

Article

The Counternarrative of Teacher Evaluation: The Kangaroo Court, the Salem Witch Trials, and the Scarlett Letter

Allison Smith

Sanford College of Education, National University, 11355 N Torrey Pines Rd, La Jolla, CA 92037, USA; asmith5@nu.edu

Received: 10 April 2019; Accepted: 23 May 2019; Published: 18 June 2019



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. With a paucity of previous research focused on the negative impacts of teacher evaluation, this study addressed the following questions: (1) *How does the Peer Assistance and Review (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?* and, (3) *How does the intended use of the PAR teacher evaluation process compare to the evident use of PAR?* 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. Vague policy language was suggested as the impetus for misuse, abuse, and biased implementation at the local level. This study suggests that policymakers 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: teacher evaluation; Peer Assistance and Review; education policy reform; disproportionate impact; race; gender; age

1. Introduction

Typically, teacher evaluation research either takes a neutral approach or focuses on the positive aspects and outcomes of teacher evaluation [1,2]. This conventional approach has led to minimal findings centered on the negative effects of teacher evaluation. When teachers have a negative experience with teacher evaluation processes or outcomes, it can be quickly assumed that they are ineffective teachers. This study explored the counternarrative of teacher evaluation by taking a deep look at the negative effects of teacher evaluation through a mixed-methods approach, with a specific focus on the teacher evaluation method of Peer Assistance and Review (PAR).

This study elucidated negative evaluation experiences and juxtaposed the intended use of evaluation policy with the evident experiences. Four profiles of teachers are shared through storytelling to better understand their negative experiences in teacher evaluation. PAR data and policy are also analyzed to better understand the context of their experiences. The Peer Assistance and Review (PAR) evaluation method was central to this study because teacher evaluation through peer review was favored in federal and state policy reform [3–5].

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? and, (3) How does the intended use of the PAR teacher evaluation process compare to the evident use of PAR?

Cochran-Smith and Lytle [6] 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” [7] (p. 285). Teachers who had negative experiences with teacher evaluation remained marginalized through the research, over many years, with scant research. This study addresses this divergence by specifically looking at negative experiences of evaluation through a sequential transformative design [8] that was framed by critical multiculturalism. This methodology and theoretical framework, as well as the data collection and analysis tool, synchronistically aimed to empower participants and inform policy makers.

2. Literature Review

With teacher evaluation policy encouraging a peer evaluation approach, it was expected that there would be a large body of research to inform policy changes. However, 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, focusing on peer evaluation, or analyzing the negative experiences of teachers. A very limited amount of studies even mentioned the negative experiences teachers had with teacher evaluation. The literature to review was extremely limited.

Among the research noting the negative impacts of an increased focus on teacher evaluation, Donaldson [9] conducted a small-scale study to help inform policymakers in the debate about teacher evaluation. Donaldson 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 negative experience with new teacher evaluation methods according to district surveys.

The majority of the findings were negative and the study was based on Donaldson’s [9] 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 found major deficiencies in teacher evaluation as a means to improve student achievement.

Hill and Grossman [10] critically revealed inadequacies of 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.

Markow, Macia, and Lee [11] 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 1000 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.

Youngs [12] conducted a study that found teacher evaluation to be a key component for the successful implementation of the Common Core standards and assessments. Youngs stated that, “past attempts to enact standards-based reform have been impeded by limitations in teacher evaluation” (p. 1). Youngs further explained ways that teacher evaluation can positively support the enactment of Common Core standards and assessments. This research appeared to support a Common Core State Standards (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 [13]. Mead 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 [13] 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. Researching the negative impacts of teacher evaluation builds on this body of research by addressing the topic at a deeper level, through a critical multiculturalism framework and transformative methodology. It answers Mead’s call and builds upon the existing literature to further look at potential mistakes in implementing evaluation systems.

In an article by Darling-Hammond, Wise, and Pease [7], “Teacher Evaluation in the Organizational Context: A Review of the Literature,” they wrote about the decade prior to 1983 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, Wise, and Pease’s report, and on conclusions from other leading teacher education and multicultural education scholars about policy impact on teacher evaluation [10,14], 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, by looking specifically at the negative experiences of teacher evaluation.

