"HOW IS THIS MAKING MY INSTRUCTION BETTER AT ALL?":
CENTERING TEACHERS' VOICES AND STRIVING FOR HUMANIZATION IN AN
INVESTIGATION OF HIGH-STAKES EVALUATIONS

By

Amy R. Guenther

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Curriculum, Instruction, and Teacher Education - Doctor of Philosophy

2019

ABSTRACT

"HOW IS THIS MAKING MY INSTRUCTION BETTER AT ALL?":
CENTERING TEACHERS' VOICES AND STRIVING FOR HUMANIZATION IN AN
INVESTIGATION OF HIGH-STAKES EVALUATIONS

By

Amy R. Guenther

This dissertation investigates teachers' perceptions of high-stakes evaluations and examines the methods used to conduct this research. While the evaluation of teacher performance has been a long-standing practice in the United States, recent education reform policies have placed a much greater emphasis on teacher evaluation (Cohen & Goldhaber, 2016). These neoliberal policies largely focus on assessing performance to hold teachers accountable (Papay, 2012) and have resulted in many states adopting performance-based teacher evaluation systems with high-stakes attached to them (Goldstein, 2014; Lavigne, 2014). These reforms have significantly changed both how teachers are evaluated and the implications of their evaluations. The purpose of this dissertation is to better understand the professional and personal consequences of these high-stakes evaluation systems on teachers, as well as how this research might be conducted in humanizing ways. Thus, I examine the lived experiences of teachers from three elementary schools in different suburban districts in Michigan, as well as aspects of humanizing research (Paris, 2011: Paris & Winn, 2014) that I incorporated into my research methods.

This three article dissertation highlights the perspectives of teachers and reveals potential reasons for the ineffectiveness of high-stakes evaluation to improve practice, as well as several harmful consequences that high-stakes evaluations can have on teachers. At the very least, this current evaluation system does not encourage teachers to work together to improve their practice.

At its most consequential, it appears to be encouraging isolationism and creating adversarial relationships among some teachers. Thus, I argue, by implicitly and explicitly discouraging collaboration, the current evaluation system is decreasing teachers' access to the social capital that could help them be more effective in their practice. Additionally, while doing little to enhance their practice, these high-stakes evaluations are negatively influencing teachers' identities. This finding is particularly significant when one considers teachers' identities have been linked to their commitment, well-being, sense of agency, and effectiveness (Day & Kington, 2008). Therefore, I argue, it is doubtful that the current evaluation system, which focuses on accountability, is producing the desired effect of improved teaching and may actually be counterproductive, negatively influencing both teachers' practice and their identities.

I contend that teachers' voices should inform necessary changes to teacher evaluation to produce evaluation systems that actually improve their practice and enhance their identities as teachers. Furthermore, in describing and reflecting upon my efforts to make my research more humanizing for my participants, this dissertation offers methods and a rationale for utilizing aspects of humanizing research amidst neoliberal policies. Such methods can implicate and counter the deprofessionalizing and dehumanizing effects of neoliberal policies on teachers.

To the teachers who go to work every day, amidst education policies that make this work so much harder, but remain committed to doing what is best for their students.

To my village, especially Jeremy,—for it definitely took a village to make this journey a reality.

To my nieces, Avery, Ainsley, and Avery, and to my nephew, Austin, you are a constant reminder of what is good in this world and why education matters. Dream big and reach far, be kind, be silly and be true to yourself, hold close and treat well the people in your life who support you, and know that I will always be in your corner.

To my advisor, Alyssa, thank you for your guidance, support, and unwavering belief in me. Your passion for an equitable and just education system and society is truly inspiring, as are your actions as a scholar, educator, and mentor. There are no words that can adequately express my appreciation for all that you have done for me. So, I will simply say thank you from the bottom of my heart.

Lastly, to my family and friends who supported me on this journey in so many ways, thank you. It definitely took a village! I so appreciate your presence in my life. To Jeremy, especially, for always believing that I could do, and should be doing, this work. And for knowing when your best option was to vacuum. I love you.

I am fortunate to have been surrounded by many strong, evolved women who have been sources of support throughout this journey and in my life. In particular, I wish to acknowledge my mom, my grandmothers, my aunts and aunties, Alyssa, Amber, Amy, Anita, Christa, Corey, Keri, Leanne, Linda, Lindsay, Marcy, Milla, and Pat. My accomplishment is your accomplishment, too.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

OVERVIEW

I was so consumed with thinking about and working through this [evaluation] process that it took the joy out of what I love to do best, which is working with my students…. Since this rating system has been implemented, things have definitely changed for the worse.

- Kindergarten Teacher, 2016

In this dissertation, I explore teachers' perceptions of high-stakes evaluations and examine the methods I used to conduct this research in humanizing ways. As a former K-12 teacher and administrator, I am intimately familiar with high-stakes evaluations. I have been evaluated as a teacher and evaluated other teachers as an administrator under this current system. I also have listened to my spouse, teacher colleagues, and friends express disappointment, frustration, and resentment over these evaluations, similar to the Kindergarten teacher's sentiments above. Thus, I have a sense of issues related to high-stakes evaluations. Because of my experiences and the relationships I have with many teachers, I approached this research with appreciation for what it is teachers do and concern for the potential consequences of the current evaluation system. I also entered into this work with a sense of responsibility to teachers, to sharing their words, to making a difference in whatever way I could. Thus, while I have existing knowledge of the current high-stakes evaluation system and a genuine concern for the well-being of teachers, it is my participants' lived experiences that are the focus of this research.

Through this dissertation, it is not my intent to show teachers know things about high-stakes evaluations—of course they do. Nor is my purpose to suggest teachers might not know the influence these evaluations are having on them—of course they do. Teachers most certainly can speak to the influence high-stakes evaluations are having on them as people, on their practice, and on their work with colleagues. Rather, my purpose is to honor teachers' voices by providing a platform for their unique and most important insights into high-stakes evaluations, insights that should be guiding the policy and implementation of high-stakes evaluations—but are not.

While evaluation of teacher performance has been a long-standing practice in the United States, recent education reform policies have placed a much greater emphasis on teacher evaluation in an effort to improve student achievement (Cohen & Goldhaber, 2016). These policies largely focus on assessing performance to hold teachers accountable (Papay, 2012) and have resulted in many states adopting performance-based teacher evaluation systems with high-stakes attached to them (Goldstein, 2014; Lavigne, 2014). Now teachers are being held individually responsible for student achievement (largely defined by standardized test scores) more than ever before (Steinberg & Donaldson, 2016)—with rewards and sanctions, such as performance pay and dismissal, often tied to that achievement (Amrein-Beardsley, 2014; Goldstein, 2014). Thus, these reforms have significantly changed both how teachers are evaluated and the implications of their evaluations. To investigate the influences of these evaluation systems on teachers and their teaching, I examined the lived experiences of teachers from three elementary schools in Michigan.

Scholars have questioned the current implementation of high-stakes evaluations and whether or not they are addressing their intended purpose of promoting teacher growth and student achievement (e.g. Darling-Hammond, 2015; Harris & Herrington, 2015) and hypothesized high-stakes evaluations may be detrimental to teachers' well-being (Lavigne, 2014), their practice (Leana, 2011), and their work environment (Johnson, 2015). Considering the prevalence of high-stakes evaluation systems across the country, the questions surrounding them, and their potential consequences, research in this area is critical. Yet little empirical work explicitly investigates why high-stakes evaluations may not be improving the actual quality of teaching or the harmful consequences teachers may be experiencing as a result of high-stakes evaluations. Furthermore, few scholars have sought the perspectives of teachers in this current

era of high-stakes evaluation. As Paufler (2018) notes, "how… teachers experience and perceive these systems and related highstakes consequences remains largely unexamined, and consequently ignored, at multiple policy levels" (p. 2). Distinctive to my dissertation, I center the voices of teachers—the individuals who feel, and can speak to, the impacts of these reforms.

High-stakes evaluations are part of a larger neoliberal reform agenda that uses education systems for economic purposes and emphasizes improved education efficiency with reduced public expenditures by introducing competition into education (Hursh, 2004; Robertson, 2008). Scholars suggest such neoliberal policies are *deprofessionalizing* for teachers, a phenomenon that has been linked to teacher dissatisfaction, frustration, and attrition (Apple, 2001; Dunn, *in press, a, b*; Ingersoll, Merrill, & Stuckey, 2014). According to Blackburn's (2014) definition of dehumanizing, it is likely these policies also are *dehumanizing* for teachers in that they actively work to take away teachers' individuality, creativity, and humanity and treat teachers like a number or an object. Because research has a history of dehumanizing individuals and groups (Freire, 2000; hooks, 1994; Tuck & Yang, 2014), researching a dehumanizing policy in traditionally dehumanizing ways would likely cause additional harm to participants. Thus, not only is research regarding the effects of neoliberal educational policies critical, so, too, are the ways in which this research is conducted.

Therefore, in my dissertation, I seek to better understand the professional and personal consequences of a neoliberal policy—high-stakes evaluation systems—on teachers, as well as how this research might be conducted in humanizing ways (Paris, 2011; Paris & Winn, 2014) through the following broad research questions:

1. What are teachers' perceptions of high-stakes evaluations?

2. How can research in the context of a neoliberal policy be conducted in humanizing ways?

By highlighting teachers' perspectives on how high-stakes evaluations influence them professionally and personally, I attempt to both complicate the neoliberal narrative and provide important insights for the design and implementation of performance evaluations. This research is needed now more than ever due to widespread use of high-stakes systems that assume holding teachers accountable will improve their quality of teaching. Considering student learning is ultimately at stake, it is important to know more about these evaluations and how they may be affecting teachers and their work. Who better to speak to this than the teachers themselves?

Summary of Study Design

In this research, I employed a qualitative approach that incorporated aspects of humanizing research (Paris, 2011; Paris & Winn, 2014) to learn about teachers' experiences with high-stakes evaluations. I conducted this research in three phases over the course of two years with teachers from three public elementary schools in different suburban districts in Michigan. These schools are a convenience sample as I previously worked in two of the school districts and maintained contact with a former colleague who worked in the third school district. While the ethnicities and socioeconomic status of the student populations varied, the ethnicities of the teacher populations were similar across schools, similarly reflecting the demographics of the national teaching population (U.S. Department of Education, 2016). (For actual demographic information, please see Article 1). As mandated by the state, all three districts employed evaluation systems that included classroom observations, measures of student growth, and the rating of each teacher on a four-point scale from highly effective to ineffective, with staffing decisions and performance pay tied to these ratings.

Beginning with an online survey, I gained a general sense of teachers' thoughts about and experiences with high-stakes evaluations from the 55 teachers who completed the survey. In the second phase of my research, I utilized a qualitative case study approach (Yin, 2013), interviewing 14 teachers with a semi-structured protocol that was informed by participants' survey responses. Through these interviews, I was able to "elicit views and opinions from the participants" (Creswell, 2014, p. 190) regarding their evaluation system and its influence on their practice. In the third phase, I facilitated focus groups informed by the praxis of culture circles to gain further insights into my research questions and afford my participants an opportunity to speak to my initial findings. In these small groups of two to four teachers, I used raw data from the initial survey and interviews and a draft of my findings as text-based think alouds, as well as semi-structured interview questions, to facilitate the focus group discussions. These groups, made up of 10 of the 14 previous participants, afforded me the opportunity to more fully understand how teachers "make sense of their lives and their experiences" (Merriam, 2009, p.23) in relation to high-stakes evaluations. As part of this final phase, I sought my participants' reflections on the focus group discussions and their overall experiences as participants in my research through an anonymous online survey.

Synopsis

This dissertation is comprised of three stand-alone articles that highlight certain aspects of my overarching research questions. The first two articles report empirical findings based on survey, interview, and focus group data. The third article is a methodological piece that draws on my reflections as the researcher, vignettes of focus group conversations, and the survey responses of the focus group participants. While I use some of the same language from the three

articles in this introduction, the language is unique among the three articles. A synopsis of each of the three articles is as follows:

Article 1: "You Don't Tell the Opposing Teams Your Plays": Teachers' Reflections on Collaboration Amidst High-Stakes Evaluations

In this article, I examine teachers' perspectives on collaboration in the context of high-stakes evaluation through the lens of Hargreaves and Fullan's (2012) theory of professional capital. Collaboration is essential for teachers' professional growth (Drago-Severson, 2012; Hargreaves & Fullan, 2012; Hawley & Valli, 1999; Kazemi & Hubbard, 2008) and critical for building social capital, a key component of the professional capital needed for highly effective teaching (Hargreaves and Fullan, 2012). However, findings reveal, at the very least, this current evaluation system does not encourage teachers to work together to improve their practice. At its most consequential, it appears to be encouraging isolationism and creating adversarial relationships among some teachers. Thus, I argue, by implicitly and explicitly discouraging collaboration, the current evaluation system is decreasing teachers' access to the social capital that could help them be more effective in their practice.

Article 2: "How Is This Making My Instruction Better At All?": Teachers' Perceptions of High-Stakes Evaluation and Its Influence on Their Practice and Identity

In this article, I draw upon survey, interview, and focus group data from ten teachers to examine their perceptions of high-stakes evaluation and its influence on their practice. I utilized theories of teacher identity and literature about the underlying purposes of evaluation and formative feedback for my analysis. Findings show that participants' desired purposes for evaluation were in stark contrast to the limitations of the current evaluation system. Teachers indicated a need for evaluation, in particular for the purposes of professional growth,

accountability, and validation. However, overall, they indicated the current system does little to support their improvement—largely due to a lack of useful feedback. Furthermore, they viewed their current system as a formality that did not accurately reflect their teaching. They also revealed some harmful influences on their identities that highlight issues with systems that serve accountability purposes. This finding is particularly significant when one considers teachers' identities have been linked to their commitment, well-being, sense of agency, and effectiveness (Day & Kington, 2008). Thus, I argue, it is doubtful the current accountability-focused evaluation system is producing the desired effect of improved teaching and may actually be counterproductive, negatively influencing both teachers' practice and their identities.

Article 3: Moving Toward Humanizing Research in the Context of a Dehumanizing Policy: Culture Circle-Informed Focus Groups

This article is a methodological piece for researchers who seek to conduct research in more humanizing ways. In it, I share aspects of humanizing research (Paris, 2011; Paris & Winn, 2014) that I incorporated into my research on high-stakes teacher evaluations. In particular, I explore the implementation of focus groups in which I drew upon the praxis of culture circles (Freire, 2000; Souto-Manning, 2007, 2010). Through vignettes from and teachers' reflections on these focus groups, as well as my own reflections, I attempt to respond to Paris and Winn's call to "provide a roadmap to foreground the worth of such processes of humanization for inquiry and society" (2014, p. xiv). I offer suggestions for enacting this methodology with teachers amidst neoliberalism and offer evidence of what this methodology can afford teachers in that context. Ultimately, I assert humanizing research can be applied, and should be considered necessary, in social science research more broadly, particularly in the context of neoliberalism in K-12 schools.

Overall Significance

The current implementation of high-stakes performance evaluations and the related debate over whether or not these models are addressing their intended purpose of promoting teacher growth and student achievement makes this dissertation research relevant and timely. Furthermore, little research has been conducted on the consequences of high-stakes evaluations on teachers and their work (Lavigne, 2014; Steinberg & Donaldson, 2016). This may be due to the following: (1) teacher evaluations have traditionally been of little consequence (Steinberg & Donaldson, 2016; Weisberg, Sexton, Mulhern, & Keeling, 2009), (2) the research has followed the policy, focusing on assessment(s) of performance (Papay, 2012), and/or (3) teacher evaluations have only recently included a rating system with high-stakes tied to those ratings. Yet considering the prevalence of high-stakes evaluation and their potential consequences, it seems particularly important to study how these high-stakes evaluations are influencing teachers and their work.

If, as many policymakers claim, the ultimate goal of teacher evaluation truly is to increase student learning, then the emphasis of teacher evaluation should be on teacher development to improve their instruction rather than holding individual teachers accountable. Research has continuously demonstrated that teachers need their colleagues, formative feedback, and a supportive environment to enhance their practice and increase student learning. As Hill and Grossman (2013) contend,

> Policy makers must resist the urge to think that simply holding teachers accountable through evaluation systems will result in the changes in teaching that are required for students to meet more ambitious standards. Instead, policy makers must engage in the kind of high-demand, high-support policies that both help teachers learn more about the kinds of instruction envisioned by new standards and to receive the feedback and professional development required to develop new knowledge and skills. (p. 382)

My research supports these claims and extends existing scholarship on high-stakes performance evaluations by identifying new understandings regarding the harmful consequences that high-stakes evaluations can have on teachers and their work. Through this research, I provide potential implications for policy creation and implementation of these evaluations.

The passing of the federal Every Student Succeeds Act provides an opportunity for state policymakers to make changes to the current teacher evaluation system. My research, and others like it, suggests teachers' voices should inform those changes to produce evaluation systems that teachers find useful to improving their practice and enhancing their identities as teachers. Furthermore, this research offers methods and a rationale for utilizing aspects of humanizing research amidst neoliberal policies. Such methods can implicate and counter the deprofessionalizing—and dehumanizing—effects of neoliberal policies on teachers.

REFERENCES

REFERENCES

Amrein-Beardsley, A. (2014). *Rethinking value-added models in education: Critical perspectives on tests and assessment-based accountability.* New York, NY: Routledge.

Apple, M. (2006). *Educating the "right" way: Markets, standards, God, and inequality*. New York, NY: Routledge.

Blackburn, M. V. (2014). Humanizing research with LGBTQ youth through dialogic communication, consciousness raising, and action. In D. Paris & M. T. Winn (Eds.), *Humanizing research: Decolonizing qualitative inquiry with youth and communities* (pp. 43-56). Thousand Oaks, CA: SAGE Publications.

Cohen, J., & Goldhaber, D. (2016). Building a more complete understanding of teacher evaluation using classroom observations. *Educational Researcher*, *45*(6), 378-387.

Creswell, J. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches*. Thousand Oaks, CA: SAGE Publications.

Darling-Hammond, L. (2015). Can value added add value to teacher evaluation? *Educational Researcher, 44*(2), 132-137.

Day, C., & Kington, A. (2008). Identity, well-being and effectiveness: The emotional contexts of teaching. *Pedagogy, Culture & Society*, *16*(1), 7-23.

Drago-Severson, E. (2012). New opportunities for principal leadership: Shaping school climates for enhanced teacher development. *Teachers College Record*, *114*(3), 1-44.

Dunn, A.H. (in press a). "A vicious cycle of disempowerment:" The relationship between teacher morale, pedagogy, and agency in an urban high school. *Teachers College Record*.

Dunn, A. H. (in press b). Leaving a profession after it's left you: Teachers' public resignation letters as resistance amidst neoliberalism. *Teachers College Record*.

Freire, P. (2000). *Pedagogy of the oppressed* (M. B. Ramos, Trans.) (30th anniversary edition). New York, NY: Continuum.

Goldstein, D. (2014). *The teacher wars: A history of America's most embattled profession*. New York, NY: Doubleday.

Hargreaves, A., & Fullan, M. (2012). *Professional capital: Transforming teaching in every school*. New York, NY: Teachers College Press.

Harris, D. & Herrington, C. (Eds.). (2015). Value added meets the schools: The effects of using test-based teacher evaluation on the work of teachers and leaders [Special issue].

Educational Researcher, *44*(2), 71-76.

Hawley, W. D., & Valli, L. (1999). The essentials of effective professional development. In
    L. Darling-Hammond & G. Sykes (Eds.), *Teaching as the learning profession:*
    Handbook of policy and practice. (pp. 127-145). San Francisco, CA: Jossey-Bass.

Hill, H.C. & Grossman, P. (2013). Learning from teacher observations: Challenges and
    opportunities posed by new teacher evaluation systems. *Harvard Educational Review,*
    83(2), 371-384.

hooks, b. (1994). *Teaching to transgress: Education as the practice of freedom.* New York, NY:
    Routledge.

Hursh, D. (2004). Undermining democratic education in the USA: The consequences of global
    capitalism and neo-liberal policies for education policies at the local, state and federal
    levels. *Policy Futures in Education, 2*(3), 607-620.

Ingersoll, R., Merrill, L., & Stuckey, D. (2014). *Seven trends: The transformation of the teaching*
    force, updated April 2014. CPRE Report (#RR-80). Consortium for Policy
    Research in Education, University of Pennsylvania.

Johnson, S. M. (2015). Will VAMS reinforce the walls of the egg-crate school? *Educational*
    Researcher, 44(2), 117-126.

Kazemi, E., & Hubbard, A. (2008). New directions for the design and study of professional
    development. *Journal of Teacher Education, 59*(5), 428-441.

Lavigne, A. (2014). Exploring the intended and unintended consequences of high-stakes teacher
    evaluation on schools, teachers, and students. *Teachers College Record*, *116*(1), 1-29.

Leana, C. R. (2011). The missing link in school reform. *Stanford Social Innovation Review*, *9*(4),
    1-11.

Merriam, S. B. (2009). *Qualitative research: A guide to design and implementation*. San
    Francisco, CA: John Wiley & Sons.

Papay, J. (2012). Refocusing the debate: Assessing the purposes and tools of teacher evaluation.
    Harvard Educational Review, 82(1), 123-167.

Paris, D. (2011). 'A friend who understand fully': Notes on humanizing research in a
    multiethnic youth community. *International Journal of Qualitative Studies in*
    Education, *24*(2), 137-149.

Paris, D., & Winn, M. T. (2014). *Humanizing research: Decolonizing qualitative inquiry*
    with youth and communities. Thousand Oaks, CA: SAGE Publications.

Paufler, N. A. (2018). Declining morale, diminishing autonomy, and decreasing value: Principal reflections on a high-stakes teacher evaluation system. *International Journal of Education Policy and Leadership*, *13*(8), 1-15.

Robertson, S. (2007). Remaking the world: Neoliberalism and the transformation of education and teachers' labor. In M. Compton, & L. Weiner (Eds.), *The global assault on teaching, teachers and their unions* (pp. 11-36). New York, NY: Palgrave.

Souto-Manning, M. (2007). Education for democracy: The text and context of Freirean culture circles in Brazil. In B. Levinson & D. Stevick (Eds.), *Reimagining civic education: How diverse societies form democratic citizens* (pp. 121-146). Lanham, MD: Rowman & Littlefield.

Souto-Manning, M. (2010). *Freire, teaching, and learning: Culture circles across contexts*. New York, NY: Peter Lang.

Steinberg, M. P., & Donaldson, M. L. (2016). The new educational accountability: Understanding the landscape of teacher evaluation in the post-NCLB era. *Education Finance and Policy, 11*(3), 340-359.

Tuck, E., & Yang, K. W. (2014). R-words: Refusing research. In D. Paris & M. T. Winn (Eds.), *Humanizing research: Decolonizing qualitative inquiry with youth and communities* (pp. 223-248). Thousand Oaks, CA: SAGE Publications.

U.S. Department of Education, Office of Planning, Evaluation and Policy Development, Policy and Program Studies Service. (2016). *The state of racial diversity in the educator workforce.* Retrieved from https://www2.ed.gov/rschstat/eval/highered/racial-diversity/state-racial-diversity-workforce.pdf

Weisberg, D., Sexton, S., Mulhern, J., & Keeling, D. (2009). *The widget effect: Our national failure to acknowledge and act on differences in teacher effectiveness*. (J. Schunck, A. Palcisco, & K. Morgan, Contributing authors.) Retrieved from The New Teacher Project website: https://tntp.org/assets/documents/TheWidgetEffect_2nd_ed.pdf

Yin, R. K. (2013). *Case study research: Design and methods*. Thousand Oaks, CA: SAGE Publications.

ARTICLE ONE: "YOU DON'T TELL THE OPPOSING TEAMS YOUR PLAYS":
TEACHERS' REFLECTIONS ON COLLABORATION AMIDST HIGH-STAKES
EVALUATIONS

Introduction

Evaluation reforms over the past ten years have significantly changed both how teachers are evaluated and the implications of their evaluations (Steinberg & Donaldson, 2016). Resting on the rationale that holding teachers accountable and rewarding or removing them based on individual performance (largely measured by student achievement on standardized tests) will improve their effectiveness (Amrein-Beardsley, 2014; Papay, 2012), these high-stakes evaluations often rate and, subsequently, rank teachers. These rankings are then used to inform compensation and personnel decisions, such as performance pay and dismissals (Goldstein, 2014). Many scholars question these high-stakes performance evaluations and whether or not they are addressing their intended purpose of promoting teacher growth and student achievement (e.g. Darling-Hammond, 2015; Harris & Herrington, 2015). Beyond this, some scholars (e.g. Johnson, 2015; Lavigne, 2014; Leana, 2011) hypothesize high-stakes evaluations could have negative consequences for teachers and their practice by encouraging isolation (Johnson, 2015), creating competition (Goldhaber, 2015), and discouraging collaborative work (Darling-Hammond, 2015). Yet little research has investigated whether or not these hypothesized consequences have come to fruition. This study begins to fill that gap by examining teachers' perspectives on collaboration in the context of high-stakes evaluation.

Collaboration is essential for teachers' professional growth (Hawley & Valli, 1999; Kazemi & Hubbard, 2008) and overall effectiveness (Hargreaves & Fullan, 2012), but what happens to collaboration when teachers feel they are being compared to or ranked against each

other through their evaluation system? When ratings are used as rankings, such systems may promote a competitive rather than a collaborative environment where how one ranks compared to another is more important than improving one's practice or that of one's colleagues. Therefore, it is vital to know teachers' perceptions of high-stakes evaluations and how these evaluations may be influencing collaboration—their opportunities to share knowledge and resources with the ultimate goal of improving teaching and learning. Without a clear understanding of this phenomenon, a system intended to improve *teacher* quality could be having negative consequences on *teaching* quality by discouraging the very collaboration that could support teachers' professional growth and student learning.

In this study, I explore whether or not a potential consequence of high-stakes evaluations has become a reality for teachers. Specifically, I examine teachers' perspectives on collaboration in the context of high-stakes evaluation in three elementary schools in Michigan, a state that mandates high-stakes evaluations. Using a theoretical framework of professional capital, literature on the connections between collaboration and professional growth, as well as literature that warns of high-stakes evaluations' potential consequences to teachers, I investigate the following research questions:

1. What are teachers' perceptions of collaboration?
2. What are teachers' perceptions of how high-stakes evaluations influence their collaboration?

The purpose of this study is to identify new understandings of teachers' perceptions of collaboration in the context of high-stakes evaluations that rate and rank them and related implications for their work with colleagues and, thus, their professional growth. By identifying new understandings regarding the consequences that high-stakes evaluations can have on

teachers' collaboration, this study offers implications for the implementation of these evaluations. At a time when the use of high-stakes evaluations is widespread and often implemented uncritically, such research is critical.

Context for the Study

In 2009, the federal government launched Race to the Top, a competitive federal grant program that sought to reform K-12 education through monetary incentives for adopting certain educational policies, one of which was performance-based teacher evaluation systems (Goldstein, 2014; Lavigne, 2014; U.S. Department of Education, 2014). To compete for this grant money, two-thirds of states changed their laws regarding public school teacher evaluation and "eleven states moved to end 'last in, first out,' the policy that requires districts to lay off inexperienced teachers before tenured teachers, regardless of performance" (Goldstein, 2014, p. 215). Michigan, the site of this study, was one of these eleven states.

Through legislation in 2010, 2011, and 2015 (Michigan Act 451 of 1976 §380.1249, 2010; Michigan Governor's Office, 2011; Michigan Department of Education, n.d.), Michigan established a new high-stakes teacher evaluation system, mandating annual evaluations that include measures of student growth and classroom observations. Furthermore, all public school teachers must be rated as highly effective, effective, minimally effective, or ineffective. While the meanings of these rating terms are not defined in the legislation, they have significant implications for teachers because they must be used to determine performance pay and considered in staffing decisions. With ratings tied to staffing, and seniority no longer guaranteeing a position, these new high-stakes evaluations now have potentially life-changing implications that make a teacher's performance rating, and those of her colleagues, significantly more important, particularly in those districts that face layoffs due to declining enrollment. Quite

simply, a rating of highly effective versus effective could mean the difference between keeping one's job and losing it.

Theoretical Framework

The theory of professional capital (Hargreaves and Fullan, 2012) synthesizes the roles of individuals and groups to effectively enact and enhance practices within their profession. I chose this theory for my theoretical framework because it underscores the kinds of individual and collective resources teachers might leverage to improve their practice in the context of evaluation. Hargreaves and Fullan (2012) describe professional capital as "the systematic development and integration of three kinds of capital—human, social, and decisional—into the teaching profession" (p. xv). In their model, they define human capital as "the talent of individuals," social capital as "the collaborative power of the group," and decisional capital as "the wisdom and expertise to make sound judgments" (Hargreaves & Fullan, 2013, p. 37). In the context of teaching, Hargreaves and Fullan explain human capital as knowledge and skills related to content, pedagogy, learners, and communities. Social capital is the interactions and social relationships among teachers that "affect their access to knowledge and information" (Hargreaves & Fullan, 2012, p. 90). Decisional capital is the ability to make wise, in the moment, judgments within the classroom.

Professional capital is a collaborative investment where the reciprocal interaction of human, social, and decisional capital transforms teaching (Hargreaves & Fullan, 2012, 2013). Teachers increase their human capital when social capital is in place because it gives them access to their colleagues' knowledge (their human capital). Decisional capital grows as teachers build their ability to use their human capital in discretionary ways which is fostered by their social network (their social capital). Hargreaves and Fullan's model emphasizes the critical role that

17

social capital plays in the development of human and decisional capital, as well as professional capital overall. They contend that professional capital is important for professional work, professional capacity, and professional effectiveness.

The theory of professional capital highlights the importance of collaboration in professional growth and teacher effectiveness. As suggested by Hargreaves and Fullan and the literature below, collaboration builds social capital, which positively influences human and decisional capital and, thus, increases one's professional capital and overall effectiveness. Conversely, it stands to reason, if collaboration is reduced, then social capital does not increase nor lend itself to increasing human and decisional capital, which means professional capital and overall effectiveness likely do not increase. In my analysis, I explore how high-stakes evaluations may be diminishing teachers' collaboration and, thus, their social capital and overall effectiveness. The literature presented in the next section delves into the potential consequences of high-stakes performance evaluations on teachers' collaboration and the relationships between collaboration, social capital, and professional growth.

Review of Literature

While scholars (e.g. Johnson, 2015; Lavigne, 2014; Leana, 2011) suggest high-stakes evaluations may have harmful consequences for teachers, there appears to be little research that explores their hypotheses or, more specifically, such consequences on teachers' collaboration. This may be due to the following: (1) teacher evaluations have traditionally been of little consequence (Steinberg & Donaldson, 2016; Weisberg, Sexton, Mulhern, & Keeling, 2009) and/or (2) teacher evaluations have only recently included a rating system that inherently, or overtly, ranks teachers and, thus, potentially influences collaboration. Therefore, to provide context for this study, I first review literature that suggests high-stakes evaluations could be

having a negative impact on teachers and their social capital. Then, I summarize literature that denotes the importance of collaboration to teachers' professional growth. Collectively, this literature highlights the key ideas that informed this study.

The Collision of High-Stakes Evaluations and Collaboration

While there appears to be almost no research that has investigated the relationship between high-stakes performance evaluations and teacher collaboration, scholars and national associations (e.g. American Education Research Association, 2015; American Statistical Association, 2014; Darling-Hammond, 2015; Harris & Herrington, 2015; Johnson, 2015) have warned about the potential consequences of value-added measures (VAMs)—a component of teacher evaluations now required in many states, including Michigan—that has potential implications for teacher collaboration. For example, Johnson (2015) contends,

> When policymakers use VAMs to identify, reward, and dismiss teachers, they may perpetuate the egg-crate model of schooling and undermine efforts to build instructional capacity schoolwide. At any time, in any school, some teachers are more knowledgeable, experienced, and skilled than others. Schools function best when they continuously leverage teachers' expertise so that all students in all classrooms are well served. (p. 117)

In other words, such performance evaluations may encourage teachers to teach in isolation to the detriment of their practice and their students.

Additionally, Leana (2011) contends that, by emphasizing individual performance, high-stakes evaluations seemingly discount the importance of collaborative work for teachers' growth and effectiveness. She argues that value-added models, which focus only on human capital, greatly undervalue the benefits of social capital in improving teacher practice. Similarly, Galosy and Gillespie (2013) note that current reform efforts to improve teacher quality are "overly

19

focused on the individual teacher at the expense of developing professional community" (p. 215). Thus, they propose that reform efforts are having a negative impact on social capital and hampering teacher growth. Furthermore, Darling-Hammond (2015) suggests, "Because the technology of VAM ranks teachers against each other relative to the gains they appear to produce for students, one teacher's gain is another's loss, thus creating disincentives for collaborative work" (p. 135). In these systems, where teachers are evaluated relative to one another, the focus on the individual teacher may result in competition rather than collaboration, undermining "the collective enterprise of teaching (Murnane & Cohen, 1986)" (Goldhaber, 2015, p. 89).

These scholars are not alone in their concerns; many educational leaders are concerned that high-stakes evaluations have the potential to encourage competition in a profession that needs more collaboration, resulting in teachers "competing against other members of their teams, instead of working together to share best practices" (Goldstein, 2014, p. 210). Indeed, in one, if not the only, study of its kind, Marcos, Machado, and Abelha (2015) found that teacher evaluation actually had an "inhibitory effect" on collaborative practices, encouraging isolationist and competitive behaviors. Notably, their study was of a newly-implemented teacher evaluation system that was touted as reinforcing teacher collaboration. While the current high-stakes performance evaluation system is meant to improve teaching, it is potentially discouraging the very collaboration that research says is essential for professional growth.

Professional Growth through Collaboration

Scholars have long argued that collaboration is essential to teachers' professional growth (Drago-Severson, 2012; Galosy & Gillespie, 2013; Hargreaves & Fullan, 2012; Hawley & Valli, 1999; Kazemi & Hubbard, 2008). They suggest that the knowledge and skills of educators (and

the learning of students) can be improved substantially when they engage in "collegial opportunities to learn that are linked to solving authentic problems defined by the gaps between goals for student achievement and actual student performance" (Hawley & Valli, 1999, p. 127). These opportunities are essential to professional development because "learning is as much a socially shared undertaking as it is an individually constructed enterprise" (Alexander and Murphy, 1998 as cited by Hawley & Valli, 1999, p. 39). In other words, teachers need opportunities for collaborative problem solving to improve their individual practice and develop professionally.

Literature related to effective professional development also emphasizes the role of collaboration in teacher learning. The acquisition of professional knowledge involves communities of learners working together over time (Wilson & Berne, 1999). As Feiman-Nemser (2001) states, "Professional development takes place through serious, ongoing conversation…. By engaging in professional discourse… teachers can deepen knowledge of subject matter and curriculum, refine their instructional repertoire, hone their inquiry skills, and become critical colleagues" (p. 1042). This highlights not only the importance of conversation in professional development, but also the need for collaborative relationships. A key component of learning communities that foster teacher learning and instructional improvement is the "willingness of community members to assume responsibility for colleagues' growth and development" (Borko, 2004, p. 6). In other words, collegial relationships, where teachers are willing to collaborate for the benefit of each other, are critical to build their individual and collective capital.

Hargreaves and Fullan's (2013) work related to professional capital also supports the need for collaboration. They view collaboration as critical to building social capital, a key

component of the professional capital needed for highly effective teaching. Relatedly, Leana (2011) has found links between teachers' human and social capital and student achievement. Students whose teachers had both high human capital and high social capital outperformed students whose teachers had less of one or the other or both types of capital. However, more significantly, