Graduate Theses and Dissertations   Graduate School

January 2013

Social Studies Teachers in an Evaluative Role: The Peer Evaluator Experience in the Accountability Era

Martha Barnes Ford
University of South Florida, mbfordy40@gmail.com

Social Studies Teachers in an Evaluative Role: The Peer Evaluator Experience in the

Accountability Era

by

Martha B. Ford

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
in Curriculum and Instruction
with an emphasis in
Social Science Education
Department of Secondary Education
College of Education
University of South Florida

Co-Major Professor: Bárbara Cruz, Ed.D.
Co-Major Professor: J. Howard Johnston, Ph.D.
James Duplass, Ph.D.
Arthur Shapiro, Ph.D.

Date of Approval:
May 28, 2013

Keywords: Observation, Rating, Evaluation, Reform

Dedication

For my family, especially my William: The most important thing in life is to be a good man and the bigger man when you have to. Walk calmly in your path, know who you are, and never give up.

Acknowledgments

I am grateful to my faculty committee for their superhuman patience with me during this prolonged exercise in perseverance. In particular, the co-chairs of my committee have repeatedly legitimized their status as legends in social science education. Dr. J. Howard Johnston has been my academic idol for over 20 years. Despite his broad expertise, he retains genuine curiosity and enthusiasm for academic exploration. Dr. Bárbara Cruz is simply an unstoppable force for good, and represents everything that is right with education and the world. I join countless other students who feel they could never thank her enough for her relentless support.

Table of Contents

List of Figures

Abstract

As the evaluation of teachers becomes prominent in the current climate of educational reform, the details of teacher evaluation systems become important. The goal of the research was to gain perspective about a little-studied group key in education reform efforts to improve teacher evaluation—the peer evaluator. Teachers serving as peer evaluators were interviewed to reveal their perceived lived experiences before undertaking the role, benefits and challenges they perceived during the role, and their perception of how the experience will impact their future as educators. Extensive profiles of three peer evaluators were crafted based on in-depth interviews. The themes revealed from phenomenological reduction analysis included Idealism, Non-Collegial Reality, Valued Experience and Residual Optimism. The study found common elements in the backgrounds and interests of the participants, including previous leadership roles and a shared a sense of idealism entering the role. The study also found that participants' perceived expectations of being able to help fellow teachers were not fully met as they progressed in the role. Among the perceived challenges and benefits of being a peer facilitator, the study found that time constraints and dealing with non-receptive teachers were perceived as the most challenging for participants, while they valued seeing a "bigger picture" beyond a single classroom as beneficial. As such, the ability to see the bigger picture, along with being a good time manager and communicating well with teachers were perceived by the participants as being most useful in performing their jobs. Finally, the study found that the participants predominantly found their experiences as peer evaluators to be affirming of their previously held

educational perspectives, and that the overall experience would impact their future performance as educators in positive ways.

Chapter 1: Overview

Introduction

 In retrospect, it is still hard to gauge all the ways that Coach B was such a bad teacher. Since

he so seldom taught at all, identifying his ineffective instructional practices is difficult. He

usually spent the bulk of his instructional time reading the sports section of the newspaper, or

casually discussing recent sporting events with some of the students in the class. As 8th grade

students, we were ecstatic at the prospect of an easy A in social studies. Coach B was cool, and

in our immaturity, we were glad to have him even if the textbooks we barely opened were the

same we had used and completed in 7th grade. Coach B did not even bother getting the correct

books to assign his students, and as long as the athletic teams kept winning, no one else seemed

to mind, either.   If anyone checked student performance, we all had great marks on the few tests

he administered, mainly because the tests were not difficult and we had often seen the questions

beforehand.

 Coach B is not a fictitious teacher; he was my social studies teacher in the 8th grade. I can

attest that the preceding description is not an exaggeration and that there was no outrage

expressed at the situation. Coach B was extremely popular with the parents and administration.

No one else stepped through the classroom door besides students, and we all gave him glowing

recommendations. What if someone else had stepped through the classroom door? What if a

fellow teacher sat through his one of his lessons and made some notes on what was seen? How

would that information have been used? What changes, if any, could the information have

prompted? Would anything be different for Coach B, his students, or even the teacher who observed his lesson?

Background/Rationale

The situation described with Coach B has been an unfortunate reality in education. There have been and continue to be some ineffective teachers. Identifying them and dealing with them has always been a challenge, but one that goes to the heart of the problem underlying this research study.   A robust educational accountability movement has arisen wherein teacher quality is demanded, and the intent is to cause teachers like Coach B to become more effective or leave the profession. Exactly how to make that happen is the directive of this educational accountability reform, and the consequences of that reform – who and how the accountability impacts – should be thoroughly researched. From beyond the realm of education and educational research, politicians and businesspeople are looking into teacher quality issues, proffering up their own opinions or solutions, and connecting them to policy decisions and funding. In a summer 2010 speech explaining what the current federal administration was doing about educational accountability, Secretary of Education Arne Duncan triumphantly declared:

> Included in the Recovery Act was—by the standards of Washington—a relatively small provision authorizing the Department of Education to design and administer competitive programs aimed at improving education in four core areas of reform: standards, teachers, data and school turnarounds. With a budget of just $5 billion dollars—less than one percent of total education spending in America—this minor provision in the Recovery Act has unleashed an avalanche of pent-up education reform activity at the state and local level. (Duncan, 2010, p.455)

The avalanche was in response to the challenge to all states to adopt better teacher evaluation systems through presenting or withholding funding through what are known as *Race to the Top* initiatives. Those systems for evaluation of teaching mostly involve some combination of a few key components in an overall reform plan, of which two are most pertinent to this study: student achievement outcomes and classroom observation ratings by a superior and/or someone in non-supervisory role. It is important to note that this broad generalization of components is necessary to structure the description of what can be wildly variable in teacher evaluation systems. In a report from the Brookings Brown Center Task Group on Teacher Quality, the authors explain:

> Because of the immaturity of the knowledge base on the design of teacher evaluation systems and the local politics of school management, we are likely to see considerable variability among school districts in how they go about evaluating teachers—even as most move to new systems that are intended to be more informative than those used in the past. (Croft, Glazerman, Goldhaber, Loeb, Raudenbush, Staiger & Whitehurst, 2011, p.1)

Many plans call for gauging teacher effectiveness by the amount of learning displayed by the students taught. The terminology for this kind of measure includes "student achievement" or "student performance." The tool of choice for producing data for this category is through formal testing of students and using the student scores on such tests as proof of teacher performance. The tests can be either norm-referenced or criterion-referenced, but in either case, the scores ultimately are used to compare student performance to, in turn, compare teacher performance. So pervasive and important is this practice, that some researchers call standardized achievement scores the "coin of the realm" in public education in the United States (Haladyna, Nolen, & Haas, 1991, p.2).

As described by several researchers, other teacher evaluation plans choose to factor in some kind of external evaluation through observation of teachers at their craft (Karnes & Black, 1986; Lieberman, 1998; Stodolsky, 1990). In other words, someone will observe a teacher in the classroom while teaching, and from that observation derive data. Such data come in a variety of formats from highly narrative and subjective to what is perceived to be more objective and quantifiable by adhering to some kind of rubric or other observation instrument. This kind of evaluation has most typically been conducted by an administrator or other trained observer, but almost always in a supervisory role. A relatively newer approach in teacher evaluation that includes this kind of external evaluation is to still conduct such observations, but to assign the task of observation to someone in a non-supervisory role. For this study, the term "peer evaluator" will be used to generally describe that role, but it is important to note that in various evaluation plans and in different school systems, a myriad of titles has been used. Also, for this particular study, it is prudent to recognize that the most widespread use of such peer evaluation until recently was not exactly between peers, but between brand new hires and the more veteran counterparts assigned to mentor and help acclimate them to the profession. As such, peer evaluation is not exactly novel, but recent evolutions of peer evaluation --- and the purposes for using it --- are at the current forefront of educational research and policy.

Purpose of the Study

It can be generalized that the focus of all accountability models is ultimately students. Educational accountability has to do with ensuring quality education for all students and improving public education systems for the benefit of students. Teacher quality and accountability is directed toward improving teachers and, in turn, improving student learning. So it is natural that most research about accountability models revolves around the impact on

students or teachers evaluated or more rarely --- in the cases where only administrators do the evaluating --- the administrators. The consequences of implementation of these models are mostly measured or examined for those three groups, but the consequences for the peer evaluators is largely ignored.

As different states and school systems work to develop accountability models, some researchers remind them,

> We cannot expect that work to lead to a perfect assessment and accountability system with no negative consequences, but we can seek to minimize those negatives while maximizing some of the intended benefits by giving careful attention to a wide range of intended and plausible unintended consequences in the design of the system. (Linn, 2010, p. 146)

This study was conducted in that spirit of exploring unintended consequences, whether positive or negative. Since peer evaluators are not the main intended target of reform in accountability systems, or are at best lumped in with the general teaching population, the impact of that system on them can be seen as an unintended, or at least unexamined, consequence. Because peer evaluators can have such a key role in accountability systems, their experience should be examined, not just for testing the design of accountability systems, but on a more human level to see what, if any, transformation occurs when a teacher becomes a peer evaluator.

The purpose of this study is to examine the lived experiences of peer evaluators, who are teachers embarking on a new role. It is of particular interest to this researcher to look specifically at social studies teachers who become peer evaluators. As a social studies supervisor in a large, urban school district, the quality of social studies instruction and curiosity about what social studies peer evaluators encounter naturally interests me as part of my job function. But

beyond my own curiosity or professional interest, there are other factors which point to this particular group as a population worthy of study. First, the inaugural cadre of peer evaluators for the teacher evaluation system implemented in this district included many social studies teachers. The district website features brief biographies of the peer evaluators, and of the original 115 peer evaluators, 21 indicated either recent social studies teaching assignments or a professional background in social studies education. (Per the conditions under which approval for this study was obtained from the school district in which the study took place, the reference for this website could not be disclosed without also identifying the district, so it has been omitted.) Considering the myriad of educational preparations teachers could represent, having such a large portion of the original peer evaluators come from a social studies background would seem significant. Second, those social studies teachers that became peer evaluators have been assigned to observe teachers inside and outside of the subject area. Is there something in the preparation or experience of social studies teachers that may make them adept as external evaluators of teacher quality, even outside their own subject area? Examining how these teachers make sense of their new roles may provide such insight.

Theoretical Framework

The theoretical framework underpinning this study is constructivism. Constructivism is built on the "thesis of ontological relativity, which holds that all tenable statements about existence depend on a worldview, and no worldview is uniquely determined by empirical or sense data about the world" (Patton, 2002, p. 97). Perceptions, emotions, and experience contribute to shaping each person's conception of reality. Put simply by Larochelle and Bednarz (1998), for constructivists, knowledge is "constructed, negotiated, propelled by a project and perpetuated for as long as it enables its creators to organize their reality in a viable fashion" (p. 8).

As such, the appeal for this study is that constructivism recognizes the temporal and situational limits of reality, but that recognition gives power or voice to the individual experience. Among the assumptions of constructivism, Guba and Lincoln (1989) explain that "phenomena can only be understood within the context in which they are studied; findings from one context cannot be generalized to another; neither problems nor solutions can be generalized from one setting another" (p.45). This does not necessarily denigrate the context, it simply gives no higher value to context over experience; it takes both to make up knowledge and reality.

Phenomenology, as a constructivist method of investigation, builds up knowledge as a process of development. This phenomenological study intentionally brings focus onto peer evaluators by the investigator conducting in-depth individual interviews to reveal the lived experiences of peer evaluators in a large, west-central Florida school district. That district has implemented a teacher evaluation system that determines 60% of a teacher's evaluation rating based on a combination of observations by administrators and peer evaluators, while 40% is based on student performance on assessments using a value-added model. More about this district, model and study population will be discussed in Chapter Three, but it is useful to note the weighting of observation ratings to help explain this study's theoretical framework as it operates on just one concept—the weighting of inputs.

Initially, the weight of the peer evaluator's input on a teacher quality rating seems neutral—it may seem like just a percentage. Considering that one data point against the context of the study, however, may reveal it as a potential influential concept. Before the implementation of the district's new model for teacher evaluation, there were no peer evaluators, and only personnel in a supervisory role ever had input into a teacher's performance rating. In light of that information, the concept of weighting now may seem to accord peer evaluators an unprecedented

power and responsibility with respect to their fellow teachers that they would not have encountered in their previous role as classroom teachers or that could have occurred in previous systems. It indicates that role transformation may be pertinent to the lived experiences of the participants, which is explored through thoughtful structuring of interview questions and careful analysis of the interview answers.

As explained succinctly by Swenson (1996), a qualitative approach to investigation is appropriate when the "precise nature of the problem investigated is fluid" (p. 188). This condition applies with this investigation because the background of the problem shows that education is in a time of systemic flux --- where problems arise, the solutions may create other problems. Of additional utility to the investigator employing a qualitative approach is that it can allow for a purposeful sample. In this case, that purposeful sample is that subgroup of one district's cadre of peer evaluators that comes from a social studies teaching background.

Research Questions

Although the peer evaluator position, in its current form, came about because of large scale reform, the evaluation of such reform has centered on effects on students and the classroom teacher, not the peer evaluator. As such, there is a dearth of hypotheses or research about the peer evaluator. In the concept example used previously, some basis was established for asserting that the peer evaluator is a pivotal role worthy of investigation. Research questions were framed toward the end of illuminating that role, impacts on it, and impacts of it. This study addresses the following research questions:

1. Are there common perceived, self-reported elements in the backgrounds, interests or motivations of social studies teachers who elect to become peer evaluators?

2. What perceived expectations do these teachers report that they hold before becoming peer evaluators, and how are these perceived as met or not met as they progress through their roles?

3. What perceived challenges or benefits do the peer evaluators report that they encounter in performing their duties?

4. How do peer evaluators perceive their experiences in the role will impact their own performance as social studies educators?

5. What behaviors or abilities do peer evaluators perceive to be most useful in performing their jobs?

6. Does being a peer evaluator change a person's perceived self-reported educational philosophy or perspective?

Overview of Research Methods

In order to address these research questions, the investigator conducted in-depth interviews with volunteers from the target study population of social studies peer evaluators. The initial answers drove and transformed follow-up questions. The data derived from the transcribed interviews were analyzed using phenomenological reduction. More detail about the study methods is described in Chapter 3.

Statement and Analysis of the Problem

The specific problem being addressed in this study is the lived experience of social studies peer evaluators as they are impacted in that role. The broader underlying problem prompting this study is the conundrum of teacher evaluation and its impact in an era of accountability. Discussions about teacher quality have historically engendered controversy and charged emotions, mostly because of the perceived inefficiency and shortcomings of previous systems.

"Our system of teacher evaluation…frustrates teachers who feel that their good work goes unrecognized and ignores other teachers who would benefit from additional support" (Duncan, 2010, p. 458).

Yet not everyone embraces reform of teacher evaluation and accountability policies. It would be hard to argue, however, that when it comes to the good of children or the investment of the public's money, accountability should be ignored. Clearly, the opposite is imperative—that there be some way of indicating that what is being done in public education is somehow effective and fund-worthy, particular when it comes to teaching. If it could be supposed that a large portion --- perhaps the even the majority --- of the teaching force is well-educated, prepared and motivated to do a good job, then why would some teachers or other entities fear evaluation? As noted by Hoggarth and Comfort (2010), it is because there is so much potential for evaluation to be done the wrong way and "there are certainly plenty of wrong reasons for investing in it" (p. 19). On the list of inappropriate reasons, they include management trying to avoid responsibility or trying to find justification for not funding. Any worker aware of the potential for nefarious motivation or bad implementation would be justifiably leery of evaluation processes. And if a worker has additionally experienced unfair or next-to-meaningless systems of evaluation, the worker's consequent feelings toward evaluation would be negatively impacted with emotions ranging from apprehension to frustration.

Charlotte Danielson's (2010) critique of traditional evaluation systems marks their weaknesses as 1) using outmoded evaluation criteria, usually in the form of checklists; 2) using simplistic evaluative comments with no consistency as to what comments mean; 3) applying the same evaluative procedures for both novices and veterans; 4) lack of consistency among evaluators and 5) top-down communication that makes the evaluation feel punitive. Her critique

is echoed by many researchers (Marzano, 2007; New Teacher Project, 2010; Weisberg et al., 2009) who paint a dismal picture of how teacher evaluation has been conducted in the past, bemoaning, as Bridges does, that if criteria are vague or too subjective, "teachers will be struggling blindly to meet undefined and unknown performance objectives (Bridges, 1992, p. 34). Stressors such as these will be explored later in a review of the available literature, but clearly, there is much stress and confusion relating to teacher evaluation and changes to it.

Ultimately, almost any educational research study becomes a study about change, or more likely, research into the response to some change effort or reform. As external accountability in public education increases, it precipitates changes at a speed heretofore not common in education, an institution where real and sustained change has traditionally been slow. Paradigms and roles in education consequently shift perhaps uncomfortably quickly, and a profession which has been characterized by unique stressors can become even more stressful. What are these lightning changes? How are they impacting the different participants in the educational system? Put simply, what does change do to and for the people involved? Research toward the purpose of illuminating such change is merited.

Much educational research related to accountability topics has used or uses the familiar process-product approach. It mirrors the industrial approach to evaluation where a process is instituted with a particular population with the expectation that the population will produce something. This study instead goes down to a personal level within the larger context of teacher accountability to explore the *unintended* consequences of teacher evaluation reform on the trench-level enforcer of those reforms --- the peer evaluator. As such, it is not so much a product as a perspective that is the goal of this research. Perspective can belie the lived experiences of the participants, and vice versa. Perspective gives meaning to experience by its very subjective

uniqueness. This is put more eloquently by the investigator's favorite science fiction author in a passage where an alien character explains to an earthling character about an impending experience:

"…you'd better be prepared for the jump into hyperspace. It's unpleasantly like being drunk."

"What's so unpleasant about being drunk?"

"You ask a glass of water" (Adams, 1979).

Assumptions

This study posits that there is a context of confusion and possible stress due to changes precipitated by newly implemented teacher accountability models. This assumption has three critical components—change, stress, and accountability. First, in considering change, it is recognized that education as an institution has been characterized by periods of change and transformation. Second, whether change-initiated or inherent to the institution, education has its own particular stressors that interplay and affect its participants. Finally, accountability itself is challenging to all participants, not just those who are being held accountable.

In addition, this study makes the assumption that there are unintended consequences for almost any plan of action. As such, it is not always possible or necessary to know ahead of time what those consequences will be. With proper investigation of different perspectives, such consequences may be revealed. So, although perspective is the goal of the study, it is assumed that such perspective will reveal an unintended consequence of the overall study context.

Definition of Terms

In this study, the most important term is the title of the key figure, the *peer evaluator*. For this study, the peer evaluator is an educator who has sought out the role to be an external

evaluator of fellow teachers' classroom practices. The peer evaluator is released from classroom teaching duties during their tenure as an evaluator, and evaluates teachers at possibly several different schools. The peer evaluator, once hired, is trained in using a rubric for conducting observations, and mandated further and continuing training in being calibrated and conducting feedback sessions. In the literature, this term has been used in different accountability models that do not necessary accord the same duties, and is sometimes confounded with *peer reviewer* or *mentor teacher*. To be clear, other than offering cursory "next steps" suggestions to teachers following a classroom observation, these peer evaluators are not used in a mentoring role. A separate cadre of teachers was recruited to mentor new and struggling teachers in the system where these peer evaluators work.

Other terms associated with this role are explored throughout the rationale and the literature review for this study. Overall, the context of accountability requires terminology that reflects and differentiates the stakes behind the terms. For one example, *teacher quality* may seem interchangeable with *teacher effectiveness*, but such comparison may be woefully misguided if one term is associated with professional reflection, and the other associated with performance ratings. Connotation of these terms is married to the example or system in which they are used, so it is imprudent to establish generalized definitions here.

Limitations

The size of the sample population naturally limits how the findings of this study can be generalized to other situations. Because the context of the study problem is in a state of fluidity, the problem may also be in a similar state. The effects of both of these limitations are hoped to be ameliorated by the use of a phenomenological approach and very careful attention to study methodology to increase validity. Concern about the investigator being an involved participant

is only a limitation to the extent that the interviewees have had previous contact with the researcher in that the investigator is a *curriculum supervisor* in the district of the study population. A curriculum supervisor interacts with classroom teachers, principals and other administrators, but has no supervisory capacity over the peer evaluators once they are not classroom teachers. Even when they have been in or return to classroom teacher roles, the supervisor does not make personnel decisions that impact those teachers. In that sense, and because their participation is strictly voluntary, it was not anticipated that any relationship between the investigator and volunteers would significantly influence the data, but it is noted as a possible limitation nonetheless.

Chapter 2: Literature Review

Introduction

There is little available literature thus far on research relating to peer evaluators in the role characterized by the population of this study. For this study, then, it becomes even more important to review the existing literature for elements that help to describe the broad context in which peer evaluators work so that the more nuclear perspective and experience of the study population can be triangulated against it. This approach is recommended by Shenton (2004) for a qualitative study to help ameliorate any transferability issues. While many states and districts are moving forward with accountability schemes, they are likely to run into real problems with design or interpretation that compromise and distort their impact. In fact, while it seems natural to measure outcomes and hold schools responsible for them, research has shown the reality is much more complicated (Hanushek & Raymond, 2001). However, the investigator's goal with this literature review is not to explain the entire universe of accountability and reform, just what elements are impacting the galaxy in which the peer evaluators operate.

The Accountability Movement

Almost since the beginning of American public education in the nineteenth century, there have been dilemmas associated with determining the quality of that education. Parents and policymakers have a vested interest in student learning and teacher competency. Meeting the need to reassure everyone that students are getting a good education, that teachers are effective, and that funds are spent wisely is not a task left to a single entity. Families at home, teachers in

the classroom, school leaders, district administrators, policymakers at the state and national level, and colleges of education all share a piece of the responsibility.   Increasingly, almost everyone else who does not shoulder that responsibility proffers opinions since, by their own extrapolation, the entire society is impacted by the quality of a country's education system.

To mangle an old adage, this makes for quite a few cooks who could potentially spoil the broth. It is not surprising, then, that efforts at quality control in public education have been historically confounded by mixed masters, political influences, and social climate. At the inception of public education, exposure to formal learning for the masses was a novel concept. Funding new schools and someone to teach in them was a local effort town by town, and almost every town approached the task differently (Sedlak, 1989). Introducing organization beyond that level became the stuff of educational legends like Horace Mann, who in the 1840s developed some standardization at the state level in Massachusetts (Asp, 2000).

Standardization was the mantra of quality throughout the rest of the 19th century and well into the 20th century. State departments of education and professional societies worked to standardize requirements for high school graduation, teacher education programs, high school curriculum, and organizational structure (Arends, 2006). Once standardizing the setting and preparation became entrenched, new focus on output and results surfaced mid-20th century. The concepts of external accountability and standardized testing emerged then fully bloomed toward the end of the century. Information from nationally standardized tests was used in funding decisions. The Elementary and Secondary Education Act of 1965 provided federal funds to support reading and math instruction for low income students and, of most significance to the trend that would follow, introduced emphasis on program evaluation with objective measures tied to student learning.   "The resources provided for evaluation energized the educational

research community, and many of the early evaluations of federal programs were the beginning

of the process–product research that shaped the view of teaching and teacher education for nearly

3 decades" (Arends, 2006, p. 7).

The evolution of process-product research and national scales in education reflect many

aspects of the evolution of the United States. There was a shift from regional interests to a

unified nation competing internationally. Transportation and technology did not just unite, but

bound citizens across the country. The ideas of industry and mass production spread into all

areas of society: "Sustaining and intensifying the influence corporate leaders exert today in

education are the assumptions that the practices and discourse of business are not only objective,

rational, and applicable to any organization, but also fundamental to running schools and

providing education" (Taubman, 2009, p. 99).

Education itself prompted evolution. Once it was offered to the masses, it caused the masses

to become people with firsthand experience and expectations of education for their own children.

All of this added up to education policy and practice becoming something that nearly everyone

both judged and felt qualified to judge, and the judges demanded results. Further, the notion that

the results should be gauged to some mass scale as had proved useful in our industrial

development seemed imperative.

This imperative characterizes the era of accountability that began at the end of the end of the

20th century to current times. The pessimistic report, *A Nation at Risk*, included statistical

claims such as the "average achievement of high school students on most standardized test is

now lower than 26 years ago once Sputnik was launched" (National Commission on Excellence

in Education, 1983, p.8). This put U.S. public education in the spotlight at the national level, and

perhaps has permanently relegated it to being in crisis status, as evidenced by the subsequent

analyses, policies, and legislation that labeled it so: Goals 2000, Education Act of 1994, and No

Child Left Behind, to name three key initiatives. Taubman (2009, p. 8) refers to the current

educational climate as our "audit culture." In this study, the phrase "imperative context" will be

used to indicate the sense of urgency that overlays that audit culture in education that can further

propagate and sustain its prolonged crisis status.

Change Management

Much has already been noted about a sense of urgency in the imperative context of the

current educational landscape. The success formula for many action movies is to introduce

change and speed in plot lines, often to dizzying effect. That same formula, however, introduced

into the management of educational systems could prove disastrous. How many innovations can

be adroitly handled and monitored at once? Education systems are first and foremost human

systems. That automatically ensures huge variability in the swiftness at which the systems and

members of the systems can accommodate, implement or accept change. Considering the

assumptions about change and change-induced stressors in the context of this study, at least a

superficial exploration of the field of change management literature is indicated.

The now-classic parable *Who Moved My Cheese?* (Johnson, 1998) encourages readers to

identify with characters as they approach big changes in their path to getting what they want. In

the forward to the book, Kenneth Blanchard explains the purpose of the parable: "As you know,

living in constant white water with the changes occurring all the time at work or in life can be

stressful, unless people have a way of looking at change that helps them understand it" (p. 17).

Change presents the potential for stress to any organization. It may be oversimplifying to

think that understanding a change or making a conscious effort at controlling attitudes about

change will reduce the stressors. But the simple idea marks the notion that within the umbrella

of management, change should be addressed, and this is the central theme of several journals on change management. Not all of those journals' contents will be reviewed here, but there are some concepts to be culled that are pertinent to this study.

The first pertinent concept beyond recognizing that change is a key factor in management is that it is not always clear who should or can best manage change, or who it impacts. Scholars in *The Journal of Change Management* assert, "Arguably, change practitioners (i.e. managers and consultants) are often immersed in the everyday life of organizations to an extent which makes it difficult to see beyond organizational ends" (Todnem, Burnes, & Oswick, 2011, p. 2). These same change management theorists posit that the evolution of their own field shows movement from thinking that change management is only an activity for managers: "We could reframe change management as a micro-situated, everyday distributed practice rather than perpetuating the dominant perspective which treats it as a strategic tool deployed by key actors" (p. 3). That makes it important to consider perspectives and input from inside stakeholders that are outside of immediate management, and even from external observations like those afforded by this study. This idea that change can be at once massive but also micro-situated further supports examining a system in flux from the perspective of individual members of the system.

The second pertinent concept from current change management literature introduces the limitations of change management based on evolving conceptions of change. Todnem et al. (2011) explain that there exists a shift in the approach of change management research away from a "problem-centered approach and focus on changing tangible processes" (p. 3). That approach is termed *diagnostic* change; the newer approach is termed *dialogic* change:

> Dialogic changes center on the processes of social construction and systems of meaning-
>
> making with a view to changing mindsets rather than changing more concrete phenomena

(e.g. behavior, procedures or structures). The real-time social negotiation of meaning associated with dialogic change offers a significant challenge to the manageability of processes of change management insofar as it involves 'coordinating' and 'facilitating' change conversations in the moment and on a largely improvised and unscripted basis rather than engaging in more established forms of planned change. These developments, and the disjunctures in change management that they reveal, clearly undermine conventional wisdom and established practices. (p. 3)

If it can be asserted that education reform, although operating on a major societal institution, is in a state of dialogic change, then fundamental management practices could be questioned. Some change management theorists question the ability to meaningfully manage change or to manage it at all, particularly in light of the inherent complexity of organizations (Hughes, 2006; Shaw, 2002; Stacey, 2001). Other researchers (Taneja, Pryor, Humphreys & Singleton, 2013) concede that "the biggest challenge for organizations is to function within a culture of organized chaos" (p.123).

Those same researchers, however, believe that chaos theory and systems theory can work together, and they promote strategic management of chaos which includes:

- Taking an expansive look around at what's going on outside the organization and how it might affect the organization (an environmental scan) and measuring the extent and possible impact of existing or potential disruptive complexities;

- Assessing the internal environment of the organization for the causes of existing or potential turbulence;

- Aligning the strategic management of chaos with the organization's mission, vision, core values, and goals in order to strategically manage the existing turbulent environment and/or to invent the organization's future; and

- Establishing performance excellence expectations and measurements as a result of existing or potential turbulence inside and outside the organization. (Taneja et al., p. 122).

The aforementioned two key concepts of change management theory resonate well with the imperative context of education reform. The key seems to be to go back to the simplified idea that acknowledging chaos or change can help to manage it. To consider another metaphor, recognizing that the sea will be turbulent does not calm the sea, but it can help you adjust your boat for a more efficient sail even though you cannot predict exactly how rough the waves will be. Reforms and innovations will inevitably cause some organizational turbulence. Turbulence cannot simply be managed in a traditional sense with only hierarchical leaders involved in the management. It takes a whole crew, even the newest mate, to get a boat through a storm. The captain that ignores that can stress the boat to the point of foundering. That is why the lived experiences of the participants in this study matter: success or failure of a teacher evaluation system may partly depend on those lived experiences amid the change and turbulence.

Educational Stressors

Education, like any career, has stressors and stressed-out people working in it. As the demographics of the teaching population change, it is likely that more stress will continue to cause high turnover (Ingersoll & Merrill, 2010). The topic of teacher stress has become so large and pervasive, that full discussion of it would overshadow the overall purpose of this study. The investigator is not including stress in the literature review to cast a cloud over what might be revealed in this study's interviews or to bias the outcome. The focus on stressors is included

because, as previously noted, the peer evaluator is a relatively new role in the district in which the study will take place. This is only the third year of the plan's implementation, and it is logical to accept that with any transition in roles, there may be some stress.

In general, much of what happens in accountability could cause stress to the educators involved. When stakes are high, someone will feel the pressure. Beyond stress reaction to change as a whole, teachers under scrutiny may react to an age-old role duality that resurfaces in recent discussions of teacher quality (Lambert & McCarthy, 2006). The duality has to do with whether teachers are accorded professional status.

Throughout history, occupation structure has been linked to privilege and social rank. The occupations historically given the prestigious designation of professions included government, law, church, and army, but expanded over time to include medicine, engineering, and other sciences. What set these occupations apart was the education needed for entry to them. Those within professions, with their acquired knowledge, were considered learned and therefore capable of self-regulation.

> The union of education and occupation is expressed by speaking of the learned professions. These professions formed a hierarchy of status that both paralleled and cut across the hierarchies of birth on the one hand and economic success on the other hand. These ranking systems co-existed and overlapped. Each provided the possibility of a degree of social power over other men. (O'Boyle, 1978, p. 241)

It is ironic that the very agents of learning, teachers, have not traditionally been accorded the same status as the very professionals they helped to prepare. At the post-secondary level, there are some programs and individual educators that may be held in high regard, be paid commensurately, and are allowed to some degree to self-regulate. But public school teachers

have found it necessary to rely on collective bargaining to safeguard rights and improve their

working conditions and wages. Reliance on unions is more characteristic of labor and trades, as

are the lower wages that teachers can earn. This duality has led some researchers to study if

teachers even see themselves as professionals (Martinez, Desiderio, & Papakonstantinou, 2010),

and led teacher unions to grapple with their own existence as champions of workers while trying

to also be champions of professionalism (American Federation, 1999; Kelly, 1998). Other

researchers have come up with definitions of professionalism as it applies to teachers:

1. Professionals develop a specialized knowledge base from which decision can be made on

    behalf of clients;

2. Professionals have the ability to apply that knowledge in individual, non-routine

    circumstances;

3. Professionals have a strong ethical commitment to what they do (Wasley, 1991).

Not to ignore the list, but the crux for teachers is that even if these three things are in place, they

are not seen as professionals until someone legitimizes them as such. Unfortunately for teachers,

their professional status has not been helped by the teacher evaluation systems that could have

helped to confer that legitimacy.

   A review of some criticisms of teacher evaluation systems will help to illustrate how it is that

the systems can undermine both prestige of the profession and public confidence in the

profession. The New Teacher Project is a nonprofit organization founded by teachers in 1997.

According to its website, the New Teacher Project (TNTP) "works to end the injustice of

educational inequality by providing excellent teachers to the students who need them most and

by advancing policies that ensure effective teaching in every classroom." In 2009, TNTP

released an influential report titled *The Widget Effect: Our National Failure to Acknowledge and*

Act on Differences in Teacher Effectiveness. The report's strong criticism of evaluation systems in U.S. education explains the "Widget Effect," that is, the idea that somehow teachers are interchangeable because school systems wrongly conflate educational access with educational quality. Because teacher evaluation systems can be so simplistic or not truly reflective of performance, and principals often may have had little or no training in evaluation, nearly everyone could receive good ratings. "Our research reflects that there is a strong and logical expectation among teachers that they will receive outstanding ratings" (Weisberg et al., 2009, p. 22).

Rating systems or scales have indeed seemed to be oversimplified or carrying no truly useful information (Goldstein, 2007). The vagueness is evident in some key phrases from an unsatisfactory teacher performance rating: "inability to attain and maintain a level of acceptable teaching performance" and "failed to create an appropriate classroom atmosphere to develop the students' interests in an attitude for learning" (Lawrence, Vachon, Leake & Leake, 2001, pp. 81-82). How would someone specifically go about improving on such vague directives? At the other end of the spectrum, how would someone provide proof of excellence, and what incentive previously existed to do so? Some salary schedules can prescribe the same pay for any teacher who meets a minimum standard. Some salary schedules involve pay bumps for having advanced degrees and years of experience. But researchers have done many analyses that point to those factors as having inconsequential impact on student achievement, making the professional achievement and/or pay differential also meaningless (Borman & Kimball, 2005; Ehrenberg & Brewer, 1994; Goldhaber & Brewer, 1997; Hanushek & Raymond, 2001; Stronge & Tucker, 2000).

To address some of these issues, and restore some incentives in pay reflecting quality or performance, there are school systems where "career ladder" plans are devised. Career ladders are promoted as being predominantly rewarding vs. punitive. Arizona experimented with career ladders beginning in the 1980s. Although it eventually lost funding, the Career Ladder Program allowed for participating districts to pay teachers at different compensation levels based on their performance, which included measures like professional development, assumption of extra roles and monitoring student achievement of benchmarks (Sawchuk, 2010). This was the setting of the single study the investigator found that approximated the effort of this study. Done for a dissertation project, Samuel McClung (1993) was himself one of the Career Ladder Program's cadre of evaluators. In that role, he noticed that many former peer evaluators were moving into supervisory or other administrative leadership positions rather than to return to teaching class, and he wondered about their experiences in the role that prompted such career moves (McClung, 1993). While the study had many limitations, including having a small sample and the investigator being an involved observer, the study is pertinent to this one because of the focus on the peer evaluator as the target of study.