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 [1]. 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 has not been understood because it had not been the focus of prior research. Through a critical multiculturalist framework, this issue was directly addressed in this study.

3. Methodology

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. The Complementary Analysis Research Matrix Application [15] was an orienting tool designed to juxtapose data collected in interviews with public data sources and was utilized for this study. Triangulation of these data sources increased reliability and provided depth to the findings. CARMA was utilized throughout data collection and analysis, with the aim to empower participants to share their experiences and “use the research to improve their educational space” (p. 131).

The CARMA matrix [15] consists of four phases (see Table 1, below). The first phase provided the existing state and district policy to better understand the intent of PAR. In phase two, district data of teachers referred to PAR over a ten-year period was used to evaluate and explore the evident use of the PAR policy. Evident use was also explored in four individual interviews. The third column of CARMA was, “the point where NoteTaking changes to NoteMaking, and the interpretation of the data is made” (p. 2). This phase juxtaposed the intended use of PAR, noted in PAR policy, with the evident use of PAR, noted in district data and individual interviews. Congruence or divergence between notes in the first two phases was noted in the third phase. The final phase consisted of conclusions and recommendations based on the analysis of the first three phases. 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 [16].

Table 1. 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 and Recommendations
Methodology: Sequential Transformative	Phase 1	Phase 2	Interpretation	Results
Data Source Collection and Analysis	PAR Policy	District Data Individual Interviews (n = 4)	Juxtapose notes from Phase 1 and Phase 2 Co-construct results with participants	Co-construct the recommendations to modify or maintain program with participants

Initially, participants were recruited through purposeful unique sampling targeting 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 [16]. The participants referred other teachers they knew who had negative experiences with teacher evaluation.

District Data

In the district data set, four variable categories were chosen for analysis based on their historically marginalized status and their potentially disproportionate representation in PAR: African-American teachers, female teachers, teachers over the age of 55, and pay scale classification of teachers. 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. This showed the degree of randomness that teachers within each variable were referred to PAR. 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 [17].

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 master’s degree with 36 additional units, or a doctorate degree.

The final demographic indicators analyzed were sex and age. The district-level age data were not made available, despite multiple requests under the Freedom of Information Act [18]. State-level data were used in lieu of hidden/missing data sets about age.

Initial analysis consisted of open coding of data. Then open codes were grouped 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. Each data set was analyzed individually and then coding was merged to capture recurring patterns across all quantitative and qualitative data [16].

4. Findings

The intended use of PAR was established by reviewing the PAR policy. Peer Assistance and Review (PAR) was a program for teachers who received ‘unsatisfactory’ or ‘needs improvement’ marks

on their performance evaluation. After receiving a low evaluation, teachers were 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 [19]. The major finding, upon analysis, was that verbiage in the PAR policy were overwhelmingly nonspecific (see Table 2, below), leaving many areas open for varying usage (or misuse).

Table 2. Education Code Analysis.

Education Code	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?
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”?
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 be limited to, interviews or surveys of the program participants	Why is the program evaluation component not explicit and statewide?

4.1. District Data

The district dataset revealed a disproportionate number of teachers of each variable in PAR, thus negatively impacting specific groups of teachers. Of the teachers referred to PAR, from 2002 to 2012, 24% were African American. Within the entire district, African Americans accounted for 6.8% of the total number of teachers [19,20]. 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 0.027% 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 pay level demographic data showed that 35 of the 41 teachers placed in PAR were in column six or seven on the pay scale, the highest salaries for teachers. 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 female 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 over the age of 55 was a random incident. The analysis indicated that the probability that this happened at random was 0.00000013% (see Table 3, below). The state-level demographics on age indicate that teachers older than 55 represented 21.5% of the state teacher population. In this district, 90% of women placed in PAR are over the age of 55.

Table 3. Binomial Probability.

Variable	African American	Women Aged 55+
Probability of Success		
Percentage of variable within the school district	6.8%	21.5% *
Number of Trials		
Total number of teachers referred to PAR	41	41
Number of Successes		
Teachers within the variable referred to PAR	10	19
Binomial Probability		
Probability that the number of successes happened at random or by chance	0.027%	0.00000013%

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

4.2. Individual Interview: Reggie

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. 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. This year, Reggie was appointed to mentor three teachers in the state-mandated Beginning Teachers Support and Assessment (BTSA) induction program that all new teachers must complete to receive their full teaching credential.

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.

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."

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.

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

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 induction program for three teachers that year. He pondered how he could 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." 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 that 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." For Reggie, evaluation became less helpful to the teachers and more to fulfill an agenda.

4.3. Individual Interview: Tina

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 minoritized backgrounds, prepare for entrance to universities [21]. Tina also participated in other programs that supported college readiness for Students of Color.

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.

Tina was 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.

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 that their recommendation for Tina's termination was because she showed the movie *Gone with the Wind* [22]. 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.

Tina also explained that 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.

4.4. Individual Interview: Ana

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.

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.

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.

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 that a script be followed but clearly had a format of expectations that were not made apparent to teachers until after evaluations were completed.

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.

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 that a change in district administration left her unknowing about the future.

Ana suggested that 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.

4.5. Individual Interview: Dawn

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.

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.” 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 diagramming. 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.

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.

“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?

Dawn suggested that 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.

4.6. Emergent Themes in the Findings

As the data were collected and analyzed, three major themes emerged (see Figure 1, below). 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* [23]. 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.

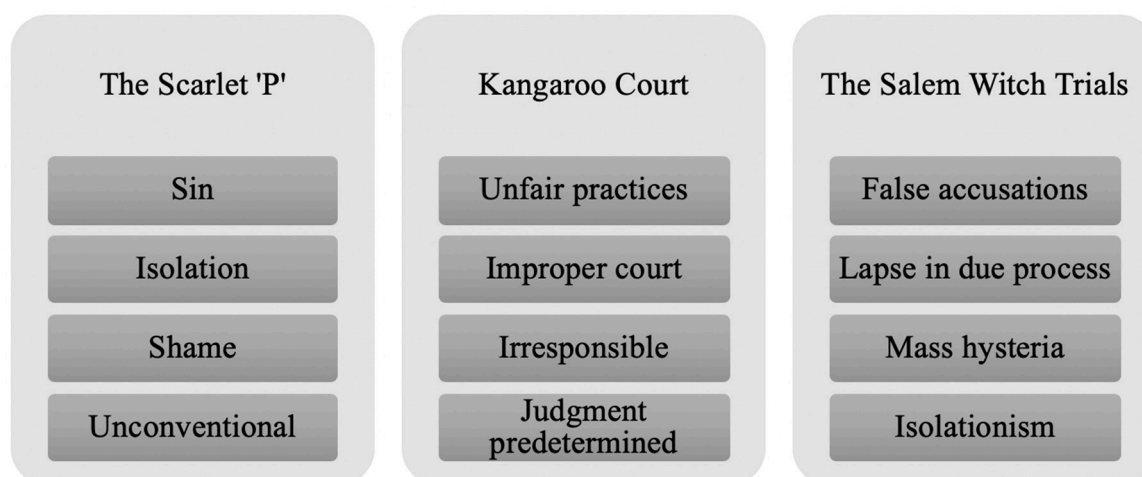


Figure 1. Emergent themes and their major attributes.

The Scarlet P theme addressed the research question: *How does the PAR teacher evaluation process negatively impact teachers?* 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.

The second theme that emerged was the Kangaroo Court. Merriam-Webster [24] defines the term Kangaroo Court as, “a court that uses unfair methods or is not a proper court of law” (para. 1). The findings of this study suggest 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 Kangaroo Court theme addressed the research question: *How does the intended use of the PAR teacher evaluation process compare to the evident use of PAR?* It was also 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. Participants noted administration change as the point of negative change in their evaluation experience.

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. The findings of this study suggest the same tendencies within teacher evaluation processes. Additionally, court case decisions, such as *Vergara v. California* [25,26], 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 Salem Witch Trials theme addressed all three research 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? and, (3) How does the intended use of the PAR teacher evaluation process compare to the evident use of PAR? Bias revealed in the district data illustrated a disproportionate referral to PAR from specific demographics (race, age, gender, and pay scale). Teachers reported having to “play the game” and other survival techniques. Interviews revealed that teachers felt their teacher evaluations were predetermined, regardless of their performance or improvements. Participants speculated that 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 [27] 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” (p. 723).

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.

4.7. Alignment of Findings to Research Questions

There were three major findings that aligned with the three 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.

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, higher education levels, 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 into 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?”

Ladson-Billings and Jackson [28] also explained that state and federal reform undermined the goals and aspirations that have historically drawn African Americans 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 accounted for 6.8% of the teaching staff in the school district [19,20].

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. The interviews and district data 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.

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 support Goldstein’s [1] finding that program ambiguity was a challenge to implementing PAR. This could have contributed to the amount of confusion and frustration that participants expressed throughout their interviews. The evident use of 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 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. Past research has also made similar conclusions about the challenges of implementing PAR. One major challenge reported was program ambiguity [1].

5. Conclusions

Based on the findings of this research, prior research about teacher evaluation, and evaluation policy, punitive teacher evaluation implementation must be halted. Other systems to improve teacher quality and student achievement should be considered. 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. Teacher evaluation does not appear to be the most effective way to increase student achievement and teacher quality, as previous research would corroborate.

An important outcome of this research was a deeper understanding that African-American teachers are negatively impacted by teacher evaluation processes and policies at disproportionate rates. Education scholars have reported concerns about the systematic pushing out of African-American teachers [28–30] and this study supports those concerns. Nieto [30] 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) [30] 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 [19,20]. Teacher education must take a more concerted effort to increase the number of Teachers of Color. Evaluation systems at the K–12 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.

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 PAR policy should be further analyzed. This would provide more information about the parallel traits that exist among those referred to PAR and provide a deeper understanding about the negative impacts of teacher evaluation methods.

Overall, 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. 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. Evaluation systems that prioritize teacher learning over accountability are integral to successfully improving student achievement.

Funding: This research received no external funding.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Goldstein, J. Making sense of distributed leadership: The case of peer assistance and review. *Educ. Eval. Policy Anal.* **2004**, *26*, 173–197. [CrossRef]
2. SRI International. Evaluating Teachers: New Study Underscores Effectiveness of Rigorous Peer Assistance and Review (PAR) Programs. 2012. Available online: www.sri.com (accessed on 19 November 2014).
3. U.S. Department of Education (USDOE). Race to the Top Program Executive Summary. 2009. Available online: <http://www2.ed.gov/programs/racetothetop/index.html> (accessed on 19 November 2014).
4. U.S. Department of Education (USDOE). Teacher Incentive Fund. 2012. Available online: <http://www2.ed.gov/programs/teacherincentive/index.html> (accessed on 19 November 2014).
5. U.S. Department of Education (USDOE). ESEA Flexibility. 2013. Available online: <http://www2.ed.gov/policy/elsec/guid/esea-flexibility/index.html> (accessed on 19 November 2014).
6. Cochran-Smith, M.; Lytle, S.L. Troubling images of teaching in No Child Left Behind. *Harv. Educ. Rev.* **2006**, *76*, 668–726. [CrossRef]
7. Darling-Hammond, L.; Wise, A.E.; Pease, S.R. Teacher evaluation in the organizational context: A review of the literature. *Rev. Educ. Res.* **1983**, *53*, 285–328. [CrossRef]
8. Creswell, J.W. *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*, 3rd ed.; Sage Publications, Inc.: Thousand Oaks, CA, USA, 2009.
9. Donaldson, M.L. *Teachers’ Perspectives on Evaluation Reform*; Center for American Progress: Washington, DC, USA, 2012. Available online: www.americanprogress.org (accessed on 19 November 2014).
10. Hill, H.C.; Grossman, P. Learning from teacher observations: Challenges and opportunities posed by new teacher evaluation systems. *Harv. Educ. Rev.* **2013**, *83*, 371–384. [CrossRef]
11. Markow, D.; Macia, L.; Lee, H. Challenges for School Leadership: The MetLife Survey of the American Teacher. 2013. Available online: <http://www.metlife.com/teachersurvey> (accessed on 19 November 2014).
12. Youngs, P. *Using Teacher Evaluation Reform and Professional Development to Support Common Core Assessment*; Center for American Progress: Washington, DC, USA, 2013. Available online: <http://www.americanprogress.org> (accessed on 19 November 2014).

13. Mead, S. *Recent State Action on Teacher Effectiveness: What's in State Laws and Regulations?* Bellwether Education Partners: Sudbury, MA, USA, 2012. Available online: <http://bellwethereducation.org> (accessed on 19 November 2014).
14. Markley, T. Defining the effective teacher: Current arguments in education. *Essays Educ.* **2004**, *11*, 1–14.
15. Putney, L.G.; Wink, J.; Perkins, P. Teachers as researchers: Using the Critical Action Research Matrix Application (CARMA) for reflexive classroom inquiry. *Fla. J. Teach. Educ.* **2006**, *IX*, 23–36.
16. Merriam, S.B. *Qualitative Research: A Guide to Design and Implementation*; Jossey-Bass: San Francisco, CA, USA, 2009.
17. Biddle, R.E. *The Role of Two Statistical Approaches in EEO Cases*; Biddle & Associates: Sacramento, CA, USA, 1995. Available online: <http://www.biddle.com/documents/twostats.htm> (accessed on 19 November 2014).
18. U.S. Department of Justice (USDOJ). The Freedom of Information Act. 1966. Available online: <http://www.foia.gov/about.html> (accessed on 19 November 2014).
19. California Department of Education (CDE). California Laws & Codes. Available online: <http://www.cde.ca.gov/re/lr/cl/> (accessed on 19 November 2014).
20. California Department of Education (CDE). DataQuest. 2013. Available online: <http://www.cde.ca.gov/ds/> (accessed on 19 November 2014).
21. U.S. Department of Education (USDOE). *Advancement via Individual Determination (AVID)*; U.S. Department of Education (USDOE): Washington, DC, USA, 1998. Available online: <http://www.ed.gov/pubs/ToolsforSchools/avid.html> (accessed on 19 November 2014).
22. Selznick, D.; Fleming, V. *Gone with the Wind [Motion Picture]*; Metro Goldwyn-Mayer: Beverly Hills, CA, USA, 1939.
23. Hawthorne, N. *The Scarlet Letter*; Ticknor, Reed & Fields: Boston, MA, USA, 1850.
24. Merriam Webster. Evaluate. Merriam-Webster.com. 2014. Available online: <http://www.merriam-webster.com/dictionary/evaluate> (accessed on 19 November 2014).
25. *Vergara v. California*, 246 Cal. App. 4th 619; Docket No. B258589S (Cal. Ct. App. 2016). Available online: <https://www.leagle.com/decision/incaco20160414027> (accessed on 30 May 2019).
26. Education Counsel. TLE Legal Update: *Vergara v. California* Court Invalidates State Tenure, Dismissal, and Last-in-First-out Statutes. 2014. Available online: <http://www.educationcounsel.com> (accessed on 19 November 2014).
27. Papay, J.P.; Johnson, S.M. Is PAR a good investment? Understanding the costs and benefits of teacher peer assistance and review programs. *Educ. Policy* **2012**, *26*, 696–729. [CrossRef]
28. Ladson-Billings, G.J.; Jackson, J.F.L. *The Fragmented World of the Social: Essays in Social and Political Philosophy*; SUNY Press: New York, NY, USA, 2007.
29. Ladson-Billings, G.J. Is the team all right? Diversity and teacher education. *J. Teach. Educ.* **2005**, *56*, 229–234. [CrossRef]
30. Nieto, S. Solidarity, courage and heart: What teacher educators can learn from a new generation of teachers. *Intercult. Educ.* **2006**, *17*, 457–473. [CrossRef]



© 2019 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).