything bad I can find.” | Reggie's interview | Subjective Standardization Union Judgment predetermined Ultior motives |
| Reggie was suggesting that money had an influence on how policy and procedure played out in the school. He continued to describe a weak administration at the school that did not work with teachers to support student success and speak from a uniform front to parents. | Reggie's interview | Privatization Subjective |
| Reggie explained, in a frustrated manner, that his evaluator had no content knowledge and merely focused on procedural aspects of teaching. Reggie also felt that the administrator was not experienced in an administrative role or with proper evaluation procedures. | Reggie's interview | Standardization Subjective |

| Data | Data Source | Notes |
|---|--------------------|--|
| It was as though no matter what Reggie did, they would still find fault in his work as a teacher. | Reggie's interview | Subjective Judgment predetermined |
| The result of the evaluation was that Reggie needed to change procedural items, like closure activities and classroom management techniques of off-task students. His evaluating administrator told him to address off-task students through humiliation. The administrator told Reggie he should compare student achievement and participation publicly. He also insisted that Reggie use the Gradual Release of Responsibility (GRR) model. | Reggie's interview | Standardization Subjective |
| Later that day he received a message to go to the principal's office. At this meeting he was told he did not follow advisory protocol when he played music to engage students. When he asked for direct feedback about how to improve, the principal would not give him directives. At that point, Reggie felt that, "he [principal] just didn't like me." | Reggie's interview | Ambiguous Subjective Unfair practices Judgment predetermined |
| For Reggie, evaluation became less helpful to the teachers and more to fulfill an agenda. | Reggie's interview | Ulterior motives Judgment predetermined |
| He felt that the administration had been, "hijacked by special interest groups making demands" | Reggie's interview | Privatization |
| Reggie explained that the administration was misusing the evaluation process but, "they are following what they are being told from the state - to toughen it up on the teachers." He also briefly mentioned that the union was on a downfall because he did not feel the teachers were supported by their union. Instead, Reggie felt that the union and administration were working together. | Reggie's interview | Union Standardization Subjective Unfair practices |
| For Reggie, he thought, "they [administrators] were coming after me because I'm an older teacher who gets paid more." | Reggie's interview | Ulterior motives Judgment predetermined Subjective |
| For Reggie, he knew administrators, "come in with a negative attitude toward older teachers, teachers of color, toward anybody who's gotten negative complaints from parents." | Reggie's interview | Ulterior motives Judgment predetermined Subjective |
| Tina explained the power shift she noticed within the district. She explained there previously was a strong union and veteran teacher presence in the district. When those teachers retired there was a vacuum that occurred and, "there was no longer a partnership between teachers and administrators." She also explained that the union was dismantled and working with the administration against the interests of the teachers. | Tina's interview | Union Unfair practices |

| Data | Data Source | Notes |
|---|------------------|---|
| Through the evaluation process Tina received marks of proficient or distinguished until a change in administration. At that point her evaluations became a very negative experience for Tina. “So I understood that she [the administrator] was coming in not to help me, not to present, um, constructive criticism, but to find areas in which she could present a negative evaluation in my pedagogy. She, um, rarely had positive comments to make. They were very subjective.” | Tina’s interview | Subjective Ambiguous |
| Unbeknown to Tina, the union had gone to the administration to help them file the negative evaluation and place her in PAR. Tina was faced with a principal who had gone back on his word and a union that had operated behind her back to support the administration rather than a teacher. She was placed in PAR and given a consulting mentor teacher who was a former administrator in the district. | Tina’s interview | Union Unfair practices |
| Despite the improvement the consulting teacher had reported, the panel recommended Tina be placed back in PAR for a second year. “The consulting teacher was shocked when she heard that I was put back in PAR.” | Tina’s interview | Unfair practices Subjective |
| The second year in PAR ended with a recommendation to terminate her employment. The consulting teacher changed her recommendation to the PAR panel and recommended Tina be placed in PAR for a third year. The panel explained their recommendation for Tina’s termination was because she showed the movie Gone with the Wind. She showed this movie in the history course she taught while they were studying the Civil War and Reconstruction. | Tina’s interview | Unfair practices Subjective Judgment predetermined |
| “PAR is a program that is meant to come in and support teachers. The way that we use PAR is not to support teachers but to target teachers for termination or to harass teachers to the point where they quit.” | Tina’s interview | Unfair practices Ulterior motives Judgment predetermined Subjective |
| Tina also explained she was not made aware of or trained in the new expectations. | Tina’s interview | Ambiguous |
| “it’s anything but constructive and you feel like they have already made up their mind before they even come in” | Tina’s interview | Judgment predetermined |
| She felt the PAR panel was out of compliance because the administrators outnumbered the teachers. | Tina’s interview | Unfair practices |
| Ana attributed the subjectivity of evaluation, under the current system, to the abstract policy language. | Ana’s interview | Ambiguous Subjective |

| Data | Data Source | Notes |
|--|-----------------|----------------------|
| "I think that with this evaluation, which would have been in the 2011-2012 school year, that this was with a shift in standards and I would say the standards are written in an abstract language that can shape an evaluation when an administrator, an evaluating administrator, has come in and is essentially doing snap-shots of what goes on in your class. And I think the concept of snap-shot is important because I don't think they see the overall picture. You are to demonstrate certain things and you are to demonstrate them in ways that they think, um, create the evidence that they want specifically" | Ana's interview | Ambiguous Subjective |
| That following year, Ana was placed in PAR because of her timing and classroom management techniques. | Ana's interview | Subjective |
| The coach was not quite sure why Ana was in PAR, "she would just shrug her shoulders because in other districts the people who were getting coaching needed it extremely and I don't think I was that needed in the extreme." | Ana's interview | Subjective |
| Ana indicated that the district did not require a script be followed but clearly had a format of expectations that were not made apparent to teachers until after evaluations were completed. | Ana's interview | Unfair practices |
| She surmised that she was pushed out of the district for financial reasons communicated to her directly and indirectly through policy implementation. Ana explained her systematic understanding, "It makes sense monetarily for the district to expedite getting teachers out who are close to retirement because of the training aspect of Common Core. They would have to train another teacher in delivering CCSS in the district very soon after training the older teacher." Ana also explained a meeting with the principal of her school, "My principal even said at one point, do you know how much this district has spent on in-servicing you? I don't think she meant just me, but all teachers." These experiences contributed to her negative experience with teacher evaluation, "That definitely kind of lowers the ceiling on you or puts a cloud over your head" | Ana's interview | Ulterior motives |
| She explained that the union was there to help but they did not always provide recourse to personal attitudes and bias toward teachers, "There's some union politics that I'm not understanding these days." She also mentioned a change in district administration left her unknowing about the future. | Ana's interview | Union Subjective |
| "it isn't a process we ever dignified completely" | Ana's interview | Subjective |

| Data | Data Source | Notes |
|--|------------------|---|
| Ana continued her discussion with her doubt in evaluators coming in to the classroom and checking off competencies and procedures. She felt that she was being observed through a general evaluation with very specific criteria. For her it felt like the evaluator was just looking for proof to uphold what they had already decided. | Ana's interview | Judgment predetermined Subjective |
| As the union site representative, she had nobody to go to for assistance. Once she was referred to PAR the union stripped her of the position and the president of the union would not engage in conversation. | Dawn's interview | Union Unfair practices |
| The school referred Dawn to PAR for having slow transitions, not intervening when a student with special needs was doodling while Dawn was teaching, and the timing of instruction. The observers did not have conversations with Dawn about the changes needed. Instead, she was written up and referred to PAR. | Dawn's interview | Unfair practices |
| "I tried to get help from the union. The only thing the union will help with on is if it's a timeline thing in the contract" | Dawn's interview | Union Unfair practices |
| Dawn felt the evaluation system was, "abused and misused. I think the evaluations are useless because they are so subjective and because there is nobody to defend a teacher who feels that they have been unfairly treated" | Dawn's interview | Subjective Unfair practices |
| She felt nervous about the process and equated it to The Gulag | Dawn's interview | Unfair practices |
| "It was like an execution thing. It's like 'Prisoner A' can come in now." At that point the mentor teacher explained to the panel how Dawn, the participating teacher, improved (or did not improve) throughout the year in PAR. Dawn's PAR coach was very supportive and reported that Dawn had improved through the year in PAR, as documented in the data she collected and presented to the PAR panel. However, the PAR panel asked the mentor teacher probing questions about Dawn's use of a technique called equity sticks. Dawn used many techniques, but had not used this specific technique using names on popsicle sticks to ensure randomly selecting students. This 'failure' led her to a second year in PAR. | Dawn's interview | Unfair practices Subjective Judgment predetermined |
| "I thought I was going to have a heart attack when I found out that I wasn't exiting and I had to go through this again. I had no idea why because I had jumped through all the hoops. I had done everything." Dawn attributed her confusion about remaining in PAR to the PAR panel's private conversations. She was not sure why there were so many conversations about her performance that did not include her input or allowed her to even hear what was said about her performance. | Dawn's interview | Ambiguous Subjective Unfair practices |

| Data | Data Source | Notes |
|--|------------------|---|
| Dawn was not allowed to have the same mentor coach the second year. She was comfortable with the previous mentor, so this inconsistency between years added stress to the situation. | Dawn's interview | Unfair practices |
| "You're guilty before you go in" | Dawn's interview | Judgment predetermined |
| "I don't believe in the evaluation process. I believe it's a sham. I believe it's a total worthless sham because, um, because of what happened to me and because I can see how they can write anything they want and there's nobody going to listen to you. Nobody going to protect you. Maybe if you get a lawyer, but otherwise they can do anything they want and guess what, if PAR doesn't kill you, you know, and get you out of there, the stress will. | Dawn's interview | Subjective Unfair practices |
| In Tina's response, she explained how she was not supported by the union when she told the union president that the principal had decided to "throw out the evaluation." She explained that, "I asked them to stay out of it and they went behind my back and helped my evaluator so that when I returned to work I was placed in PAR, and told by the union VP that it would be good for me." Instead of the union following the wishes of the union member, they pushed the PAR agenda. Tina explained, "I felt totally betrayed." | Group interview | Union Unfair practices |
| Alex stated, "I do think that it [PAR policy] conflicted with our contract which states that the purpose of evaluations is to improve a teacher's practice." Alex explained that his experience was not supportive, "there was nothing helpful offered in the evaluation, and it caused me to feel anxious." Like Tina, Alex felt that the union was not supporting its membership. | Group interview | Union Unfair practices |
| Alex stated, "I believe that PAR came about as an instance of collaboration between unions and districts." This cooperation between unions and districts was a mystery to Alex. Alex furthered this statement by stating, "teachers' unions are overdue for an investigation as to why on a statewide if not nationwide basis, they began promoting PAR and healthcare caps to locals at around the same time." His overall feelings were that the PAR policy, healthcare caps, and other "anti-teacher policies" were disastrous and divisive of teachers. One of the major foundational issues that Alex emphasized in the group interview was that he was present when the district adopted the PAR program for evaluation, nearly twelve years ago. He stated that "PAR was not presented by union leaders as a program for improving teaching, as many now claim it to be." He felt that PAR was actually, "a superior pathway to dismissing bad teachers." | Group interview | Union Unfair practices Ulterior motives |

| Data | Data Source | Notes |
|--|-----------------|--|
| Ana felt that she was negatively evaluated simply because she was a year away from retirement age. Additionally, Ana stated, "I wondered also whether the principle had been given directives by the district offices." Ana surmised the reason she was targeted was for financial investment reasons, especially with the mass implementation of the Common Core State Standards (CCSS). She explained that since she was in her mid-sixties, and if she did not retire, the district would have to pay for her to be trained on the CCSS's. She would only use this training for a couple years because she was at retirement age and the district would need to train her replacement. Ana's responses brought forth a focus on a master narrative that suggested the timing of teacher evaluation reform and the introduction of CCSS's was not by chance. Instead, teacher evaluation reform appeared to support the implementation of the CCSS's. Ana did feel that the PAR policy was consistent, but wondered about the local interpretation aspect. | Group interview | Subjective Unfair practices Ulterior motives |
| All respondents reported inconsistencies and unclear expectations in their evaluations. Reggie completely dropped out after the individual interview and stated being tired of, "administrators that keep changing", and that he has to, "jump through a new set of inconsistencies all over again" | Group interview | Ambiguous Subjective |
| Tina concurred with Reggie when she described her negative evaluation experience. She stated, "my evaluator came in with the intent of not providing constructive criticism, but rather subjectively citing petty scripted parts of the lesson from the administration that I was unaware of" | Group interview | Subjective Ambiguous Unfair practices |
| "I felt that he was paving the way for multiple 'improvement needed' at the next evaluation." This was foreshadowing an eventual referral to PAR | Group interview | Subjective Unfair practices |
| Ana agreed with Alex about the concern about all the standards being interconnected and rated on each one together rather than individually. Ana stated, "I felt this bias, too" and she, "felt strangled by this perception which left little space to clear individual hurdles." Ana also discussed the way standards were written abstractly, which led to various interpretations by local evaluators. Ana explained that she, "wondered how subjectively they were being interpreted by my principal" | Group interview | Subjective Ambiguous |
| Reggie did not respond to the questions, but emailed me directly explaining, "I am very busy and tired of fighting the 'unnamed' administrators that keep changing every time the water gets hot. Then we have to jump through a new set of inconsistencies all over again." The participant did not respond to any further emails and dropped out of the study. | Group interview | Unfair practices |