Peer Evaluators

The term peer evaluator has already been initially described in this study as a key term and pivotal role. It is a role expected to be carried out in systematic fashion after training, and within clear guidelines and procedures, yet the small amount of literature thus far on the experience of individuals in that role leaves much ambiguity in its description. The peer evaluators in this study walk on a cloudy path between truly being peers to the observed teachers as the title implies, and not being peers since their evaluation of their peers inherently sets them apart. For the peer evaluators in this study, that path between gets even longer since the peer evaluators'

observations produce ratings that then count toward the appraisal of the observed teachers' job performance.   As pointed out, the professional literature mostly shows that the peer evaluator term has been used in different accountability models that do not necessary accord the same duties, and is sometimes confounded with other roles to the effect that there is very little research to be found that focuses on the kind of peer evaluators to be examined in this phenomenological study.   With that in mind, the following section of this study will elucidate the confounding variety of accountability terms and forms that, in effect, surround the peer evaluator concept. By exploring the literature in this manner, the reader can build his or her own concept of the role of the peer evaluator.

Accountability Terms and Forms

Accountability systems utilize varied terminology for related concepts. One such term from the literature is "appraisal." Appraisal exists in other fields besides K-12 public education. It is a basic human resources function in business and industry, where management wants to measure how effective its workers are. Miller (2001) conducted an exploratory field study to see how receptive employees of five different business firms were to different kinds of appraisal. More specifically, Miller's study tested the assumption that individuals are more satisfied with the appraisal process when they have an opportunity to evaluate their peers, supervisors and themselves. Two different paper-and-pencil instruments were administered. After analysis of the results, Miller came away with two findings of interest to this study: 1) satisfaction with an appraisal matters because if employees are not satisfied, they are less likely to improve their performance, and 2) the appraisal source's impact on satisfaction matters when employees are evaluating themselves or a leader, but there was no impact on satisfaction when rating a peer. Another study from the business world confirms that peers do not particularly like evaluating

each other, especially if any policy decision may result (Long, Long, & Dobbins, 1998). Miller interpreted this phenomenon to indicate that trust is very important when implementing multi-rater appraisal systems, and she came to the conclusion that multi-rater appraisals should be offered to employees as an optional tool, and even so be gradually adopted.

Miller's findings may highlight a subtle difference in meaning between evaluation and appraisal. In a book about a two-year study of teacher appraisal in England (Wragg et al., 1996), researchers offered conventions which included the following descriptors:

1. Evaluation – a general term used to describe any activity by the institution or LEA where the quality of the provision is the subject of systematic study.

2. Review – a retrospective activity implying the collection and examination of evidence and information.

3. Appraisal – emphasizing the forming of qualitative judgments about an activity, a person or an organization.

4. Assessment – implying the use of measurement and/or grading based on known criteria.

These conventions may be useful in that they show how many purposes are sometimes conveyed under the single term "evaluation," or how the terms are mingled without thought to nuance of purpose. For the English study, the emphasis was clearly on teacher appraisal, which the researchers note is "in theory, a self-evident, neutral notion that can acquire overtones of retribution or support, depending on the individual's viewpoint" (Wragg et al., p.7). Since the scope of the English study was wide, and it had too many varying and confounding components, its context and results will not serve to inform this study, but the distinction of conventions merited its inclusion.

With these conventions in mind, and considering that some accountability terminology is neutral until it is assigned overtones by those involved, it is interesting to look at some examples in the available literature where the overtone of collegiality reigns.   Wilkins and Shin (2011) introduce yet another term, "peer feedback" in describing a study where teachers in pairs engaged in reciprocal teaching. They then observed one another to promote "professional development, collaboration, and self-assessment" (p. 50).   Given the opportunity to collaborate in such a manner provided members of the pairs a way to gain valuable insights into their teaching practices in the relatively safe environment of a partnership. Such collegiality would seem to alleviate the isolation that teachers can feel and would promote professional dialogue.

That kind of collegiality is what has been promoted at the higher education level with its tradition of peer review (Fernandez & Yu, 2007). But even at that level, as pointed out by Chism and Stanley (1999), the process can be fraught with challenges that make its implementation difficult and the results less than stellar. According to them, real peer review is "informed peer judgments about faculty teaching for either improvement or judgment purposes" (p. xi). In this work about peer review at the college level, one of the most useful things addressed is the role of the peer reviewer. To that end, the authors point out that it is actually a multiplicity of roles including information gatherer, staff developer, or judge, depending on whether the reviewer is engaged in formative review or summative review, an important distinction.

The distinction becomes critical because there is a line drawn between what can be considered formative evaluation with its adherent connotations of informality and collegiality, and summative evaluation which implies external evaluation with higher stakes and judgments. As an example, Brigham Young University's Center for Teacher & Learning features on its website the following overview of peer review of teaching at the university level:

- *Formative Review 1.* Between the first and second year, a peer mentor (in partnership with a teaching and learning consultant if desired) would review the new faculty member's teaching through evidence of student learning to provide formative feedback regarding teaching and student learning to the faculty member.

- *Summative Review 1 (3rd Year).* Peer reviewers from the department's rank and status committee would review the faculty member's teaching through evidence of student learning to make recommendations for the faculty member's status with the university.

- *Formative Review 2.* Between the fourth and fifth year, the peer mentor (in partnership with a teaching and learning consultant if desired) would again review the faculty member's teaching and would focus, in particular, on any recommendations from the third year review.

- *Summative Review 2 (6th Year).* Peer reviewers from the department's rank and status committee would review the faculty member's teaching through evidence of student learning to make recommendations for the faculty member's status with the university. (Center for Teaching and Learning, 2013)

From the example, the descriptions of the two formative reviews reveal that the process at those points is to provide feedback primarily to the faculty member. The description of the two summative reviews, however, reveal that at those points, evidence about the faculty member's teaching will be used to make recommendations about the faculty member's rank and status with the university. In other words, summative reviews in this example impact employment; the stakes are raised.

The Evaluation of Teachers

Those higher stakes and judgments found in such summative reviews are representative of the direction in which teacher evaluation is heading at the K-12 level. There is growing emphasis on observation and evaluation of teachers for evidence toward making rank and status decisions vs. observation and evaluation for reflection or growth. In the description of her own novel plan for observing teachers at her school, Principal Jenne Colasacco (2010) relays her satisfaction at the depth of information she could obtain about teacher performance while spending one week in each teacher's class. With pre-observation conferences and mid-point consultation, a dialogue evolved with the observed teacher that Colasacco found valuable. At the end of the school year, however, she noted the time constraints of observing so methodically. She also still missed the unannounced visits she would make to classes before trying the week-long approach, and intended to move on to a blended model of the two kinds of observations for the next school year. Apparently, the unannounced visits to classes provided a draw for the principal, who, after all, is in a supervisory role. This draw was delicately described as providing confirmation that what was discussed for improvement during the week-long observation would be evidenced when there was no guarantee the principal would be there for oversight. Regardless of the principal's intent, there is still the implication that she is looking for a "gotcha" moment, that is, trying to catch teachers off-guard or looking for something to criticize. Fear of such "gotcha" accountability, regardless of who is doing the evaluating is a common complaint of teachers. Are the fears well founded? The following case presents what happened in Washington, D.C. schools.

Michelle Rhee made headlines when she was made chancellor of Washington, D.C.'s schools and promptly fired administrators, closed schools and took other actions to follow through with

her plans to improve the capital city's schools. Her professional background as she came to the role of chancellor adds to her notoriety:

> Michelle Rhee had been a surprise choice to lead the schools. After college, she joined *Teach for America* and taught for three years in a low-income school in Baltimore. After earning a graduate degree in public policy at Harvard, she took over a fledgling non-profit that recruits mid-career professionals into teaching, The New Teacher Project. In that role, she eventually ended up supervising 120 employees. As Chancellor, Rhee would be managing a school system with 55,000 students, 11,500 employees and a budget of nearly $200 million. (Merrow, 2013, section 2)

Once in the role of chancellor, she surrounded herself with a team of people who also had no experience running a school system. This team included her best friend (with similar experience) being named deputy chancellor, and a teacher straight from the classroom tasked to create a new teacher evaluation system and union contract (Merrow, 2013). What followed was a period of upheaval in the District of Columbia school system.

The firing of teachers during this upheaval peaked in 2010: 302 employees total were fired, of which 241 were teachers under Rhee's new evaluation system for school-based personnel called "IMPACT." Of the total employees fired, 226 were terminated for poor performance, and 76 terminated for other problems like not having the licensing required by the No Child Left Behind Act. Besides the 241 teachers, those dismissed were librarians, counselors, custodians and other employees. 737 other employees were rated "minimally effective," meaning they had only one year to improve or be terminated (Lewin, 2010, p.A8).

Because of controversy about the IMPACT program, and in part because the mayor who hired her lost his bid for re-election, Rhee resigned in the fall of 2010. In her press conference

announcing her resignation she steadfastly defended her reforms, and vowed to continue reform efforts on a broader scale (Moroney & Young, 2010). After Rhee's departure from the position, Washington, DC schools continued to rely on the IMPACT system of evaluating teachers that relies on data from student test scores and classroom observations of teaching. Those observations are carried out by roaming evaluators called "master teachers" hired by the city to observe its 4,200 teachers. The IMPACT program has been tweaked since its inception, but it has been held up as model to many school systems wanting to compete for *Race to the Top* funding.

In the district which is the context of this study, a program similar to IMPACT has been devised. Teachers must be observed both formally and informally by an administrator and peer evaluator at least once each school year, but potentially several times throughout the year, depending on previous performance appraisals. The lower the performance appraisal rating, the greater the number of observations of that teacher that will be required in the subsequent school year. During observations, the observer documents evidence of effective teaching behaviors then compares the documentation against an established rubric to derive ratings of the observation. Those ratings factor in, along with other data like student performance, to give a teacher an overall performance rating.

Student Performance and Assessment

Because the concept of student performance and assessment is so massive and complex, it would detract from this study to delve too far into this topic. But it is necessary to elaborate on its form somewhat to understand the terminology for and place within accountability discussions. As noted previously, some recent plans for teacher accountability use student performance on standardized tests as part of a teacher's overall quality rating. The trend has been to include

student performance into a teacher's evaluation through a value-added model. The actual

mathematical calculations needed to derive value-added measures are complex, but DiCarlo

(2012) gives an effective overview of the value-added concept:

> Value-added models are a specific type of *growth model*, a diverse group of statistical
>
> techniques to isolate a teacher's impact on his or her students' testing progress while
>
> controlling for other measurable factors, such as student and school characteristics, that
>
> are outside that teacher's control. Opponents, including many teachers, argue that value-
>
> added models are unreliable and invalid and have absolutely no business at all in teacher
>
> evaluations, especially high-stakes evaluations that guide employment and compensation
>
> decisions. Supporters, in stark contrast, assert that teacher evaluations are only
>
> meaningful if these measures are a heavily weighted component. (p. 33)

Regardless of what assessment model is used, how can a test for students be used to gauge how

teachers are doing? Towards an answer, a familiar performance measure will be discussed to

show how there has come to be more than one purpose for student assessment.

One nationally administered test clearly illustrates the multiplicity of purposes in student

assessment without even having to make the more complex connection to teacher quality. The

description of the SAT (Scholastic Aptitude Test) from the College Board website includes:

> The SAT and SAT Subject Tests are a suite of tools designed to assess your academic
>
> readiness for college. These exams provide a path to opportunities, financial support and
>
> scholarships, in a way that's fair to all students. The SAT and SAT Subject Tests keep
>
> pace with what colleges are looking for today, measuring the skills required for success
>
> in the 21st century. (College Board, 2012)

From that short description, the SAT is first supposed to gauge the potential for a student to be successful academically in college. It is also supposed to help ensure the fairness of opportunity for financial support for attending college. Finally, the SAT is ostensibly to measure student acquisition of a set of necessary skills. This is a large amount of extremely important information to be garnered from one test. Consequently, it is not surprising the amount of anxiety the test inspires in high school students about to take it, much less the large and lucrative industry dedicated to preparing students for it both within and outside of the school day. As such, the SAT epitomizes what has become known familiarly as "high stakes testing."

High stakes tests can be identified as any student performance measure where behind the initial or superficial intent of the test, there are significant other purposes for or uses of data obtained from the test scores. Some researchers reported finding up 29 uses of standardized tests (Haladyna, Nolen, Bobbitt & Haas, 1991). It can be assumed that many student assessments are initially for the purpose of measuring student learning or mastery of some content. That initial purpose, however, transforms under high stakes accountability into using the score to compare students against each other for ranking or other identifying purposes. Further, such rankings or categorizations on high stakes tests can be used to compare classes of students or entire schools. Even further, entire districts and states can be compared by such data.

When public relations, policy decisions, and funding hinge on such comparisons, the stakes are incredibly high for everyone involved. States and districts have taken on such stakes, continually redesigning and increasing the number and kinds of tests that will be administered to students. Policymakers are attracted to using standardized tests because they are relatively easy and inexpensive to implement (Linn, 2010). Even so, other researchers find that due to the burdens high stakes tests put on teacher and students, they are not worth the time and money

spent on them (Boardman, 2004; Falk, 2002; Paris & Urdan, 2000). Regardless of the viewpoint, it has become practically entrenched to include some measure of student performance within the teacher accountability systems.

Summary

   Overall, the review of literature reinforces the idea that there is enough going on in the context of accountability to indicate the issue of peer evaluators is worth studying. Accountability in general is evolving quickly, and while some of its components are being studied empirically, other components have been virtually ignored. Change management theory posits that looking at all aspects and entities involved in a change can help manage the impact of change. The lack of available research on the experiences of peer evaluators reflects both the newness of the latest incarnation of the role and the idea that peer evaluators are not the target or the intended outcome to be impacted by accountability designers. As part of the process of accountability, their experiences may have simply been overlooked; this study seeks to rectify that omission.

Chapter 3: Research Methods

Study Overview

Much of the rationale for this study comes from the complexity and dissonance of teacher accountability systems and the myriad of competing models being implemented. It is hoped that taking a qualitative approach of phenomenology with this study will help convey the lived experiences of a select, understudied group of individuals who are operating as key players within one of those systems. Through interviews and thorough analysis of the collected data, the researcher developed profiles that serve to answer the study's research questions.

Research Questions

1. Are there common perceived, self-reported elements in the backgrounds, interests or motivations of social studies teachers who elect to become peer evaluators?

2. What perceived expectations do these teachers report that they hold before becoming peer evaluators, and how are these perceived as met or not met as they progress through their roles?

3. What perceived challenges or benefits do the peer evaluators report that they encounter in performing their duties?

4. How do peer evaluators perceive their experiences in the role will impact their own performance as social studies educators?

5. What behaviors or abilities do peer evaluators perceive to be most useful in performing their jobs?

6. Does being a peer evaluator change a person's perceived self-reported educational philosophy or perspective?

Theoretical Framework

The theoretical framework underpinning this study is constructivism.   As pointed out by Fosnot (2005), "From a constructivist perspective, meaning is understood to be the result of humans setting up relationships, reflecting on their actions, and modeling and constructing explanations" (p. 280). With the relative newness and consequent contextual flux of the role of the peer evaluator, this study's participants are positioned as described in the preceding quote; they are building meaning as the inaugural group to undertake the role of peer evaluator. To capture and illuminate that meaning, the researcher implemented this phenomenological study with careful attention to procedures that both reflect the constructivist framework and are accepted as appropriate to maintaining potential for validity and reliability. Thorough descriptions of these procedures are provided in the following sections.

Institutional Review Board

The researcher secured approval to conduct the study both from the school district of the study participants and from the university's Institutional Review Board (IRB). Completion of a certificate of study in Protection of Human Subjects was undertaken by the researcher as one of the prerequisites to seeking IRB approval of a study. IRB approval requires strict adherence to protection protocols. These protocols are evidenced throughout the following description of the research procedures, and in the informed consent documents in Appendix A.

Research Procedures

1. Recruitment: Using procedures prescribed by the IRB and the district in which volunteers were recruited, the investigator made a short oral presentation to the group of

potential participants at one of their regular monthly meetings. Informed consent documents were circulated to all potential volunteers. (See Appendix A). More about recruitment procedure follows this list.

2. First interview: The investigator conducted and digitally audio recorded an initial face-to-face interview with volunteering participants. (See Appendix B for interview protocol).

3. Transcription and analysis: The interviews were transcribed.

4. Member-check: The investigator submitted the transcripts of audio interviews via email communication to participants to begin member-checking and verify accuracy of transcripts.

5. Second interview: The investigator conducted a second face-to-face interview consisting of the remaining questions.

6. Transcription and second member-check: After transcription, the investigator submitted the transcript of all interviews via email to participants for a final member-check.

7. Analysis of data: The investigator utilized phenomenological reduction to complete a coding and analysis of transcripts. This process is detailed later in this section of this proposal.

8. Peer-check: The investigator had a peer review the collection and interpretation of data. The peer is a colleague in social studies supervision who has earned his doctorate in social studies education within the last decade.

Recruiting for volunteers was straightforward. The investigator, with permission from the district:

- Attended a monthly meeting of the entire pool of potential participants.

- Made a brief presentation to whole group about study parameters and expectations.

- Explained "informed consent" document and procedures.

- Made selection of participants objectively from those who volunteered, with the only priority being that the peer evaluator participant was not already retired from the school system.

- Individually reviewed study description and informed consent documents with each volunteer.

- Obtained informed consent documentation.

The investigator made an oral presentation to the group and conducted one-on-one individual consultations with each volunteering participant before obtaining an informed consent signature from each. Peer evaluators meet monthly, in a large group, so it could be assumed that the largest possible pool of potential participants was present. Of the whole group, peer evaluators with a social studies background are the target study population. The district did hire some retired educators to become peer evaluators. Because of the focus of the study and the planned interview questions, it was necessary to only exclude that particular group from participation in the study.

Essentially, the investigator hoped to gain partial insight into the lived experiences of participants in the role of peer evaluator while protecting the participants that are living that role. The informed consent document follows IRB template recommendations and is written in plain language avoiding jargon or technical language. Peer evaluators agreeing to participate signed an informed consent form (see Appendix A). Participants were given two copies of the informed consent. They signed and returned one copy and kept the second copy for their records. The subjects' voluntary participation in this research posed neither physical, biological, or emotional

risks beyond those inherent in normal daily activity. Each study interview was done outside of the participant's work hours at a site of the participant's choosing. In other words, because peer evaluators work at a myriad of sites within the district on any given day, it may have been most convenient for them to meet at a work site. Regardless, if the participant chose to meet at a district site because of convenience, the interview took place outside of the participant's paid duty time with the district. Other communication with participants was conducted via personal email. Every effort was made by the researcher to create a climate in which the participants felt comfortable, and it was made clear to participants that they could drop out from the study at any time with no fear of recrimination.

As part of the informed consent process, participants should understand if there are any benefits to volunteering for the study. Although participants in this study did not directly benefit from participating, it was explained that their participation may generate information to help school systems grapple with the role and duties of the peer evaluator.

Nothing other than voluntary withdrawal by the participant was expected since qualitative data may be derived from even partial participation in the earliest stages of the study, regardless of completion. Since the benefits described above are intangible at best, they cannot realistically be weighed against any apparent risk to the participant. The interview questions were screened by university faculty and other reviewers to insure they addressed central research questions. Interview questions were written in clear language allowing for divergent answers appropriate to qualitative research. The interview protocol is provided in Appendix B. The member-check structure was in place to have participants verify accuracy of transcripts of interviews and correspondence. As such, data were monitored at least two times during the data collection portion of the study. In addition, a peer-review procedure was in place to verify that the

investigator collected and interpreted data correctly. The primary consultant for this peer review

is a fellow social studies supervisor who has already earned his doctorate in the field. The

researcher's doctoral committee, of course, has a part in the validity of the research by giving

sage counsel on the research design, research and interview questions, and data analysis.

To preserve the anonymity of subjects participating in this study, names have been concealed.

Furthermore, the participants were assigned a pseudonym for the entirety of the research process.

To report results, the researcher has used these pseudonyms and generalizations.   Although it

did not occur, a participant could have requested not to be kept anonymous.

Throughout the research process, all data have been kept at the investigator's private

residence, and when not being used for analysis or drafting results, the materials remain locked

in a file cabinet. Electronic records are kept on a removable "thumb" storage drive which is

stored along with the documents in the locked cabinet. A back-up storage drive is kept in a

separate locked cabinet. Audio recordings, transcripts, electronic records on thumb drives of any

files including e-mail correspondence and any other data gathered during the course of the study

will remain confidential and in a locked file cabinet at my residence to which only the

investigator has access. After the completion of this study, all audio recording, electronic files,

and any printed data will be maintained and discarded in compliance in the IRB regulations.

Participants

The pool of potential participants is made up of employees of a large, west-central Florida

public school district. There are 250 school sites serving over 200,000 students in the district.

40% of the students are identified as White, 29% are identified as Hispanic, 21% are identified as

African-American, and 3% are identified as Asian. The district can be classified as urban and

features a major city within it, but also contains large agricultural areas. 12% of students are identified as English learners, and 57% of students qualify for free or reduced lunch.

As noted earlier, there are approximately 115 teachers currently in the peer evaluator position; that number varies based on need. They circulate the district to conduct multiple observations on over 15,000 teachers. While there are subgroups within that population that represent almost any niche, level or content area, peer evaluators may be assigned to evaluate teachers outside their content or grade-level expertise. Even so, the desired participant pool consists of those peer evaluators that self-reported a social studies background, preparation or teaching assignment prior to becoming a peer evaluator. From the group of volunteers, 3 peer evaluators were recruited for the study. Some generalizations about the possible volunteers could already be made before selection just from the peer evaluator job description which requires them to be professional educators with current state teaching certification, possess five or more years of teaching experience, and have had some previous experience in mentoring. The definition of such mentoring has not been specified more clearly, but this experience could include having supervised intern teachers, mentored teachers under the district's older beginning teacher induction process, or serving on a site support team for a newly hired teacher.

Interview Process

Since there had not been a cadre of such evaluators before in this district, this study conducted at this time presents the unique opportunity to capture the experience of this "freshman" group. The power of this opportunity is described by Seidman (2006):

> By presenting the stories of participants' experience, interviewers open up for the reader the possibility of connecting their own stories to those presented in the study. In connecting, readers may not learn how to control or predict the experience being studied

or their own, but they will better understand the complexities. They will appreciate more

the intricate ways in which individual lives interact with social and structural forces and,

perhaps, be more understanding and humble in the face of those intricacies (p. 52).

To capture that experience, the phenomenological approach of in-depth interviews is ideal.

After recruitment, the researcher conducted in-depth interviews with the participants. The

researcher undertook the interview process as an adaptation of Seidman's (2006) recommended

three-interview structure, with the first interview focusing on background and current

experience, and the second interview re-visiting the current experience then exploring

transformations or future plans (see Appendix B, Interview protocol). While an interview

protocol had been established beforehand, it was with recognition by the researcher that probes

or explorations or other "navigational nudges" (Seidman, p.79) may have been necessary.

Possible re-statements of interview questions and follow-up probes are listed in the interview

protocol document.

Data Analysis

For this study, the data analysis follows the steps set forth in Seidman's (2006) *Interviewing

as Qualitative Research: A Guide for Researchers in Education and the Social Sciences*. The

instructions are paraphrased into step-by-step format here from the sections on "Studying,

Reducing and Analyzing the Text," (pp. 117-118); "Sharing Interview Data: Profiles and

Themes," (pp. 119-125); "Making and Analyzing Thematic Connections" (pp. 125-128) and

"Interpreting the Material" (pp. 128-130):

1. Read transcripts of interviews and mark with brackets passages that are interesting.

2. Take all the bracketed parts and cut/paste them into a new document (reduction).

3. Read the new reduced version of transcripts and underline most important parts.

4. Re-write the transcript in first person using ellipses for gaps and brackets for when you change words.

5. Re-read this new version and begin to label categories and portions, making sure to code each portion to remember where it came from in the transcript.

6. Cut/paste these portions and re-organize into files by category.

7. Re-read transcripts now by file and begin to write the explanation of what meaning this has for you.

Obviously, these instructions reflect the researcher's intent to conduct the data reduction and analysis manually, although there do exist software programs that ostensibly can assist with categorizing, coding and re-organization. By the time these seven steps were completed, the researcher was able to produce output in the profile format recommended by Seidman (2006):

> I have found crafting profiles, however, to be a way to find and display coherence in the constitutive events of a participant's experience, to share the coherences the participant has expressed, and to link the individual's experience to the social and organizational context within which he or she operates (p. 119).

Validity and Reliability

In the description thus far of the steps to be taken and the intended output of the analysis, there remain some considerations which should be highlighted to fully convey the study design's commitment to valid phenomenological investigation and revealing the lived experiences of the study participants. As noted by Seidman (2006), the phenomenological interviewer's task is to present the experiences of the participants in "compelling enough detail and in sufficient depth that those who read the study can connect to the experience, learn how it is constituted, and

deepen their understanding of the issues it reflects" (p. 51). Toward that end, the process of phenomenological reduction must be gauged carefully so as to not eliminate important information. The researcher must also, however, achieve a balance between recognizing what is important against trying not to inject their own prejudices and interests in finding something important that is not. Seidman explains that "the repetition of an experience that was already mentioned in other passages takes on weight and calls attention to itself" (p. 127). Even so, Seidman admits that he looks for certain "things": conflict, hopes, beginnings, middles, ends of processes, frustrations and resolutions, isolations or collegiality, class, ethnicity, gender, hierarchy and power. (p. 118)

Considering the work of Guba (1978) and his adoption of terms like *convergence*, the researcher approached analysis of the interview data set with three key guidelines for determining if a resulting classification system is working or not working and whether the analysis is effective. These guidelines represent this researcher's internalization of what has been learned about qualitative study and validity throughout the doctoral program. The first guideline involves the usefulness and completeness of the categories derived—are they consistent and sufficient to provide a complete picture? To test that guideline, too much overlap or too many "left over" or "left out" data points would indicate insufficient categories. This guideline logically derives from the concepts of convergence and divergence (Guba, 1978) or internal homogeneity and external heterogeneity (Patton, 2002), among others. The second guideline essential to this study will require that the categories derived are identifiable by someone else—can a second observer of the data affirm that categories make sense and that data has been correctly assigned in these categories? This guideline is tested through the peer review process where a colleague will affirm (or challenge) the researcher's initial creation of categories

and assignment of data to categories. The third guideline the researcher may have found

essential in data analysis involves the participant giving credibility to the derived categories—

does the participant find the categories and assignment of data to categories representative and

sufficient? In other words, it could have conceivably become necessary to verify certain

categories through a member-check process, perhaps in the second interview or in an additional

interview, however, that did not become necessary in this study.

Many of the same procedures discussed so far for increasing validity will help build toward

reliability as the concept can be applied within qualitative research. Reliability, especially in a

phenomenological study, can be a challenge since the goal of that research is to reveal the lived

experiences of individuals without necessarily having the goal to generalize the research to a

larger population. As the field of qualitative research has evolved, that distinction and different

conceptions of knowledge among qualitative researchers has led to differing opinions of what is

reliability. "The definition of reliability in qualitative research differs between positivists

(traditionalists and modernists), constructionists, and the critical researchers, but there is

concurrence in the need for trustworthiness, accuracy, and dependability of research findings"

(Lewis, 2009, p. 7).

Role of the Researcher

For a researcher engaged in qualitative study, Patton (2003) recommends that the investigator

"carefully reflect on, deal with, and report potential sources of bias and error" (p. 51). He goes

on to list several other techniques that help buoy reliability or dependability including systematic

data collection procedures and triangulation. These techniques work to put judgment of

reliability onto the reader (Creswell, 1994), which makes the requirement of thick description

and reporting of biases and assumptions critical in order for that reader to have the proper

information to make that judgment. Some strategies for trustworthiness, accuracy and dependability have already been addressed here. To work toward a high degree of dependability, the researcher is including a brief biographical sketch in Appendix C that provides information about the role of the researcher. This inclusion acknowledges the constructivist perspective that in the role of the principal investigator who has collected and analyzed data, the researcher – much like the participants—is influenced in that role by her own contextual reality. That, in turn, can affect the endeavor of investigation. In phenomenological research, the term *epoche* refers to a practice that "requires a researcher to explore his or her assumptions about the topic to be aware of their influence in the research" (Romanyshyn, 2010, pp. 280-281). Creating the biographical sketch allowed the researcher in this study to undertake this epoche process, which was essential in subsequently making study design decisions, creating the interview protocols and explaining the study's context.

In the spirit of epoche, the biographical sketch is written in narrative format to more fully and personally reveal who the researcher is and to disclose what connection the researcher has to the content, context, and participants in the research. In the biographical sketch, the researcher identifies some of her assumptions, predispositions, and potential biases. Reporting that information within the sketch and throughout the study should serve to help readers judge for themselves the reliability or dependability of the research.

Special Considerations

As asserted earlier, this study, with its safeguards in place, ostensibly posed no risk or threat to the specific participants. A reader may note, however, lack of depth in description of the study setting compared to what might be recommended by qualitative study theorists. Keeping in mind the high-stakes nature and funding issues associated with accountability systems, it is

reasonable that a school district may want a degree of anonymity to avoid risk.   During the process of gaining approval to conduct the study through that district's established procedures for research studies, discussion with district personnel on the issue led to a careful approach, in particular to protect the participants in the study. Thus, the needed reference to the district's website has been omitted earlier in this document, and later, some selections from a participant's transcript have been unattributed even to a pseudonym for the protection of the participants. Ultimately, however, these protections do not detract from a detailed description of the lived experiences of the peer evaluator, and the goal of richness of description is met through the participants' words, emotions and perceptions.

Chapter 4: Results

Organization

This study focuses on the lived experiences of social studies teachers serving as peer evaluators.   To reveal those lived experiences, in-depth interviews were conducted. The interview questions were crafted to parallel the central research questions. The questions are:

1. Are there common perceived, self-reported elements in the backgrounds, interests or motivations of social studies teachers who elect to become peer evaluators?

2. What perceived expectations do these teachers report that they hold before becoming peer evaluators, and how are these perceived as met or not met as they progress through their roles?

3. What perceived challenges or benefits do the peer evaluators report that they encounter in performing their duties?

4. How do peer evaluators perceive their experiences in the role will impact their own performance as social studies educators?

5. What behaviors or abilities do peer evaluators perceive to be most useful in performing their jobs?

6. Does being a peer evaluator change a person's perceived self-reported educational philosophy or perspective?

Just as is reflected in these research questions, the interview questions led participants to generally go from their experiences before becoming peer evaluators through to their predictions

of the future after completing their tenure as peer evaluators. To maintain coherence for the reader, the semi-chronological organization is retained here in the presentation of abridged profiles that were crafted from the interview answers. These profiles will assist in corroborating the subsequent themes identified by the researcher as the outcome of phenomenological reduction. Those themes are presented after the profiles. Finally, those themes are used to assist in presenting the findings of this study in regards to the research questions.

Profiles

While using third person to introduce the participant and separate the significant portions, the bulk of each profile is written in first person as recommended by Seidman (2006) because "using the third-person voice distances the reader from the participant and allows the researcher to intrude more easily than when he or she is limited to selecting compelling material and weaving it together into a first-person narrative" (121).

The profiles reveal much about each participant, but some demographic data about the group may be helpful for the reader to fully identify with the participants' lived experiences. It has already been noted that the participants all have at least five years teaching experience and have some prior leadership experience as part of the application requirements for the peer evaluator role. None of the three has yet reached twenty years in education. There are two females and a male in the participant group, which reflects the larger possible participant pool where females consistently outnumber males. Two of the participants are Caucasian-American and one is African-American. Their exact ages are not revealed to protect their identities. All three of the participants have taught classes with a specific social studies designation, meaning that all three have experience in middle or high school which is when such designations begin, but one participant had elementary teaching experience as well. All three reside in the district in which

they work, came to the peer evaluator role from within the district, and among the three of them, there is a parent who has at least one child attending school in that district. All three have been peer evaluators since the first year of the program, which is now in its third year.

Participant A: Abigail. Abigail had teacher leader experience at her school. In that role, she came across some difficult personalities and wanted experience in how to work with them. Abigail envisioned that the peer evaluators would take ideas from the vast number of classrooms they visit, and be able to share them with other teachers. Abigail explained that she has a "helping nature," and expected that would serve her well in the role of peer evaluator.

I felt that the impact I was able to have in my building could be strengthened by having more opportunities to practice my skills at coaching and evaluating other people as well. The difficult personalities in the building really made me uncomfortable, and I figured it was easier for me to learn how to work those if I had more practice at it…The idea that I always had was that it was a district-wide PLC (Professional Learning Community) and that the [peer] because they see so many classrooms would be able to take ideas and share them with other classes….I like helping people anyway.

Once she was in the role, Abigail found that the collegial, friendly system she anticipated did not materialize. Having too many people to observe and evaluate presented time constraints so that coaching could not really happen. There was some animosity, especially from teachers who knew Abigail before she entered the role of peer evaluator. In some cases, the animosity even provoked attacks on her credibility. In her second year as a peer evaluator, Abigail worked less with people she already knew. She sees that as a benefit, because it helps her in observing most objectively and in being able to provide feedback.

I thought it was going to be a little friendlier...I didn't think initially that we were going to be met with as much resentment and that people would be willing to listen to some of the advice and feedback that we were giving....Because of the format of everything at the beginning, we had so many teachers we couldn't have the conversations that we needed to, and we didn't get a chance to build rapport...In the second year, I think I had an advantage when the people [don't] know who I am. If they know who I am, then they already have preconceived ideas about who I am and what my beliefs are, and it changes that role of mentor and coach into something different.

In her three years as a peer evaluator, Abigail found that multitasking skills were most essential to fulfilling her role, but that being a reflective person also helps. Being a peer evaluator affords Abigail a view of a bigger picture and connections she had not seen before. Those connections just reinforce the holistic approach to teaching that Abigail believes will help her upon return to the classroom, although she is not sure if she will return to the classroom when her term as a peer evaluator concludes.

If you can't multitask, you can't get anything done. There's too much to do and no work space and you're running around a lot so you have to be able to work in a bunch of different locations...I've been thinking about the social impact that I have on things like how can I get people to hear me. I can't get them to hear me if I just come in like here are the facts and this is the evidence and so I have to figure out how it makes them work and being reflective is something that's really key, I think. I think that more successful people are more reflective than not...I got exemplary in reflection.

I just see how things are connected a little differently than I did before. I didn't really see how we all played into the bigger picture but now I do and how each teacher and

each classroom has a role that's bigger than them...I like the holistic idea of teaching a

child...I'm going to be a better teacher. I'm taking that holistic approach to what I do in

lessons now...Never teaching anything in isolation, which I might have been guilty of

before...but I don't know what I'm doing (in the future). I always joke that I don't know

what I want to be when I grow up.

Participant B: Belle. Belle was a classroom teacher looking for something else, somewhere else in her career when an administrator she consulted suggested the peer evaluator role. This administrator convinced her that the peer evaluator job would provide her with unique experiences and exposure to help Belle meet her professional goals. Before beginning as a peer evaluator, Belle anticipated working with teachers to help them be better, but she also knew that the job would entail evaluation. She viewed it as a leadership role with more autonomy than that of a classroom teacher. Belle had mentored many beginning teachers through existing district programs at her site, and had conducted staff development trainings. She believed both those experiences would help in the peer evaluator role.

I was actually looking to apply for another job in [another city]... [My administrator]

suggested that I apply for [the peer evaluator position]. All I saw was announcements

about applying for this position, and I didn't really know what path that would lead to so

I wasn't really interested….After talking with him, he really laid it out for me what this

job could lead to, the experience I could get from it, the exposure. I trust him personally

as a friend and professionally, and I went for it and so that's what drew me to it. It was

the advice of someone that I trust.

I anticipated working with teachers to help them to be better teachers. I anticipated

more of a coaching role. I knew I would be evaluating so I definitely knew that was part

of what this job would entail. I anticipated having some autonomy and not being driven by a bell schedule...and I anticipated this being somewhat of a leadership role within the district. I've done a lot of mentoring of new teachers so I thought...this would be a good fit for me. I consider what we do with the rubric as professional development and I do enjoy that part of what we do when I get to train so I thought that was a good element. I was in the midst of going through [earning an advanced degree] so [I thought] about setting myself up for what could be next after I'm done with school... I thought this would be a good fit for leadership.

While in the role of peer evaluator, Belle has found some of the autonomy she expected, and she feels she is getting the leadership experience she had hoped for, but does feel limited in her ability to coach and help teachers. Negative perceptions from those she has had to evaluate are a challenge for Belle in performing as a peer evaluator. But she also sees a benefit for herself in that challenge, and believes it is strengthening her leadership skills. Additional benefits of the job are getting to see so many good teachers and getting to see a bigger picture view of teaching and the district as a whole that she did not have from within her classroom.

I definitely have gained an even deeper sense of autonomy through this job...It's really stretched me [with] little things like paying attention to details, managing my time. It's strengthened me in terms of leadership, how to deal with difficult people and how to have difficult conversations.

I sometimes don't feel like I'm as much of an asset to teachers as I [thought] I would be going into this....This job is evaluative so that hasn't quite met my expectations and sometimes that does carry into my overall satisfaction of the job. I have to feel like I am making a difference. On some days, I feel like I really really am, and then there are days

when I see a struggling teacher and I want to give so much more but the job limits me...I want to do more, but I can't. This job isn't meeting my expectation in that way of coaching and mentoring teachers.

Sometimes it's the image that people have...There's a perception about our jobs as peer evaluators and that's a huge challenge because I have to sometimes fight against that in order to get teachers to hear me. One thing I can say, though, is a lot of teachers are very open and transparent with me about their negative perception, and I think that sometimes...becomes a benefit for me because when they are real and they're transparent, I can make that human connection with them...It's almost like a Catch 22. I want teachers to be open and transparent with me about their perceptions because it...guides me in how I interact with them, but then sometimes it's like I know that's what I'm fighting against and I have to spend an extra 15 minutes just to listen to them getting it off their chest or there's a wall built up...That is probably the biggest challenge is the perception that we're fighting against.

One of the benefits that I often share with teachers, and I truly do mean this, is all the good teaching that I get to see and I kind of see it as a pay it forward job where I can see good teaching and I can pay it forward and share it with another teacher...They do sometimes seem a little bit more receptive to that because it's not me walking in as the boss saying this is what you need to do, but it's like, hey, I'm paying it forward...It is a benefit to get to see cool things in the classroom and it helps me to grow professionally.

I think sometimes in classroom teaching you're kind of like [tunnel vision]. You only kind of see what's going on in your classroom and then your school building, but as a peer, what I've seen is I do get some of the big picture. I like that...it broadens my

perspective that it's so much more than my classroom at my school. My perspective has been opened so much more as to the big scale district level… and I know it's because of this job.

As a peer evaluator, Belle found that being very organized and a good time manager are essential because of the workload and scheduling issues associated with the role. She has always been a good researcher, and used that skill in earning advanced degrees. She thinks that has been an asset because she is always searching for more strategies to help teachers. In order to help teachers, Belle has also relied on her ability to read people in order to conduct the communication the role requires.

> *I know my type A personality really helps me a lot, so being organized has really helped a lot. I thought I was good at managing my time before becoming a peer, but this has really bumped it up to another notch. Time management is everything…it affects your productivity…if you're not good at managing your time, then it affects your effectiveness in terms of getting feedback to teacher [in a timely way], in terms of scheduling, when your choose to go where you choose to go, how long you're going to choose to sit on a conference when you know you have stacked up so much more for the rest of the day. ..I think if you haven't experienced a job like this then time management sounds kind of tedious, but it's everything.*

> *I don't mind learning and researching, and I really do think that helps me, my desire to go find new information to research…because it deepens my toolbox so that I can be able to provide more resources and help for other teachers. I think [it's] my natural ability to want to learn more and research.*

I think the ability to read people and to connect is again one of those things that sound tedious, but it is so critical. You have to know how to read body language. You have to know when enough is enough and you just kind of end the conversation if it's a difficult conversation. I feel like I've minored in psychology or behavior analysis in these three years of being a peer because I'm constantly reading people and I'm constantly listening to key words and using that as an element to build on a conversation...I don't know if that's an ability I always had or if I just picked that up as a result of doing the job, but being able to read people and connect with people, it's huge, and I think that's helped.

Three years of experience as a peer evaluator has reinforced Belle's philosophy and goals to be more of an advocate for education—for children and teachers. She feels some frustration at wanting to do more for both. Belle believes that her experience as a peer evaluator will make her a much better teacher and teacher leader, but going back as a classroom teacher is just one option among others.

It's just really...sparked up even more of a desire for me to want to do what's right for kids. At the end of the day, this job has put me in a position of being uncomfortable because I want to do more, and I want to do more in the district but even beyond the district...I had never thought about rights for teaches and supporting teachers on the level that I have now as a peer. I don't think my philosophy has changed much. It's just really become that much stronger and sometimes even to the point that it frustrates me because I want to do more and I think some things should be different.

A lot of people ask, "What are you going to do after this job?" and I've never said, "I'm not going back," because it's an option. It's a realistic option for me. I think this job will put me in a position of being a much better teacher. I know that for a fact

because I've learned so much about best practices. I also believe that it will put me in a position to be even more of a teacher leader on my campus and to not be afraid to think outside of the box because this job has given me the confidence to where I know I can do that… [Going] back to the classroom is one of my options…I could definitely see myself writing trainings and doing things for social studies as it relates to best practices in the classroom.

Participant C: Charlie. Because Charlie took a unique path into a teaching career, he had some real struggles in the first years of teaching. He was drawn to the peer evaluator position hoping he could support teachers to avoid some of the struggles he had faced. Charlie actually was more drawn to the mentor teacher position, but accepted a job as a peer evaluator still hoping that along with evaluation, he would still be able to be in a coaching role. With experience in different grade levels, staff development, and a school-based teacher leader role, Charlie thought he had a lot of experience packed into a relatively short career. He felt this would be an asset and lend him credibility in undertaking the peer evaluator role.

[Starting] off as a teacher, I did not have an education degree. I had a social science degree. I wanted to be a teacher…decided that kind of late; I was already in my major. I could have changed my major, but I was on an academic scholarship, so I [said to myself] "Well, let me just get my degree and I'll go the alternative certification route." I did that because, you know, just caring about kids. That was really my main motivation. [Coming] into [teaching], though, it was a culture shock to say the least. I had not been in [a school] since I was a student…so that was very strange… I didn't feel supported and really just felt like I was going through my first year kind of blind…not knowing what to do and not knowing how to do it, how to get better.

When [the peer evaluator] opportunity came up, I looked at it as this is a way to support teachers…so that we don't have as many people leaving the profession and so that we have people getting better not just staying sort of stagnant or plateau on their skill level. Like, [just repeating] year two over and over. That was my reason.

I applied for both [peer and mentor positions], hoping that I would be a mentor… [When] I came into the [peer] job and accepted the position…I knew the evaluation piece was there, and that I would be evaluating teachers' performance and that sort of thing, but we had also talked in the beginning about there being an opportunity to support teachers.

Even though I didn't have a large number of years of experience, I had a lot of experience in the time that I taught…I knew how to teach Reading so when I became a social studies teacher I knew how to embed comprehension thinking strategies and different things into that role. I have a Master's degree in Educational Leadership and I did a lot of work in looking at best practices…I thought I was learning a lot about leadership, and I did, but I learned a lot about teaching practices and pedagogy in that, and that made me a better teacher. [In my school leadership role] and supporting teachers and doing walk-through [observations] and giving them feedback…conducting trainings with teachers in my department and then for the district, I felt like I had, again, not a vast number of years, but a lot of experience in that short amount of time.

Now in his third year as a peer evaluator, Charlie has seen growth toward the peer evaluator system meeting his expectations of supporting teachers. While there were initial limitations and negativity, Charlie still believes he is helping teachers to reflect on their practice and improve. Being a coach and mentor as he expected, however, has not materialized, and although limited

coaching happens in post-observation conferences, Charlie now realizes that he is truly in an evaluative role more than anything else.

> *I've seen growth having done this from year one until now…where when we came in, this was not very well received by all teachers. Not by everybody; not that everyone was negative, but overall there was more of a negative connotation and a negative feel for this whole initiative when we first started…I've seen tremendous growth in that. I feel like what we've done is we're encouraging teachers to reflect more on their practice, and I learned that's what this is all about and that's what we should be doing—getting teachers to think, "Is what I'm doing effective?" [But] the piece of really being that coach and that mentor to teacher, that has not happened and that, I don't believe, is something that will happen. I think that vision is kind of…that's kind of gone. I [think] this role of the peer evaluator, I think that now it is an evaluative role. There is coaching in the post conference. After seeing the lesson…you try to get the teacher to reflect and ask the questions about, "What went well? What would you change?" and then try to hone in on specific things so that the teacher comes up with a lot of it, but I don't have the opportunity to go in and model lessons, help them plan the lesson, coach them further.*

In that role, Charlie has encountered challenges working with teachers who instead of being reflective, focused on the differences between themselves and him, questioning Charlie's ability to evaluate them. He tries to counter that challenge by using people skills and emphasizing that he is just a fellow teacher trying to share ideas and information. Garnering those ideas and information is one of the benefits of the job—getting to see a bigger picture and being able to see from others' examples where Charlie could improve himself.

That insight is also part of what Charlie perceives as key in performing the peer evaluator role. Charlie believes you have to be a good teacher who recognizes good teaching. In particular, he thinks that truly understanding learning and student engagement helps a peer evaluator be effective in rating teachers and giving feedback. In order for teachers to be receptive to the feedback, Charlie believes that several characteristics help establish rapport: being personable, empathetic, honest, and approachable.

> *It's all fine if they agree with the feedback…but as soon as they think that something is graded lower than their expectation or their evaluation of it, then it's, "Oh. Well, you're not qualified because you've never taught this" or "you've never been in this type of a school," or "you've never experienced this." [In this] position…you have to constantly be aware that you're not their boss but you have to convey that to them. They need to feel like you're not coming in on a power trip, as a know-it-all. I work really hard to…convey to them that I'm a teacher, I just get to watch a lot of other teachers teach and steal ideas from them and share them with you…Doing that has helped me overcome some of those issues that could have been issues. I've gotten better at it with time.*

> *I've learned that experience does not necessarily equate to years. I have probably gained in a year doing this job five years of teaching experience…understanding the bigger picture of how things are [that] from the perspective of a teacher I could not understand…The first thing you have to [do is] build rapport with teachers up front, so you need to be…someone who's friendly, someone who's approachable. I don't think that teachers are willing to listen to you if you don't convey that. If they don't feel like you're there to support them and to help them, they're not going to open up and reflect honestly, and they're probably not going to be receptive to your feedback…I think you*

need to be a good teacher to do this job. You have to really know instruction. You have

to have a good understanding of what is learning and what is engagement.

Three years as a peer evaluator has mostly affirmed the educational philosophy Charlie had going into the position. In working with struggling learners, Charlie always felt that he should retain high expectations for his students. He believes all children can learn, and this experience has reinforced that students need rigor, challenge and differentiation. Added to Charlie's existing philosophy, he now sees assessment as key in assuring that all students are learning. That will influence him when Charlie returns to a role as a classroom teacher, where he intends to focus on student self-assessment and critical thinking. He is not sure, however, if he will return to his previous teaching assignment or explore other levels or content areas.

It's more affirmed feelings that I had. I've worked in Title 1 schools, and I've worked

with ESE students, and I've worked with very challenging groups of low-level readers.

I've always felt like I don't need to dumb it down for them. I need to give them support

so that they can be successful, but I need to have high expectation and I need to challenge

them to think. In this job, I've seen that is really what needs to be happening. [I have

observed] a lack of rigor, a lack of expectations...Teachers tricked [into] thinking,

"They're busy. They're answering. They got all those questions right. Yes, they were

engaged. They can answer every question right." Really, what does that say? It says

something about the questions we're asking them. If they can answer every one right and

they're never wrong and they never really have to think, they aren't really challenged.

Everything that we do, we need to be really evaluating. What I've really learned

about assessment is that where we ultimately need to go with it, it's not just about the

teacher knowing. The kids need to become more aware of their own learning and that we

need to do more self-assessment...I think doing this job honestly makes me want to go

teach other topics and content areas, like in [another level]. I love watching [other

level] classes. The content really interests me and it makes me want to go teach it even

more.

Themes

Several themes emerged from the interviews as illustrated in the profiles. Since the interview

protocol followed a chronological pattern, the themes are illuminated best linked to the

chronology of the peer evaluators' experience. Specific excerpts from participants' interviews

are used here to emphasize or serve as evidence of the themes. To assist in developing these

overarching themes, the principal investigator took the significant portions that emerged from the

transcripts and laid them out across a ten-column matrix. Each column represented the gist of

the interview questions, for example: "How Drawn to Job," "Expectations," "Background or

Qualifications," "Challenges," etc., as seen in Figure 1.

Laid out in that fashion, the portions were easier to essentially portion again to perceive a

theme. It was also possible to portion the columns together into a chronological organization

that solidified the themes. In looking at the peer evaluators' attraction to the job, background

experience, and expectations, the phrases captured from the transcripts led to the theme of

Idealism. Looking at the perceptions of how expectations were met or not met and challenges in

the role, the transcript portions led to the theme of *Non-Collegial Reality*. By examining the

portions of transcripts relating to benefits of the role, and those relating to assets to performing in

the role, the theme of *Valued Experience* emerged. Finally, in looking at the portions of

transcripts relating to any changes in the participants' philosophy of education and future plans,

the theme of *Residual Optimism* emerged. An abridged version of the matrix is provided in

Figure 1, since the actual matrix involved a huge swath of butcher block paper hung on the wall with pieces of transcripts taped to it.

How drawn to job	Expectations	Background or Qualifications	Expectations met	Expectations not met	Challenges	Benefits	Asset or Skill	Philosophy	Future
Helper	Mentor	Varied experience	Seen good teaching	Did not get to coach	Teacher negativity	Big picture	Time management	High expectations	Maybe go back to classroom
Support	Coach	Advanced degrees	Autonomy in scheduling time	Thought teachers would be more receptive	Time	See good teaching	People skills	Holistic	One option of many
Exposure	See good teaching	Site leader	Better in second year	Want to be more of an asset	Wanting to help more	Pay it forward	Be a good teacher	Use assessment more	Teacher leader
Learn to deal with difficult people	Help teachers	Trainer		Teacher resentment	Frustrated	Sharing good ideas	Rapport		Go to other level
	Evaluating	Mentoring new teachers			Too many teachers to build rapport	Reflect on my own teaching	Organized		Help more
	Leadership role	Like helping people			People who knew me were worst	Better at difficult conversations	Type A personality		
	Share ideas					Objectivity	Multi-tasker		
							Reflective		
Entering the role: Idealism			In the role: Non-collegial reality			Effects of role: Valued Experience			Future

Figure 1: Matrix of Theme Development

Idealism. The term "idealism" is used here as objectively as possible. The principal investigator makes no judgment on whether idealism is proper or foolish. It is used simply to indicate that a preponderance of phrases and portions of the transcripts about what the participants perceived the role of a peer evaluator was to be illustrates that the participants had an "ideal" in mind for the role, and that ideal included a degree of professional altruism. All three participants indicated in some way that they were drawn to role or felt they would be suited to it because they had an ideal of being able to support or coach teachers.

Abigail's idealism is most evident when she contrasts it to her reality, which will be explored in the following section. She had a vision of a kind of "district-wide PLC (Professional Learning Community), where peers and teachers would take ideas and share them, but then: *I found myself in a different role and trying to figure out how to convince people that what I was going to help them with would be a good thing.* Abigail imagined people wanted help, because she herself welcomed it. She saw help as a "good thing," and liked to stretch professionally: *I felt that the impact that I was able to have in my building could be strengthened by having more opportunities to practice my skills at coaching.* While Belle saw leadership opportunity in the peer facilitator role, her anticipation at the outset was more of a coaching role: *I would anticipate experience working with teachers to help them be better teachers…I thought that peer evaluators would truly be a resource for teacher and could get in more of that coaching part…I have to feel like I am making a difference.* Charlie held to the ideal of coaching and supporting even after taking on the role of evaluator: *I even attended the mentor training the first year because…there was a group of us that were possibly going to…not only evaluate some teachers but also go in and mentor and coach other teachers.* Even with past leadership experience at their site as a common element in their backgrounds, the perceived ideal of all three study participants was that they would be helping their colleagues.

Non-collegial reality. Dreams acted upon often come up against the reality of action. If the theme of idealism has been affirmed as participants sought out and began the peer facilitator role, then the subsequent theme of what they experienced in the role can be affirmed as non-collegial reality. The participants' responses in how their expectations were met or not met, and what challenges they faced undertaking the peer facilitator role illustrate that the collegial and helpful ideal they envisioned for themselves was at least to some degree transformed in action.

To avoid too much foreshadowing or indications of bias, the principal investigator has reserved portions of the transcript indicating negative receptivity from teachers until this point, and they are included not to judge the success of the peer evaluation program, but only to illustrate the current theme.

Belle explains: *Probably after six months of the job and reality set in, and we realized we weren't doing much coaching and offering as much assistance…[We were] wanting to be able to help more and there's just some limitations to the job where you can't.* Negative teacher reception and time are the two major elements in this theme. Both Belle and Charlie noted that it takes additional time or effort just to overcome negativity in working with a non-receptive teacher. Belle notes: *I have to spend an extra 15 minutes just to listen to them getting it off their chest.* As noted by Abigail: *We had so many teachers we couldn't have the conversations that we needed to, and we didn't get a chance to build rapport. You can't really coach without a rapport with somebody. The first year was kind of a wash.*

Negative reception from teachers proved to be personally upsetting to some of the peer evaluators, so the following excerpt remains unattributed here to protect the participant who is describing the challenges:

> *People thinking you're out to get them, and that I take as a personal attack because I don't want to hurt anybody. I want everybody to be good. I want everybody to get those "Accomplished" and "Exemplary" [ratings], but if I offer advice, it's seen as detrimental to them instead of "Hey, I've seen this somewhere else and it's really cool and a great idea and you might want to try it."*

When asked to elaborate on how it was a personal attack, the participant shared:

Body language, not listening, attempting to damage my credibility with things that have not actually happened, and they portray it as something that was truthful, calling supervisors, talking about me in public places…all not nice things. Nothing I have control over. I can't change those things.

Explained in practical terms, Belle points out the difference in coaching and evaluating in teacher receptivity:

In my past, a lot of times when I'm coaching and mentoring teachers, they want to be in that position. They want to receive the feedback. Now in my role as a peer [evaluator], they're not always on the [receptive] end, but that's the job that I have to do. That's the way it goes. That's tough, too, when they don't want to receive it and they're just kind of there going through it because they have to.

Valued experience. Despite the challenges, the participants' responses about the benefits of their experience and what skills they perceive as being essential to the peer evaluator role indicate they each find growth in what they've encountered. This illuminates the theme of valued experience. It is important to note the use of "valued" instead of "valuable" experience. To avoid the appearance of bias in assigning value, the principal investigator chose the term "valued" to better indicate that it was of value only through the lens of perception that the participants used to describe the value.

Keeping in mind that Charlie noted earlier, "*I have probably gained in a year doing this job five years of teaching experience,*" it is clear that he has valued his experience. He expresses satisfaction with the resilience he has acquired in the peer evaluator role: "*Doing [this] has helped me overcome some of those issues that could have been issues. I've gotten better at it with time. It was more difficult in the beginning of the process…but now it's just more natural.*"

Charlie valued the opportunity to look from a different viewpoint which led to "*understanding the bigger picture of how things [are] from the perspective of the teacher I could not understand [before]."* A standout concept from this new viewpoint was how he used and understood assessment in the classroom:

> *I don't feel like I was very proficient at assessment and using ongoing assessment every day in my classroom as a teacher. I think I did more assessment of the class as a whole, and I didn't really understand what that meant until I took this job and I got to sit back and watch and see how you need to be watching what kids are doing throughout the lesson…I've kind of gotten a better understanding of how the way that we plan, how we prepare for lessons, how much of an impact that has and how it connects to what we actually do in the lesson and just the bigger picture in general.*

The valued ability within the peer evaluator role to see things from another perspective or the "big picture" is echoed by the other participants. Abigail notes, "*I just see how things are connected a little differently than I did before. I didn't really see how we all played into the bigger picture, but now I do and how each teacher and each classroom has a role that's bigger than them."* Belle places value on this as well: "*Sometimes in classroom teaching you're kind of like [tunnel vision]. You only [see] what's going on in your classroom and [your] school building but as a peer…I do get some of the big picture….I like that…it really broadens my perspective."* For Belle, that bigger picture included a broadened perspective about the classroom teachers' association. "*I never paid much attention to the [union] like I do now. I know it's because I'm a peer and I see everything the [union] does…if I were in the classroom, I don't know if I would have paid attention that much."*

Residual optimism. The participants in this study echo their original idealism and remain optimistic as they describe how the experience will influence them after their tenure as peer evaluators ends. In the matrix in Figure 1, this optimism resides loosely in the "Future" column, although there is some overlap with "Valued Experience."

This optimism and excitement about what comes next is revealed by Belle when she explains, "*I had never thought about rights for teachers and supporting teachers on the level that I have now as a peer. I don't think my philosophy has changed much. It's just really that I've become that much stronger...*" Her experience seems to have left her invigorated about her career:

> *I also believe that it will put me in a position to be even more of a teacher leader on my campus and to not be afraid to think outside of the box because this job has given me the confidence to where I know I can do that. I know it will make me a better teacher and has even strengthened my confidence in being a teacher leader.*

Charlie almost seems impatient to take what he has gleaned and put it in to practice, optimistic that strategies can be implemented to improve student learning.

> *When it comes to my educational philosophy or kind of what I believe that all kids can really learn, but that we do need to provide more, maybe like focused instruction for kids, in different groups and really differentiate what we're doing...Get kids to become more aware of their own learning because that's how they're going to initiate improvements...Just going back and make sure that I'm planning engaging activities for all kids...Yes, they need to be interested and then they also need to be asked to think.*

Perhaps residual optimism is summed up best by Abigail's future –oriented comments:

> *We're all in it together, but I didn't really see it until this...I'm going to be a better teacher. I'm taking [that] holistic approach to what I do in lessons now. Thinking*

through future lessons, how is it going to impact my kids? Also, how are they going to be

able to connect my skills to other skills that they're going to face?

Findings

The first research question of this study is, "Are there common perceived, self-reported elements in the backgrounds, interests or motivations of social studies teachers who elect to become peer evaluators?" The professional backgrounds of the participants were similar, but that could partly be attributed to the job credentials required to become a peer evaluator, i.e. previous mentoring experience and at least five years of teaching experience. Evidence from the interviews, however, most supported that there were common perceived elements in interests and motivations of the participants as they undertook the role of peer evaluator.

The general theme that emerged and revealed the commonalities was *idealism*. The participants reported idealized, even altruistic, interests and motivations to improve their own teaching and the teaching of others. All three participants already had some experience in mentoring teachers previous to entering the peer evaluator role. The study revealed their common perceived interest in continuing such mentoring through the role. Terms associated with an interest in mentoring that repeatedly appeared during interviews included "helper" and "coach."

The second research question was, "What perceived expectations do these teachers report that they hold before becoming peer evaluators, and how are these perceived as met or not met as they progress through their roles?" With the idealism noted above, it is not surprising that the participants held rather idealistic expectations before entering the role of peer evaluator. Although aware of the responsibility for evaluation indicated by the role's title, the participants shared perceived expectations that the actual job would entail more coaching or mentoring in the

spirit of professional reflection and improvement. As the lived experience in the role progressed, participant responses indicated they faced a common theme of *non-collegial reality*. Rather than being embraced by the teachers they were observing and rating, the participants found they were no longer seen as colleagues, and their idealized vision of mentoring was not going to materialize.

While the less-than-welcoming response from some of those they observed could somewhat be explained to a human reaction to something new, the interviews revealed that the study context was indeed influenced by an element of non-collegiality that did not meet the participants' initial expectations. Simply put, the participants had wanted to be more of an asset to teachers, but teacher attitudes and some constraints of the evaluation system itself prevented the participants from fulfilling their expected coaching role. A list of paraphrased parts of the interviews where non-collegial reality surfaced would include "not getting to coach," "thought teachers would be more receptive," and "teacher resentment."

The third research question was, "What perceived challenges or benefits do the peer evaluators report that they encounter in performing their duties?" Despite reporting challenges like less-than-welcoming interactions and time/workload stressors, the theme that emerged from participant responses was *valued experience* because the participants shared the reported perception that even negative experiences --- like being verbally attacked, for example --- would ultimately be valuable and help them grow professionally. That perception harkens to the old adage that what doesn't kill you makes you stronger. While they cope with a range of negative receptivity from the teachers they observe, and juggle intense workload schedules, the peer evaluators reported valuing the beneficial experience of getting to see teaching strategies they

would like to use or share, and getting to see the bigger picture of what happens outside their own classrooms.

The phrases from the interviews that through repetition most supported the perceived challenges encountered by peer evaluators include "not enough time," "too many teachers to build rapport," and "teacher negativity." Phrases to represent what the participants perceived as benefits of the experience include "seeing good teaching," "paying it forward," "better at difficult conversations," and --- overwhelmingly --- "getting to the see the bigger picture." The perceived benefit of seeing the bigger picture of the system in which they work was so dominant in the participants' interviews that it merits designation as a phenomenon all by itself.

That ability to see value and benefits relates to the next research question that asks, "How do peer evaluators perceive their experiences in the role will impact their own performance as social studies educators?" The best way to describe the study's findings relating to this question is to first recognize again the value each participant placed on seeing the bigger picture. For some participants, this prompted a sense of connectivity that they plan to foster in their futures as educators.

The participants reported that seeing how what happens in one classroom connects to another classroom or course has put different elements of education into focus. One participant reports it will impact lesson planning to address concepts holistically. Another participant reflects that will impact the approach to assessment, primarily with plans to increase use of all types of assessment. Finally, one participant sees participation in professional organizations as key to helping remain connected.

The theme of *non-collegial reality* is revisited and illuminated in response to the research question, "What behaviors or abilities do peer evaluators perceive to be most useful in

performing their jobs?" Taking in account the non-collegial reality and workload stressors, a shared perception by the participants is that time management and organization are key. Participants perceived that learning to deal with difficult people and quickly building rapport are among the most useful behaviors and abilities in performing their jobs, which entails overcoming negativity.

More specific to the daily performance of their duties, it is important to note that the interviews revealed that time management was both a challenge and a benefit to the participants and a key skill to performing the peer evaluator job. The study reveals that in the first year, the peer evaluators perceived themselves to be overscheduled with not enough time to complete paperwork or build relationships effectively. Facing that challenge, they reported becoming adroit at trying to make the schedules work. This ability was a benefit in subsequent years when they perceived a change to more reasonable schedules and workloads but already had strategies to be time efficient.

"Does being a peer evaluator change a person's perceived self-reported educational philosophy or perspective?" For this last research question, participant interview responses revealed a theme of *residual optimism* that helps answer this research question. In effect, these peer evaluators seem to hold fast to the idealism and even altruism that they reported was present when they first started. Despite some daunting challenges, they commonly report a perspective of wanting to continue to strive to better the profession and do more for students and teachers. Their own philosophies and perceptions are only changed in that their previously held convictions seem deepened. They are enthusiastic about strategies they witness, and want to share or utilize those strategies in their own practice. Their plans on doing that, however, vary.

The peer evaluators share the perception that there is indeed a bigger picture, which indicates other opportunities for them besides going back to their previous teaching assignments. Key phrases from the interviews that indicate the experience as a peer evaluator may have the participants entertaining thoughts of expanded opportunities include "maybe going back to the classroom," "it's one option of many," and "going to another level." The findings of the study most support that the participants' educational philosophies remain intact, but their perspectives about their own future in education have absolutely been impacted by their lived experiences as peer evaluators.

Chapter 5: Conclusions and Recommendations

Summary

Teacher evaluation figures predominantly among the major issues of current education reform. The teacher evaluation programs being put in place vary from school system to school system, with one variation featuring value-added measures of student performance along with ratings on observations of teaching to derive an overall evaluation of a teacher. In the large school system examined in this study, such teacher observations for evaluation purposes are conducted by both site administrators and a cadre of trained teachers, using a common rubric of effective teaching indicators.

Recognizing that most research to date focuses on the effects of accountability reform on students or teachers, this study took a phenomenological approach to illuminate the experience of the little-studied group of peer evaluators—those teachers who, along with administrators, are charged with conducting observations of their fellow teachers. Extensive interviews were conducted with three peer evaluators to reveal their perceived lived experiences as peer evaluators. The interviews provided qualitative data to gain perspective about the role of a peer evaluator. The research questions were:

1. Are there common perceived, self-reported elements in the backgrounds, interests or motivations of social studies teachers who elect to become peer evaluators?

2. What perceived expectations do these teachers report that they hold before becoming peer evaluators, and how are these perceived as met or not met as they progress through their roles?

3. What perceived challenges or benefits do the peer evaluators report that they encounter in performing their duties?

4. How do peer evaluators perceive their experiences in the role will impact their own performance as social studies educators?

5. What behaviors or abilities do peer evaluators perceive to be most useful in performing their jobs?

6. Does being a peer evaluator change a person's perceived self-reported educational philosophy or perspective?

The themes that emerged were *idealism* coming into the role of peer evaluator, followed by *non-collegial reality* in undertaking the role. While the peer evaluators shared altruistic notions of an idealized role helping their fellow teachers, negative receptivity from the teachers being evaluated and time constraints characterized the reality they faced in the role. The theme of *valued experience* emerged, despite the perceived challenges, as peer evaluators valued their experiences in the role, especially the opportunity to get to see a bigger picture outside their own classrooms. Finally, the peer evaluators retained some of their idealism and altruism as evidenced in the theme of *residual optimism*, indicating their expectation that the overall experience would make them better in their future educator roles.

Conclusions

This study sought to reveal the lived experience of social studies teachers in the peer evaluator role. From the reported lived experiences, the social studies peer evaluators hold a

degree of idealism about the role and the teaching profession in general. They prevail as idealists in the face of some challenging realities inherent in the role. They are or become skilled in time management and having difficult conversations. They value the experience of being in the role for the big picture perspective it affords, and for the strategies and exemplars of teaching that they witness.

Less easily discerned is whether there is some particular effect from their preparation as social studies teachers that would make them particularly suited to the role of peer evaluators. One participant did note a difference between social studies teachers and other content teachers who became peer evaluators. Abigail explained, *"There are math teachers who are evaluators and the evidence is the evidence, and here's the rating and whether that is going to impact whether the teacher is going to change something or not is irrelevant."* As a social studies teacher, she perceived that she was more willing to engage in a deeper conversation with an observed teacher or vary from a script in order to help a teacher improve. Idealism and altruism, however, are not characteristics unique to social studies teachers.

Much emphasis was placed by the participants on the perceived benefit or experience of being able to see a bigger picture through the peer evaluator role. Perhaps seeing the "big picture" is something for which social studies teachers have unique aptitudes, but ascertaining that is well beyond the scope of this study. Some of the assumptions in place at the outset of this research were, however, supported by subsequent findings.

Implications

It was assumed at the outset that the study context would be influential through its three components of change, stress and accountability. That assumption was borne out by the participants' descriptions of how the job conditions changed throughout their tenure. Peer

evaluators currently struggle to walk through shifting sands in performing their jobs. Their job

duties change, communication changes, even the tools (rubrics) they use change. Such fluidity

adds even more stressors, creating turbulence. More research into the details of change can help

with designing accountability systems that eliminate or minimize such stressors.

In particular, the results of this study illustrate just how important context is to almost any

issue, and how useful qualitative research can be in illuminating context. The goal of

perspective was reached in this study, and perspective on some aspects of the context of

accountability was achieved. Educational reform and accountability is a vast and almost

overwhelming concept. Perhaps the quality of education or the effectiveness of teachers should

not be viewed as problems to be solved, but rather as contexts with components that are actually

predicaments with their own particular contexts to be improved or withstood. The context of the

role of the peer evaluator can thus be seen as a predicament. As explained by Farson (1996),

> A problem is created by something going wrong, by a mistake, defect, disease, or a bad
>
> experience. When we find the cause, we can correct it. A predicament, however,
>
> paradoxical as it may seem, is more likely to be created by conditions that we highly
>
> value…Predicaments require interpretive thinking. Dealing with a predicament demands
>
> the ability to put a larger frame around a situation, to understand its many contexts, to
>
> appreciate its deeper and often paradoxical causes and consequences. (p. 42-43).

The distinction seems appropriate in light of the findings of this study. A school system

implementing a teacher evaluation program with a managerial approach that looks for problems

to solve or looks for a distinct start and finish to a problem underestimates the larger context of

accountability. The key players in that system, in this case the peer evaluators, are agents of

managing a predicament. To revisit the concepts of change management theory, the peer evaluator is an agent in the management of turbulence, and should be valued as such.

Limitations

The potential impact of interview-based research is in the very depth that the interviews allow. The depth in this study helps to reveal the lived experiences of the participants. In addition to any possible limitations noted in other chapters, it is advisable to note that the illumination of the lived experience, however, is always inherently limited by the participant's conceptual tools and frameworks to describe it. In particular, a context of turbulence can further confound the participant's ability to describe the constantly changing experience. There may be omissions of information or over-emphasis on unrelated information when participants give responses to interview questions. Sometimes in research involving interviews, participants can meander away from the topic of the question during their response. This should not automatically provide cause to discount the response, however, because much can be revealed in such meanderings. When perception of lived experience is the goal of the research, depth is absolutely necessary to provide adequate description.

Recommendations for Research

The development of educational reform systems should also include thorough research about the impacts of those systems. Because educational reform targets a human system, it is essential to take what can be vast and complex and make it understandable to any member of that system. To that end, the results of this study show that perception and depth of understanding can be enhanced by a qualitative approach to examining smaller components of a broad issue like teacher evaluation.

As noted by Sharratt and Fullan (2012), an enormous amount of data is being generated due to both the technological ability to create and distribute it and the imperative context or agenda of education reform that demands it. The very abundance of data generated to inform, however, can lead to confusion about the information. Within the realm of education reform, and teacher evaluation in particular, focusing on representative singular cases can help frame statistical data and give it meaning. In the example in this study, the profiles created of the three participating peer evaluators show the human facet of the context of teacher evaluation. That human facet conveys so much more than knowing there are 115 peer evaluators observing 15,000 teachers or how many observations each conducts, or even the average score of the observation ratings they assign. For example, the participant Charlie in this study explained his motivation to take on the role of peer evaluator was to help other teachers avoid what he experienced at the beginning of his teaching career which he describes as follows:

> *I didn't feel supported and really just felt like I was going through my first year kind of blind, wondering, 'Is this working? I don't think this is working. I don't think I'm doing a very good job' but not knowing what to do and not knowing how to do it, how to get better.*

From that one description, a reader can ascribe empathy to Charlie, as truly understanding what a struggling teacher can experience. In this manner, the impacts of policy are illustrated through participants' own words and descriptions of how they perceive they were impacted, allowing the reader to identify with a person in that role.

Further, the methods of phenomenological reduction employed in this study led to identification of the essence of the perceptions, thus preventing overload of anecdotal data which could have proven just as overwhelming as overabundance of statistical data. Charlie's lengthier

anecdote from which the above was excerpted is boiled down to its essence as an example of the theme of Idealism. That theme can be easily digested, retained conceptually and perhaps even be perceived again in other situations.

Thus, one recommendation for research into the issue of teacher evaluation is to implement a similar research approach with participants who are representative of subsequent cohorts of the peer evaluators to see if the emergent themes from this study continue to appear. Additionally, the same approach should be applied to representative participant groups of other people within this particular evaluation system who also conduct evaluations to compare results. These would include site administrators, content supervisors, and the group of teachers who have been designated as mentors to new teachers. To fully examine the impacts of the teacher evaluation system, future research will also need to focus on teachers being impacted by observation and evaluation and, their perceptions of that impact, to compare against other types of research data collected on that group.

Further study of the lived experiences of peer evaluators through qualitative research is also merited. More detail about their daily interactions and how they schedule their time could provide valuable information to minimize the impact of non-collegial reality that sets in when the evaluators begin their duties. Further illumination of the peer evaluators' perceptions could lead to better communication about how evaluation systems impact all members of an accountability system, thus improving the chances of success of accountability systems.

More research on the lived experiences of peer evaluators could provide evidence to address questions like, "When is a peer no longer a peer?" and "How can the service of the peer evaluator be organized to maximize the effectiveness of the role?" Specifically, research into the lived experiences of peer evaluators could serve to inform staff developers as they design

recruitment and training plans for future cadres of peer evaluators. It could especially help with the crafting of communication to stakeholders in any system undertaking a teacher evaluation program that includes peer evaluation. By providing insights into the lived experiences, the human resource definition of the role of the peer evaluator could be more thoroughly described to prospective evaluators and to those who will be evaluated, which could potentially ameliorate some of the effects of non-collegial reality found in this study.

A longitudinal approach to continue this study could also provide valuable information to first, find out where the role of peer evaluator led the participants professionally. The initial design of the system in which they operate indicated a tenure of two to three years for peer evaluators. As that initial period concludes, the possibility now exists for extensions of that tenure of up to five years. As participants noted, other career possibilities now seem either possible or necessary to them other than to return to their original teaching positions. As noted in the McClung career ladder study referenced earlier, participating in an evaluative role outside the classroom may be transformative to the subsequent career trajectory of the teachers performing in that role. A longitudinal approach may secondly provide insight into what particular aspects of their experiences as peer evaluators influence their performance in subsequent roles.

Recommendations for Practice and Policy

Throughout the duration of this study, teacher evaluation has remained a hot topic in the national news. In a widely distributed editorial spurred by his foundation's involvement in school reform, Bill Gates (2013) explains, "The challenge now is to make sure we balance the urgency for change with the need to ensure fair ways to develop, evaluate and compensate teachers for the work they do." More pertinent to the findings of this study, however, is his

cautionary statement in the same editorial: "What the country needs are thoughtfully developed teacher evaluation systems that include multiple measures of performance, such as student surveys, classroom observations by experienced colleagues and student test results."

This study underscores that need for thoughtful development of teacher evaluation systems. As repeatedly evidenced in this study, the state of fluidity in reform and accountability measures contributes to the stressors and negative receptivity. This sets up a context for turbulence within school systems. Setting up a peer evaluation system or any teacher evaluation system without proper planning for turbulence control seems folly. Making tweaks and adjustments are needed to stay responsive, but too much fluidity leads to difficulty with role definition, not just for the peer evaluator, but for those he or she is evaluating.

The professionals serving in roles like the peer evaluators in this study need to clearly understand the expectations, job functions and limitations of the role from the outset. The evaluation system administrators need to put in place a thorough recruitment campaign that explains to prospective candidates in writing and through training more precisely what the role will and will not entail, and update that communication --- with any changes --- at scheduled intervals. That precision must include details on time commitment, communication expected, and hierarchy within the overall system. These details, in turn, must be clearly explained in writing and through training to the teachers who are being evaluated. While a school district may undertake a campaign to set forth minimum qualifications for the job of peer evaluator, subsequent explanation of how the job gets done or changes in the initial conception should not solely be communicated to the peer evaluators themselves. That limits the teacher understanding of what the peer evaluator does to only a generalized notion of the role or to their individual experience with being evaluated.

Disconnects in understanding the role of the peer evaluator may be offset by utilizing some other title for the role. The use of the term "peer" connotes equality, collegiality, or even sameness or being average. The term "evaluator," however, already sets aside the notion that the teacher is equal to or the same as the evaluator. Simplicity is a better conduit for role clarity. The term "evaluator" alone is more indicative of the perceived role revealed in this study, and avoids setting up an oxymoron. Another more neutral term that fits is "observer." This better reflects the job function of observing a class and assigning a rating based on an objective rubric. In reality, those ratings alone do not evaluate a teacher. The observation ratings are used in conjunction with other data like ratings by other entities and student performance scores before a teacher receives an overall evaluation.

Removing the "peer" identifier altogether, regardless of role title, will underscore the reality that the professionals serving as peer evaluators are no longer peers. They no longer teach a full schedule of classes. They are ostensibly to be culled from the ranks of better-performing or better-rated teachers, and teachers with more experience. Some of the current peer evaluators were retired teachers or administrators before undertaking the role, which does not exactly make them peers of current teachers. In the original plan for the evaluation system in this study, the peer evaluators would only serve for two years then return to the classroom. That tenure has been extended, apparently due to recruitment and training considerations. As the tenure is prolonged, however, that takes the person in the peer evaluator role farther and farther from the realities of the teacher role, which include direct responsibility for the safety and education of a group of students, the experience of being evaluated while teaching, and the paperwork and other bureaucratic duties of the teacher role.

This role divergence stemming from time in the role leads to recommending setting and retaining initial plans for tenure in the role. A school system implementing teacher evaluation needs to decide at the outset if the cadre of observers is to be somewhat permanent or temporary by design. That decision will require serious consideration not just of numbers of current and projected qualified candidates, but also serious consideration of the goals of the overall teacher evaluation system. If it is assumed that one goal of education reform is to improve student learning by improving teacher performance, it cannot automatically be assumed that this is the direct goal of the teacher evaluation system. The goal of teacher evaluation systems is to first evaluate. The key actors in that system contribute to evaluation, but not necessarily to improvement.

In other words, once an evaluation is determined, how that information is used and acted upon relies upon the activity of other actors in that system, not upon the peer evaluator. As the participants in this study perceived, they did not become mentors or help teachers directly. They provided information for teachers to possibly help themselves and for administrators to construct an overall evaluation of a teacher. As such, it may be entirely acceptable to simply recruit a cadre of evaluators to fulfill a new role in the system that is understood to be something distinct from the level of classroom teachers or administrators and from which there is no requirement or recommendation to return to classroom teaching. Based on the amount of training and ongoing calibration required to undertake the role, it may simply be more cost-effective to create and maintain the evaluator rank and completely remove the connotation of the evaluators as peers. It may also increase the amount of potential candidates that could come from other career paths less directly related to the classroom but suited to objective observation. This could, in turn,

contribute to the teacher perception of objectivity in the ratings they receive, since the collegial attachment would be diminished.

The clear distinction of the goal of teacher evaluation systems compared to the broader goals of education reform is important. Where the particular evaluation system fits within the larger educational reform system to improve teacher performance must be clearly understood by all those acting in and affected by the systems. As already established, the goal of the teacher evaluation system is to first evaluate. But what becomes of the evaluation data generated? Much like student performance data, a teacher's evaluation ratings are compared to those of other teachers. After influencing factors and student performance are calculated through value-added procedures, a teacher is assigned an overall evaluation rating. The next step is ranking. As is inevitable with ranking, some teachers will appear to be above an acceptable level of performance, and others will appear to be below that level.

What sets up turbulence is the idea that the desired level of performance can continue to fluctuate. In other words, someone always has to be in the bottom ranking of teachers. If, for example, a system determines that teachers who are evaluated at the bottom two percent of all ranked evaluations need to improve or be removed from teaching, it must be understood that those who were previously proficient above the bottom two percent may subsequently find themselves in the bottom two percent even if nothing changed. The bar moved, which may seem on the surface to be an acceptable way to let the best rise to the top. An extended metaphor may illustrate how this approach could be problematic. Imagine the bottom two percent of athletes were eliminated through ranking at the Olympics. The bottom two percent of those athletes are still world class athletes who made it to the Olympics! Could we lose world class teachers, or is there an assumption that there will always be a really awful contingency of bad teachers to make

up the bottom two percent? As the bar is raised, could it become nearly impossible to for very good teachers to achieve stellar ratings? Could they become discouraged in their pursuit of excellence in a career which is already highly challenging? Perhaps allowing the bar to be set for a defined period of time is a better way to allow teachers to set and achieve goals for improved performance, then to automatically and continuously set the bar higher.

The questions about the ultimate goals of teacher evaluation systems and how they are implemented remain stuck in the mire of contentious issues like overall education reform, definitions of effective teaching and many other social, political and economic considerations. The peer evaluators operate in that mire. In some ways, they have become the "face" of the overall context of reform, particularly for those teachers they observe who do not so readily get the opportunity to see a bigger picture of that context. Any effort to more precisely define the role of the peer evaluator could help reduce any scapegoat effects of being cast as the "face" of reform. Providing opportunities for all the members of the system --- not just district administrators or the peer evaluators --- to see the bigger picture of how that system operates will lead to the desired development of more thoughtful and fair evaluation systems.

References

Adams, D. (1979). *The hitchhiker's guide to the galaxy.* New York: Harmony Books.

American Federation of Teachers. (1999). *Assuring teacher quality: It's union work.* (ERIC Document Retrieval Service #ED431757).

Arends, R. I. (2006). Peformance assessment in perspective: History, opportunities, and challenges. In S. Castle, & B. D. Shaklee (Eds.), *Assessing teacher performance: Performance-based assessment in teacher education* (pp. 3-22). Lanham, MD: Rowman & Littlefield Education.

Asp, E. (2000). Assessment in education: Where have we been? Where are we headed? In R. S. Brandt (Ed.), *Education in a new era* (pp. 42-52) Alexandria, VA: Association for Supervision and Curriculum Development.

Boardman, A. (2004). Teacher change and "high-stakes" assessment: What happens to professional development? *Teaching and Teacher Education, 20*(6), 545.

Borman, G. & Kimball, S. (2005). Teacher quality and educational equality: Do teachers with higher standards-based evaluation ratings close student achievement gaps? *The Elementary School Journal, 106*(1), 3-20.

Bridges, E. M. (1992). *The incompetent teacher: Managerial responses.* London: Falmer Press.

Center for Teaching and Learning (2013). Peer review of teaching. Peer review process at Brigham Young University. Retrieved from http://ctl.byu.edu/single-article/peer-review teaching

Creswell, J. W. (1994). *Research design: Qualitative and quantitative approaches*. Thousand

    Oaks, CA: Sage.

Croft, M., Glazerman, S., Haberman, D., Loeb, S., Raudenbush, S., Staiger, D. & Whitehurst,

    G.J. (2011). *Passing muster: Evaluating teacher evaluation systems. Report of the*

    Brookings Brown Center Task Group on Teacher Quality. Retrieved from

    http://www.brookings.edu/research/reports/2011/04/26-evaluating-teachers

Chism, N. V., & Stanley, C. A. (1999). *Peer review of teaching: A sourcebook.* Bolton, MA:

    Anker Publishing.

Colasacco, J. (2010). A week of observations. *Educational Leadership, 68*(4), 59-62.

College Board. (2012). About the tests. SAT and SAT subject tests overview. Retrieved

    electronically from http://sat.collegeboard.org/about-tests

Danielson, C. (2010). Evaluations that help teachers learn. *Educational Leadership, 68*(4), 35-39.

DiCarlo, M. (2012). How to use value-added measures right. *Educational Leadership, 70(3)*,

    32-37.

Duncan, A. (2010). The quiet revolution. *Vital Speeches of the Day, 76*(10), 455-459.

Ehrenberg, R. G., & Brewer, D. J. (1994). Do school and teacher characteristics matter?

    Evidence from high school and beyond. *Economics of Education Review, 13*(1), 1-17.

    doi:http://www.elsevier.com/wps/find/journaldescription.cws_home/743/description#descri

    ption

Falk, B. (2002). Standards-based reforms: Problems and possibilities. *The Phi Delta Kappan,*

    83(8), 612-620.

Farson, R. (1996). *Management of the absurd: Paradoxes in leadership*. New York:

    Touchstone.

Fernandez, C.E. & Yu, J. (2007). Peer review of teaching. *The Journal of Chiropractic Education, 21(2)*, 154–161.

Fosnot, C.T. (2005) Constructivism revisited: Implications and reflections. In C.T. Fosnot (Ed.), *Constructivism: Theory, perspectives and practice*. (2nd ed.). New York: Teachers College Press.

Gates, B. (2013, April 7). A fairer way to evaluate teachers. [Editorial]. *Tampa Tribune.* Retrieved from: http://tbo.com/list/news-opinion-commentary/a-fairer-way--to-evaluate-teachers-b82474781z1

Goldhaber, D. D., & Brewer, D. J. (1997). Why don't schools and teachers seem to matter? Assessing the impact of unobservables on educational productivity. *Journal of Human Resources, 32*(3), 505-523.

Goldstein, J. (2007). Easy to dance to: Solving the problems of teacher evaluation with peer assistance and review. *American Journal of Education, 113*(3), 479-508.

Guba, E. G. (1978). Toward a methodology of naturalistic inquiry in educational evaluation. In *CSE Monograph Series No. 9*. Los Angeles: Center for the Study of Evaluation.

Guba E.G. & Lincoln, Y.S. (1989). *Fourth generation evaluation*. Newbury Park, CA: Sage Publications.

Haladyna, T.M., Nolen, S.B. & Haas, N.S. (1991). Raising standardized achievement test scores and the origins of test score pollution. *Educational Researcher, 20*(5), 2-7.

Hanushek, E.A., & Raymond, M. E. (2001). The confusing world of educational accountability. *National Tax Journal, 54*(2), 365-384.

Hoggarth, L. & Comfort, E. (2010). *A practical guide to outcome evaluation.* London: Jessica Kingsley Publishers.

Hughes, M. (2006). *Change management: A critical perspective.* London: Chartered Institute of Personnel and Development.

Ingersoll, R., & Merrill, L. (2010). Who's teaching our children? *Educational Leadership, 67*(8), 14-20.

Johnson, S. (1998). *Who moved my cheese?* New York: Putnam.

Karnes, E.L, & Black, D. D. (1986). *Teacher evaluation and merit pay: an annotated bibliography.* New York: Greenwood Press.

Kelly, P. P. (1998, April). *A comparative analysis of teacher peer review programs in four urban districts: Professional unionism in action.* Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA. (ERIC Document Retrieval Service No. ED 447088)

Lambert, R.G., & McCarthy, C. J. (2006). *Understanding teacher stress in an age of accountability.* Greenwich, CT: Information Age Publishing.

Lawrence, C.E., Vachon, M. K., Leake, B.H, & Leake, D.O. (2001). *The marginal teacher: A step-by-step guide to fair procedures for identification and dismissal.* (2nd ed.). Thousand Oaks, CA: Corwin Press.

Larochelle, M. & Bednarz, N. (1998). Constructivism and education: beyond epistemological correctness. In M. Larochelle, N. Bednarz & J. Garrison (Eds.), *Constructivism and Education,* (pp. 1-20). Cambridge, UK: Cambridge University Press.

Lewin, T. (2010, July 24). School chief dismisses 241 teachers in Washington. *The New York Times,* p. A8.

Lewis, J. (2009). Redefining qualitative methods: Believability in the fifth moment. *International Journal of Qualitative Methods, 8*(2), 1-14.

Lieberman, M. (1998). *Teachers evaluating teachers: Peer review and the new unionism.* Bowling Green, OH: Social Philosophy & Policy Center; New Brunswick, N.J.: Transaction Publishers.

Linn, R. L. (2010). A new era of test-based educational accountability. *Measurement: Interdisciplinary research and perspectives, 8*(2-3), 145-149.

Long, W. S., Long, E. J., & Dobbins, G. H. (1998). Correlates of satisfaction with a peer evaluation system: Investigation of performance levels and individual differences. *Journal of Business & Psychology, 12*(3), 299-312.

Martinez, D. M., Desiderio, M. F., & Papakonstantinou, A. (2010). Teaching: A job or a profession? The perceptions of educators. *Educational Forum, 74*(4), 289-296. doi:10.1080/00131725.2010.507095

Marzano, R. J. (2007). *The art and science of teaching: A comprehensive framework for effective instruction.* Alexandria, VA: Association for Supervision and Curriculum Development.

McClung, S. A. (1993). Peer evaluator beliefs analyzed within a teacher belief framework. Ed.D. dissertation. The University of Arizona. Retrieved April 14, 2012, from Dissertation & Theses: Full Text (Publication No. AAT 9422814)

Merrow, J. (2013, April 11). Michelle Rhee's reign of terror [Web log post]. Retrieved from http://takingnote.learningmatters.tv/?p=6232

Miller, J. S. (2001). Self-monitoring and performance appraisal satisfaction: An exploratory field study. *Human Resource Management, 40*(4), 321-332.

Moroney, T. & Young, J. (2010, October 13). Rhee quits as Washington, D.C., schools chief

 amid clash with teachers. *Bloomberg News*. Retrieved from

 http://www.bloomberg.com/news/2010-10-13/michelle-rhee-is-said-to-step-down-as-

 washington-d-c-schools-chancellor.html

National Commission on Excellence in Education. (1983). *A nation at risk: The imperative for*

 educational reform: A report to the nation and the secretary of education. Washington, DC:

 United States Department of Education. Retrieved April 14, 2012, from

 http://www2.ed.gov/pubs/NatAtRisk/findings.html

New Teacher Project. (2010, October). *Teacher evaluation 2.0.* Retrieved April 14, 2012 from

 http://tntp.org/assets/documents/Teacher-Evaluation-Oct10F.pdf?files/Teacher-Evaluation-

 Oct10F.pdf

O'Boyle, L. (1978). A possible model for the study of nineteenth-century secondary education

 in Europe. *Journal of Social History, 12 (2)*, 236-247.

Paris, S. G., & Urdan, T. (2000). Policies and practices of high-stakes testing that influence

 teachers and schools. *Issues in Education, 6*(1), 83.

Patton, M.Q. (2002). *Qualitative research & evaluation methods.* (3rd ed.). Thousand Oaks, CA:

 Sage Publications.

Romanyshyn, R. D. (2010). The wounded researcher: Making a place for unconscious dynamics

 in the research process. *The Humanistic Psychologist, 38*(4), 275-304.

 doi:10.1080/08873267.2010.523282

Sartain, L., Stoelinga, S., Brown E., & Consortium on Chicago School Research. (2011). Rethinking teacher evaluation in Chicago: Lessons learned from classroom observation, principal-teacher conferences, and district implementation. (Research Report) Report of the Consortium on Chicago School Research.

Sawchuk, S. (2010). 'Career ladder' at risk in Arizona. *Education Week, 29*(23), 15.

Sedlak, M. W. (1989). Let us go and buy a school master: Historical perspectives on the hiring of teachers in the United States 1750-1980. In D. Warren (Ed.), *American teachers: Histories of a profession at work* (pp. 257-290). New York: Macmillan.

Seidman, I. (2006). *Interviewing as qualitative research: A guide for researchers in education and the social sciences.* (3rd ed.). [eBook version]. Retrieved from https://ehis-ebscohost-com.ezproxy.lib.usf.edu/eds/ebookviewer/ebook/nlebk_158421_AN?sid=5fe6b400-1649-432e-a9ee-d6aa0d938afc@sessionmgr111&vid=2&hid=101

Sharratt, L., & Fullan, M. (2012). *Putting FACES on the data: What great leaders do!* Thousand Oaks, CA: Corwin.

Shaw, P. (2002). Changing conversations in organization: A complexity approach to change. London: Routledge.

Shenton, A. K. (2004). Strategies for ensuring trustworthiness in qualitative research projects. *Education for Information, 22*(2), 63-75.

Stacey, R. D. (2001). *Complex responsive processes in organizations.* London: Routledge.

Stodolsky, S. S. (1990). Classroom observation. In J. Millman & L. Darling-Hammond (Eds.), *The new handbook of teacher evaluation: Assessing elementary and secondary school teachers.* (pp. 175-190). Newbury Park, CA: Sage Publications.

Stronge, J. H., & Tucker, P. D. (2000). *Teacher evaluation and student achievement.* Washington, D.C.: National Education Association.

Swenson, M. M. (1996). Essential elements in a qualitative dissertation proposal. *The Journal of Nursing Education, 35*, 188-190.

Taubman, P. M. (2009). *Teaching by numbers: Deconstructing the discourse of standards and accountability in education.* New York: Routledge.

Taneja, S., Pryor, M.G., Humphreys, J.H., & Singleton, L.P. (2013). Strategic management in an era of paradigmatic chaos: Lessons for managers. *International Journal of Management, 30 (1),* 112-126.

Todnem, R., Burnes, B., & Oswick, C. (2011). Change management: The road ahead. *The Journal of Change Management, 11(1)*, 1-6.

Wasley, P. A. (1991). *Teachers who lead: The rhetoric of reform and the realities of practice.* New York: Teachers College Press.

Weisberg, D., Sexton, S., Mulhern, J., & Keeling, D. (2009). *The widget effect: Our national failure to acknowledge and act on differences in teacher effectiveness. (2nd edition).* Retrieved April 14, 2012, from http://widgeteffect.org/

Wilkins, E. A. & Shin, E. (2011). Peer feedback: Who, what, when, why and how. *Education Digest, 77*(6), 49-53.

Wragg, E. C., Wragg, C. M., Wikeley, C. M. & Haynes, G. S. (1996). *Teacher appraisal observed.* London: Routledge.

Appendix A: Informed Consent

Informed Consent to Participate in Research
Information to Consider Before Taking Part in this Research Study

IRB Study # _____

You are being asked to take part in a research study. Research studies include only people who choose to take part. This document is called an informed consent form. Please read this information carefully and take your time making your decision. Ask the researcher or study staff to discuss this consent form with you, please ask him/her to explain any words or information you do not clearly understand. We encourage you to talk with your family and friends before you decide to take part in this research study. The nature of the study, risks, inconveniences, discomforts, and other important information about the study are listed below.

This research is considered to be minimal risk. That means that the risks associated with this study are the same as what you face every day. There are no known additional risks to those that take part in this study.

Please tell the principal investigator or study staff if you are taking part in another research study.

We are asking you to take part in a research study called:

Social Studies Teachers in an Evaluative Role: The Peer Evaluator Experience in the Accountability Era

The person who is in charge of this research study is Martha B. Ford. This person is called the Principal Investigator. However, other research staff may be involved and can act on behalf of the person in charge. She is being guided in this research by Dr. Bárbara Cruz and Dr. Howard Johnston.

The research will be conducted at the site of your choosing after your regular duty hours.

Purpose of the study

The purpose of this study is to:

- Examine the lived experiences of the social studies teachers who become peer evaluators.
- The study is being conducted as a dissertation study in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Secondary Education.

Study Procedures

If you take part in this study, you will be asked to:

- Participate in at least two in-depth interviews with the principal investigator. The interviews will be digitally recorded. After the initial interview, your answers will be transcribed and you will be asked to review them for accuracy. At the follow-up interview, you may be asked additional questions or to elaborate on the answers you have already given. This data will also be transcribed, and you will again be asked to review it for accuracy. It is after this review that a decision will be made as to whether to conduct further interviews or reviews.
- It is expected that each of the first two interviews will take between 60-90 minutes each. The second interview will occur no less than one day and no more than one week after the initial interview. Your review of the transcripts each time should take no longer than 1 hour, depending on the length of your answers. If an additional follow-up interview is required, it should take between 30-60 minutes. This means that your total time commitment to this study as described above could range from 3 ½ to 6 hours.
- The interviews will be conducted when and where it is convenient for you to do so outside of your normal duty time. Follow-up communication and transcripts of your interview will be sent to you via email at your choice of email address. Interview recordings and transcripts will be kept by participant code. All data will be compiled and kept in a locked file at the investigator's home. Files will be destroyed after the study is completed. Digital recordings will be erased after five years. Email correspondence will be saved electronically to a removable storage device, which will also be erased after five years. All paper documentation will be shredded after five years.

Total Number of Participants

About 3-5 individuals will take part in this study at USF.

Alternatives

You do not have to participate in this research study.

Benefits

We are unsure if you will receive any benefits by taking part in this research study.

Risks or Discomfort

This research is considered to be minimal risk. That means that the risks associated with this study are the same as what you face every day. There are no known additional risks to those who take part in this study. While you are being approached to volunteer for the study because of your job function, this study is not associated with your employer, nor will it be used in any way to affect your current employment or future employability.

Compensation

You will receive no payment or other compensation for taking part in this study.

Privacy and Confidentiality

We will keep your study records private and confidential. Certain people may need to see your study records. By law, anyone who looks at your records must keep them completely confidential. The only people who will be allowed to see these records are:

> The research team, including the Principal Investigator and Faculty Advisor

> Certain government and university people who need to know more about the study. For example, individuals who provide oversight on this study may need to look at your records. This is done to make sure that we are doing the study in the right way. They also need to make sure that we are protecting your rights and your safety.

> Any agency of the federal, state, or local government that regulates this research. This includes the Food and Drug Administration (FDA), Florida Department of Health, and the Department of Health and Human Services (DHHS) and the Office for Human Research Protection (OHRP)

> The USF Institutional Review Board (IRB) and its related staff who have oversight responsibilities for this study, staff in the USF Office of Research and Innovation, USF Division of Research Integrity and Compliance, and other USF offices who oversee this research.

We may publish what we learn from this study. If we do, we will not include your name. We will not publish anything that would let people know who you are.

Voluntary Participation / Withdrawal

You should only take part in this study if you want to volunteer. You should not feel that there is any pressure to take part in the study. You are free to participate in this research or withdraw at any time. There will be no penalty or loss of benefits you are entitled to receive if you stop taking part in this study. Your decision to participate or not to participate will not affect your job status.

New information about the study

During the course of this study, we may find more information that could be important to you. This includes information that, once learned, might cause you to change your mind about being in the study. We will notify you as soon as possible if such information becomes available.

You can get the answers to your questions, concerns, or complaints

If you have any questions, concerns or complaints about this study, or experience an adverse event or unanticipated problem, call Martha Ford at 813-900-9625.

If you have questions about your rights as a participant in this study, general questions, or have complaints, concerns or issues you want to discuss with someone outside the research, call the USF IRB at (813) 974-5638.

Consent to Take Part in this Research Study

It is up to you to decide whether you want to take part in this study. If you want to take part, please sign the form, if the following statements are true.

I freely give my consent to take part in this study. I understand that by signing this form I am agreeing to take part in research. I have received a copy of this form to take with me.

_____   _____
Signature of Person Taking Part in Study   Date

_____
Printed Name of Person Taking Part in Study

Statement of Person Obtaining Informed Consent

I have carefully explained to the person taking part in the study what he or she can expect from their participation. I hereby certify that when this person signs this form, to the best of my knowledge, he/ she understands:

- What the study is about;
- What procedures/interventions/investigational drugs or devices will be used;
- What the potential benefits might be; and
- What the known risks might be.

I can confirm that this research subject speaks the language that was used to explain this research and is receiving an informed consent form in the appropriate language. Additionally, this subject reads well enough to understand this document or, if not, this person is able to hear and understand when the form is read to him or her. This subject does not have a medical/psychological problem that would compromise comprehension and therefore makes it hard to understand what is being explained and can, therefore, give legally effective informed consent. This subject is not under any type of anesthesia or analgesic that may cloud their judgment or make it hard to understand what is being explained and, therefore, can be considered competent to give informed consent.

_____   _____
Signature of Person Obtaining Informed Consent / Research Authorization   Date

_____
Printed Name of Person Obtaining Informed Consent / Research Authorization

Appendix B: Interview Protocol

Initial Interview:

Part A: Reflects background, role expectations, self-image, past experience

1. What drew you to applying to become a peer evaluator?
    Possible probes or re-statement: Why did you want to become a peer evaluator? Why did the role of peer evaluator seem like an attractive job prospect to you?
2. What did you anticipate you would experience as a peer evaluator?
    Possible probes or re-statement: How did you imagine it would be like to be a peer evaluator when you started? How did you envision the role of peer evaluator?
3. Tell me about the elements in your background that you thought would help you undertake this new position.
    Possible probes or re-statement: What in your background did you think made you suited to being a peer evaluator?

Part B: Reflects recent and current experience, role match, challenges, benefits

4. You just finished describing some of your expectations as you undertook the role of peer evaluator. How have some of these expectations been met or not met in your experience as a peer evaluator?
    Possible probes or re-statement: Tell me about some of your expectations that turned out to be right on target and some that may have missed the mark. What experiences have you had that confirmed the expectations you held at the beginning of the enterprise? What expectations were not confirmed?
5. What, if any, are some of the biggest challenges in your experience as a peer evaluator? What, if any, are the benefits in your experience as a peer evaluator?
6. What behaviors or abilities are the most useful for you as a peer evaluator in performing your job?

Second Interview

Part A: Revisits recent and current experience

7. (Re-states and/or probes #4)
8. (Re-states and/or probes #5)
9. (Re-states and or probes #6)

Part B: Projects future experience, attitudinal or philosophical changes

10. How has this experience impacted, if at all, your educational philosophy or perspective?
11. How will your experience as a peer evaluator impact your performance as a social studies educator when you go back to the classroom?
12. What else would you like to add?

Appendix C: Biographical Sketch of Researcher

Early education and influences: I was raised primarily in Latin America as the daughter of a military attaché to the US Embassy in various countries. This means my K-12 education included a broad variety of types of schools with just as broad a variety of educational perspectives and philosophies. That experience led to an early adoption of an analytical (or critical) perspective of education in general, as I compared and contrasted systems each time we moved. From often being a relative outsider to the environment or context in which I attended schooled, I also came to thinking about how I learn (and how I learn best) from a fairly early age. For example, social studies takes on very different meaning when it is the social studies of the country you are in as opposed to your native country's conceptualization of social studies.   The significance of key figures or events in history logically differs depending on which country's history you may be studying. For me, this solidified the existence and importance of perspective, and how perspective and context transform reality from individual to individual. In a practical example, the near-constant translation of language, code switching for primary, secondary and other strata of cultures, and other communication challenges made me an active, regular transformer of reality. Overall, my early educational background gave rise to a metacognitive bent toward learning, and a constructivist outlook already declared in this study. Those are potential biases as I approach data collection and analysis.

Higher Education and Career Preparation: My first undergraduate major and original career preparation was in Political Science with the intended goal of working in the US State Department, Foreign Service or other governmental agency abroad. As I approached graduation, I developed an interest in Education as both a degree and career after the experience of providing tutoring services for Spanish and Test Preparation. I went on to earn a Bachelor's degree in Social Science Education, and state teacher certification in Social Science, grades 6-12. This preparation simply underscores my interest toward social science educators as participants for the study. There may be bias toward assigning importance to any themes or categories in the data analysis that shows some kind of civic involvement or patriotism.

Educational Career: I taught mostly middle school Social Studies in my years in the classroom, which included Geography and U.S. History. I also taught courses in Spanish, English to Speakers of Other Languages, and Chorus at various times as special needs arose in the schools where I taught. I began teaching at a seventh grade center, one of the creations of a system under guidelines for desegregating schools. I would characterize this period as very challenging, and recognize that those challenges also contributed to my current educational perspective and possible biases. I came to be very cognizant and analytical about political forces that affect educational practices at the classroom level,

such as legislative mandates or systemic reforms. I developed a curiosity about impacts of large-scale policies on the atomic, sometimes overlooked, members of groups affected by those policies. That curiosity remains to this day as is evidenced in the topic of this study.

As I continued teaching, I sought out opportunities to diversify my professional experience. I prepared for new roles by earning my Master's degree in Curriculum and Instruction. I took a leadership position as a department head in a middle school. I took on teaching special courses like ESOL or chorus. I became involved in school committees. I began to work with my district supervisor with curriculum creation and training. I went on to become a curriculum specialist at an International Studies magnet school. I went back to USF and earned my Educational Leadership Certification. Each new responsibility, role or experience confirmed for me that reality is created by the perspective of those in the experience. With this comes the firm belief that the more you experience, the more you truly can know. If the guiding principle for my career and life outlook were a bumper sticker, it would succinctly state "the more you know, the more you grow and the more you grow, the more you know." This stance of growth as an imperative and motivational drive will inevitably constitute a bias as I undertake this study and its analysis. Being a person inherently motivated to take on new challenges may bias my attitude toward the peer evaluators who have, after all, taken on what is a brand new role in our school district. I may exhibit bias toward comments that reinforce growth or change as positive traits during the interviews.

Twelve years ago, I became the middle school social studies supervisor for the district in which the study will take place. This is a relatively high-profile role, if not necessarily a high-power role. By that I mean that because of a supervisor's frequent interactions with teachers, the supervisor somewhat becomes the "face" of the district, or the personification of the "downtown" organization. While curriculum supervisors do make major program, material and curriculum decisions, they in fact have little "power" outside the purview of their job title (mine is middle school social studies), and very little power over personnel with only extremely limited input toward hiring teachers for schools or the dismissal of teachers. As such, as far as bias or impact of a position of power in conducting interviews, it is hard to gauge if there would be any significant impact on obtaining accurate, truthful data from participants. In short, they have little to fear from me professionally.

As a supervisor, however, I have experienced somewhat the role of the peer evaluator. Supervisors are trained in the same observation techniques, using the same rubric to code ratings. We are assigned to observe a certain number of teachers as well. So, while predictions of bias or other impacts on the study are not so clear from this connection, it does lend credibility to the researcher as having explored the context, or "having been on the island" long enough to provide accurate description.

Appendix D: IRB Approval

USF
UNIVERSITY OF
SOUTH FLORIDA

DIVISION OF RESEARCH INTEGRITY AND COMPLIANCE
Institutional Review Boards, FWA No. 00001669
12901 Bruce B. Downs Blvd., MDC035 • Tampa, FL 33612-4799
(813) 974-5638 • FAX (813) 974-5618

September 12, 2012

Martha Ford
14720 Oak Vine Drive
Lutz, FL 33559

RE: **Expedited Approval** for Initial Review
      IRB#: Pro00009081
      Title: Social Studies Teachers in an Evaluative Role: The Peer Evaluator Experience in the
              Accountability Era

Dear Ms. Ford:

On 9/11/2012 the Institutional Review Board (IRB) reviewed and **APPROVED** the above
referenced protocol. Please note that your approval for this study will expire on 9/11/13.

Approved Items:
Protocol Document(s):
Dissertation study proposal

Consent/Assent Documents:

Informed Consent.pdf
Please note, the informed consent/assent documents are valid during the period indicated by the
official, IRB-Approval stamp located on the form-which can be found under the Attachment
Tab. Valid consent must be documented on a copy of the most recently IRB-approved consent
form

(6) Collection of data from voice, video, digital, or image recordings made for research purposes.

(7) Research on individual or group characteristics or behavior (including, but not limited to, research on perception, cognition, motivation, identity, language, communication, cultural beliefs or practices, and social behavior) or research employing survey, interview, oral history, focus group, program evaluation, human factors evaluation, or quality assurance methodologies.

As the principal investigator of this study, it is your responsibility to conduct this study in accordance with IRB policies and procedures and as approved by the IRB. Any changes to the approved research must be submitted to the IRB for review and approval by an amendment.

We appreciate your dedication to the ethical conduct of human subject research at the University of South Florida and your continued commitment to human research protections. If you have any questions regarding this matter, please call 813-974-5638.

Sincerely,

John Schinka, PhD, Chairperson
USF Institutional Review Board