Salem Witch Trials

| Data | Data Source | Notes |
|---|--------------------|--|
| The data revealed that of the teachers referred to PAR, from 2002-2012, 24% were African American. Within the entire district, African Americans accounted for 6.8% of the total number of teachers. At the high school where the participants taught, 5% of teachers were African American. This data was further analyzed by using a binomial probability analysis to compute the probability that 24% of teachers in PAR would be African American. This analysis indicated a .022% probability that 10 out of 41 teachers referred to PAR would be African American if chosen randomly from a pool of 6.8% African American teachers. | District data | Bias |
| The highest column on the pay scale was seven, for those teachers who had a Bachelor of Arts degree with 84 additional units or a Masters degree with 36 additional units, or a doctorate degree. The demographic data showed that 35 of the 41 teachers placed in PAR were in column six or seven on the pay scale. The average step placement of all teachers in PAR was 15 years experience. The most highly educated and most experienced teachers made up 80% of all teachers referred to PAR. | District data | Bias |
| The next demographic indicator was sex and age of the teachers referred to PAR. Of the 21 women placed in PAR, 19 women were over the age of 55. The women teachers had more years of experience than their male counterparts who were placed in PAR. Another binomial distribution analysis was conducted on the probability that 19 out of 21 female teachers would be over the age of 55 was a random incident. The analysis indicated that the probability that this happened at random was .025%. | District data | Bias |
| Some did not worry about themselves, but saw their highly esteemed colleagues treated unfairly. | Teacher survey | Bias |
| “Other teachers look around and they go ‘I know that teacher. That’s a good teacher and they are being harassed by the administration who doesn’t really know what they are doing. So that means that could happen to me’. And so everyone has gotten in to a state of panic looking around at these negative evaluations that lead to potential termination or harassment to the point where teachers don’t want to work there anymore” | Reggie’s interview | False accusations Mass hysteria |
| Reggie referred to times when the administration met with parents behind the teachers’ backs, which was against the contract. He also felt that the public school was being turned in to a private school, which inherently created a segregated school. | Reggie’s interview | Privatization Segregation Lapse in due-process |

| Data | Data Source | Notes |
|---|--------------------|--|
| After this evaluation Reggie wrote a rebuttal, which is a teacher's right within performance evaluation. However, Reggie stated that the evaluator refused to consider the points that he had made about the evaluation. He was told the evaluation was done and it would not be updated to reflect anything stated in the rebuttal. | Reggie's interview | Lapse in due-process Unfair practices |
| To protect himself, he placated the administration and advocated for himself to various other leaders in the school to speak up and help him. After all, he was a veteran teacher who was also chosen as a mentor teacher for the BTSA mentorship program for three teachers that year. These were conflicting messages about performance. | Reggie's interview | False accusations Survival |
| "The term 'bad teacher' has become the national mantra and I think most people have no idea what that means." He explained that sometimes a teacher is liked merely because they are "bubbly and kids like them" and they get a good evaluation for that reason. However, if a teacher knows the subject matter but is not the likable type among students and parents, then they may get a poor evaluation. It essentially came down to a popularity contest. Reggie explained that, "the media has also heightened the few stories of bad teachers in the nation as a description of all teachers." | Reggie's interview | Mass hysteria False accusations Bias |
| He explained he knew he had to play the game and give the evaluator all the procedural steps of teaching very explicitly when he was being evaluated. It was as if he was putting on a show that addressed all of the things that Reggie guessed would be important in the evaluation, in an overly expressed manner. | Reggie's interview | Survival Mass hysteria |
| The school was vulcanized in to small schools and served as a way to segregate the school. This segregation took away power and created small schools of color within a larger school. Tina felt the district chose to, "divide and conquer the power of teachers" in this shift of power. | Tina's interview | Segregation Isolationism Mass hysteria |
| Tina's evaluation was conducted but not debriefed with her and filed with the district until after the evaluation deadline, as outlined in the teachers' contract. She spoke with the principal about the missed deadline and instructed the union to stay out of the issue because she had handled it herself. | Tina's interview | Lack of due-process |
| She was told the reasons she was placed in PAR were because she did not have a two-minute warm-up, objectives and goals were not written on the board, and she did not have an exit ticket activity. Tina explained she had never heard the terms that landed her in PAR. She did the components of those items, but she was not using them in the explicit fashion that newer teachers directly out of credentialing programs used them. | Tina's interview | False accusations Bias |

| Data | Data Source | Notes |
|--|------------------|--|
| Tina explained that the consulting teacher told her, “this has nothing to do with your pedagogy, this is all political. You met the standards. I watched you meet the standards. I was there 11 times. I am angry and I am in shock. This has nothing to do with your teaching.” | Tina’s interview | False accusations Bias |
| “nobody else had to do this, only me. I’m the only teacher at the school that has to do this.” | Tina’s interview | Bias |
| “I looked things over and I just said, you know this process is, I had a wonderful class last year, and I said I can’t take this stuff anymore. I can’t take the negativity anymore. And so I just said I don’t have to do this anymore” | Ana’s interview | Bias Survival |
| “We knew, the three teachers, three second grade teachers knew at the very beginning of the year that we needed help and support from the administration on this and we were told, ‘well you are the classroom teacher. You have to take care of this. This is your job’. Well it was each classroom teacher's job but sometimes there are situations that need more than the support of the classroom teacher and we all had a difficult year that year. Another teacher had a negative evaluation at that time and that teacher decided to retire that year and I was not ready to do that, so I didn't” | Ana’s interview | Survival Isolationism Lapse in due-process |
| Ana felt that her principal just had something against her. No matter what she tried to do, the principal would continue to target her. | Ana’s interview | Bias |
| Overall, her experience with PAR and evaluation leading up to PAR was punitive, negative, and ruled by fear. "It was like a cop giving tickets. You did this and I'm putting it in your file." It was a police state where power was exercised arbitrarily. | Ana’s interview | Mass hysteria Lapse in due-process False accusations |
| The appearance was of utmost importance, rather than a focus on quality teaching. “Your class needs to look like it's managed well, where children look like they are paying attention. Children need to look like they are engaged in their work. Needs to look like you have a structure during individual work times that accommodates any special needs. Class needs to look like it's a comfortable place for them to learn” | Ana’s interview | Survival |
| Dawn utilized her right to submit a rebuttal to the evaluation. “I spent hours writing a finely word-smithed, very revised, very factual rebuttal. Nobody read them. I spent hours writing very factual clear rebuttals.” Dawn explained that the process of rebutting evaluation was her right, but not actually utilized. Dawn felt that the PAR evaluation system was to, “get rid of dead wood and start a paper trail in the evaluation process” | Dawn’s interview | Unfair practices Lapse in due-process |

| Data | Data Source | Notes |
|---|-----------------|-------------------|
| When parents and reading specialists found out she had negative evaluations, they were surprised. Her experience was an inconsistency between perceptions of performance. | Group interview | False accusations |

Nonconfirming Data

| Data | Data Source | Notes |
|--|--------------------|--|
| Alternatively, many respondents were satisfied and happy with their evaluation process. They liked their administrator coming in to their classroom and giving them feedback. | Teacher survey | Positive |
| Some respondents reported discontent with the survey because they did not see the connection between performance evaluation and health. Some faulted the question construction and claimed inherent bias in the survey as a whole. | Teacher survey | Majority of respondents report no link between health and evaluation |
| “Are you really trying to link people’s personal health issues to being evaluated as a professional? What about if I answered positively to the previous question and am experiencing excellent health, sleeping great, and am newly in love? Is that somehow connected to my evaluation or feelings of job security? Clearly you are just trying to collect quotes to support your agenda and this ‘survey’ is intellectually dishonest. I expect that if you report the results of this ‘survey’ you will include all results. Right?” | Teacher survey | Majority of respondents report no link between health and evaluation |
| The only time Reggie remembered receiving useful evaluation feedback was when the evaluation was done by other math teachers. | Reggie’s interview | Peer evaluation helpful |
| Ana explained that the PAR coach was very supportive and nice to work with for the year. They had good conversations and the coach provided great feedback. | Ana’s interview | Peer evaluation helpful |
| Dawn was happy to receive feedback and simply wanted to have conversations with the evaluator and work together toward improvement | Dawn’s interview | Positive |

| Data | Data Source | Notes |
|--|-----------------|-------------------------|
| Tina explained that her evaluation experience was positive before the implementation of PAR. Tina did not have a negative experience with evaluation in general, it was specifically the PAR program. She also expressed a cooperative relationship between the principal and herself. | Group interview | Positive Cooperation |
| “The PAR process was positive because of my coach. She observed me eleven times compared with the three times that my evaluator observed me” | Group interview | Peer evaluation helpful |
| Additionally, in Ana’s interview, she stated that “It [PAR] seemed consistent with state guidelines” | Ana’s interview | PAR policy good |

Appendix G

CARMA Worksheets

| Program Expectations Administrators and/or support staff | NoteTaking Describe what was intended by program administrators and support staff |
|---|---|
| Identify the program managers and staff. | PAR: Panel of teachers and administrators. Must be mostly teachers. Panel chosen by district. Mentor teachers chosen by panel based on: Proven to be good communicator, skilled teachers, and worked in the district. Mentor teacher makes recommendation about participating teacher to panel. Panel recommends their suggestion to retain or terminate teacher to school board. The panel also has union president on the panel, who is de facto leader of the panel. |
| Who is intended to be served? | Teachers who are rated as unsatisfactory in their performance evaluation are referred to PAR for the following year to learn ways to teach in better ways. The district is also able to have a more accurate performance evaluation of teachers. Students receive a better education. |
| How are participants to be served? | Unsatisfactory teachers are referred to PAR to improve teaching or be dismissed after a year in the program. This is to improve student achievement. Mentor teacher does multiple observations and makes suggestions to participating teacher about ways to improve their teaching. At the end of the year the mentor makes a suggestion to the panel, who later reports to the school board to retain or terminate the teacher. |

| | |
|---|--|
| What will be produced by participants in the program? | Participants will either become better teachers, who better support student achievement, or they are terminated. Participants will learn more effective pedagogy and collaborate with their mentor and administration. |
|---|--|

| Evident Implementation Users/participants | NoteTaking Describe what is evidently happening in the program |
|--|---|
| Identify the demographics of the population served | Mostly women over 55. Mostly African American Mostly on higher pay scale - more education and years experience |
| Who are evident participants? | More creative teachers, using unconventional methods Less conforming Highly involved in the school - committees, after-school |
| How are participants using the service? | Seem to enter in to PAR negatively and fearfully Don't report improving teaching after PAR They just try to get through so they don't get fired, survival Feel anxiety and fear, deteriorating health Never know what to expect Some like collaborating with mentor teacher Like having more observations |

| | |
|---|--|
| What was produced by participants in the program? | <p>More anxiety, fear, and health problems</p> <p>PTSD</p> <p>Felt like kangaroo court and witch hunt</p> <p>4 of the 6 teachers are still teaching - scared though</p> <p>2 of the 6 teachers retired after their experience in PAR</p> <p>Participants felt isolated and unsupported by union</p> <p>Decisions seemed to be made regardless of performance</p> |
|---|--|

| Results Degree of congruence or divergence | NoteMaking Compare/contrast expectations with evident implementation. |
|---|--|
| Who are participants? | <p>Disproportionate number of older women, African American, and experienced teachers. Does not match total population. Appears to be disparate impact discrimination.</p> <p>Participants interviewed all seemed to be more creative and use unconventional teaching methods. More student-centered. All were passionate about teaching and learning. Most didn't like the CCSS movement. Didn't feel the union supported them.</p> |
| How are participants served? | <p>Although intent is to support teacher development toward satisfactory teaching, participants felt it was more punitive and targeted. Teachers did not feel they were better because of the PAR experience. Their health was deteriorating from stress and they felt they couldn't serve students to the best of their abilities under the stress of PAR. Teachers felt that decisions were decided regardless of their progression through PAR. Was not clear if PAR evaluation was accurate measure of teacher quality for the district. Not clear if teacher evaluation connected with student achievement.</p> |

| | |
|--|--|
| What has been produced or what are the outcomes? | There were negative impacts of teacher evaluation. Teachers interviewed cared about their students and the educational process. They were highly involved in school events and committees. These teachers are/were turned away and disheartened after their negative experience with teacher evaluation. Isolation. Two teachers chose to retire earlier than they intended to because of the stress of PAR. Those who continued to teach after PAR were still fearful and disheartened. |
|--|--|

| Conclusions | NoteMaking |
|--|---|
| Evaluator Interpretations | Implications for participants |
| What are the implications for who is being served? | PAR is discriminatory. Policy needs to be written more clearly without ambiguity. Ambiguous policy language may lead to misuse at local level. Commonality of participants being unconventional teachers may suggest PAR upholds ulterior motives: CCSS, union busting, standardization, and privatization. |

| | |
|---|--|
| <p>What are the implications for how they are being served?</p> | <p>PAR program created fear and punitive atmosphere instead of support and improved pedagogy. Those not in PAR feared referral to PAR. Overall the goal of improving pedagogy did not seem to be happening. Some participants liked the cooperation with mentor teachers, but did not like PAR. Outcomes seemed to diverge from initial intent to improve teaching and student achievement.</p> |
| <p>What are the implications for the outcomes?</p> | <p>The PAR program can be discriminatory and does not appear to support improved student achievement. Teachers had extremely negative experiences with PAR, even if not referred yet. PAR policy was written vaguely which allowed for local control. This has turned in to unethical usage to terminate employment of some teachers. The data showed a disproportionate referral of African Americans, veteran teachers, and women. The reasons participants said they were referred to PAR seemed like professional development could address the issues. Ulterior motives seem to be guiding the implementation of PAR.</p> |

Curriculum Vitae

EDUCATION

- Ph.D. *University of Nevada, Las Vegas*
Department of Teaching & Learning
Dissertation: *Negative Impacts of Teacher Evaluation: Federal Policy Impacts in Student Achievement and Teacher Quality*
Chair: Dr. Christine Clark (senior scholar in multicultural education)
May, 2015
- M.Ed. *California Lutheran University, Department of Education*
Thesis: *Methods Used to Integrate English Language Learners.*
December, 2007
- Teaching Credential *California Lutheran University*
Clear Single Subject Teaching Credential
Certified in Cross-cultural Language and Academic Development
Single Subject: Physical Education
Supplementary Credential: Computer Concepts and Applications
December, 2006
- B.A. *University of California, Santa Barbara*
Major: Communication
Minor: Physical Education/Sports Management
June, 2002

RESEARCH INTERESTS

Education policy, urban education, multicultural education, online education/technology integration, international comparative studies

PROFESSIONAL WORK EXPERIENCE

- Aug. 2013-May 2015 Online and Face-to-Face Instructor, Department of Teaching and Learning - University of Nevada, Las Vegas
- Sept. 2014-May 2015 Board of Directors and Grant Writer - AIP Inc. (Non-profit)
Supporting high-risk youth through performing arts programs
- Feb. 2013-May 2015 International Consultant - Oficina de Negocinhos
Rio de Janeiro, Brazil
Social responsibility through youth entrepreneurship courses
Culturally responsive curriculum development
Instructional and assessment design for multi-week courses
Instructional design for train-the-trainer courses

- Aug. 2011-Sept. 2014 Meeting Facilitator for the Anti-Defamation League
Anti-bias education promoting justice and fair treatment of all peoples - speaking to groups from 30-400 people
- Sept. 2013-Sept. 2014 Tutor - Zoom Schools of Las Vegas (grades 1-3)
English Language Development targeting Limited English Proficient Students
- June 2012-June 2013 Athletic Director - Berkeley High School, Berkeley, CA
Largest athletic department in western U.S.
Focused on a technology-enhanced athletic department through social media, organizational structures, coach education courses, athletic recruitment, website development
Urban school with a 64% minority student population
- Aug. 2010-May 2012 Online and Face-to-Face Instructor, Department of Curriculum and Instruction
University of Nevada, Las Vegas
- Aug. 2011-Dec. 2011 Site Facilitator - University of Nevada, Las Vegas
Mentoring preservice teachers in urban secondary schools
- Jan. 2011-May 2011 Hybrid Course Instructor (Online and Face-to-Face),
School of Education, Nevada State College
52% minority student population
- Aug. 2007-Aug. 2010 Teacher - Berkeley High School, Berkeley, CA
64% minority population
Physical Education Teacher: Team Sports, Soccer, Grades 9-12
Physical Education Department Head, 2009-2010
Executive Board of the Berkeley Union, 2008-2010
- Aug. 2007-June 2009 Teacher Technology Representative- Berkeley High, Berkeley, CA
Technology support for a staff and student population of 3,500
Urban school with a 64% minority student population
- Mar. 2007-May 2007 English Language Development Science Academy Teacher
Flamson Middle School, Paso Robles, CA
Migrant Program-empowering the children of migrant workers through education
- June 2007-July 2007 Summer School Teacher - Health
Paso Robles High School, Paso Robles, CA

Jan. 2007-June 2007 Substitute Teacher - Paso Robles School District, Paso Robles, CA
All subjects, all grade levels

Sept. 2006-Dec. 2006 Substitute and Student Teaching - Ventura High, Ventura, CA
53% minority student population
Physical Education, Health, and History

Jan. 2006-May 2006 Substitute & Student Teaching-Haydock Intermediate, Oxnard, CA
98% minority student population
Physical Education

July 2002-May 2005 Accounting
Brooks Institute of Photography

PUBLICATIONS

Smith, A., Crothwaite, J., & Clark, C. (2014). Institutional racism. In S. Thompson (Ed.), *The encyclopedia of diversity and social justice: Volume 2* (pp. 413-6). Lanham, MD: Rowman & Littlefield.

Smith, A., Crothwaite, J., & Clark, C. (2014). White privilege. In S. Thompson (Ed.), *The encyclopedia of diversity and social justice: Volume 2* (pp. 743-4). Lanham, MD: Rowman & Littlefield.

HONORS

Fulbright Scholarship Finalist, 2012

UNLV Graduate College Research Forum Outstanding Presentation Award, 2014

PRESENTATIONS

Peer-Reviewed International & National Conference Presentations:

Hawaii International Conference on Education, January 2014

Negative Impacts of Teacher Evaluation: The Role of Federal Policy on Student Achievement and Teacher Quality

American Educational Research Association, April 2012

Advancing Understanding of Pedagogies used in Preservice Teacher Preparation

Association of Teacher Educators Conference, August 2011

Negotiating Identity in American Schools: Looking at the Intersectionality of Identity Negotiation among Minority Students and Queer Teachers.

Association of Teacher Educators Conference, August 2011

Negotiating Identity: How Teachers Explore their Authentic Pedagogy to Combat Institutional Heteronormativity and Create Authentic Classrooms.

Guest Lectures:

EDSC 323: Secondary Teaching Methods, UNLV, February 15, 2011

Topic: Culturally Responsive Teaching in Urban Schools

EDU 201: Introduction to Elementary Ed., UNLV, April 26, 2011

Topic: Classroom Management

SPONSORED PROJECTS

\$5,000 Graduate and Professional Student Association, UNLV

Grants to present at education conferences, 2011-2014

\$1,700 Berkeley Public Education Fund

Improving achievement through community-building opportunities for high-risk students, 2007

\$1,000 Berkeley Schools Excellence Program

Technology Enhanced Physical Education, 2008

\$1,700 Berkeley Schools Excellence Program

Improving achievement through community-building opportunities for high-risk students, 2009

\$2,000 Teacher Initiated Professional Development

Peer Observation Professional Development Program, 2009

CONTRIBUTIONS TO THE PROFESSION

Founding Member of DREC: Doctoral Research and Education Collaboration

Mentorship and collaborative research initiatives to support new doctoral students. Supporting the teaching profession through collaboration with faculty and students - Department of Teaching & Learning, UNLV.

CONFERENCES AND TRAININGS

Hawaii International Conference on Education. January 2014

Presentation: How federal policy negatively impacts urban schools and minority teachers through teacher evaluation reform

American Educational Research Association. April 2012

Presentation: Advancing pedagogy methods in teacher education

Association of Teacher Educators. August 2011

Presentation: Identity negotiation for minority students and teachers (two papers)

Dr. Sharroky Hollie

The Center for Culturally Responsive Teaching and Learning focused on encouraging schools to infuse the concept of culturally responsive pedagogy and utilize its instructional methodology in all areas. 2007

BAYCES

Bay Area Coalition for Equitable Schools focuses on improving educational experiences and outcomes for under-served youth. 2009

2020 Vision: An All-City Equity Task Force

A call to action to make educational success, equity and positive outcomes for all of Berkeley's children and youth a community-wide priority, by removing barriers to learning and providing opportunities for all Berkeley children to succeed. 2009

Teaching for Social Justice: Foundations for Change

Developing empowering learning environments, more equitable access to resources and power, and realizing a just and caring culture. Specialized training on how neighborhoods influence student habits and effectively working with high-risk students. 2009

PROFESSIONAL ORGANIZATIONS

American Educational Research Association, Division K
Association of Teacher Educators
National Association for Multicultural Education

UNIVERSITY COURSES TAUGHT

University of Nevada, Las Vegas

Teaching and Learning in Secondary Education (undergraduate) EDSC 323
Fall 2014, Spring 2015
Secondary Methods Practicum 1 (undergraduate) EDSC 311
Fall 2014, Spring 2015
Introduction to Secondary Education Online Course (undergraduate) EDU 202
Spring 2014, Fall 2013, Spring 2013 (2 sections)
Introduction to Elementary Education Online Course (undergraduate) EDU 201
Spring 2012
Classroom Management Online Course (undergraduate) EDSC 408
Fall 2010; Spring 2011; Fall 2011, Spring 2012
Classroom Management Face-to-Face Course (undergraduate) EDSC 408
Fall 2010; Spring 2011; Fall 2011; Fall 2013

Nevada State College

Foundations in Education: Elementary and Secondary (undergraduate) EDU 250
Spring 2011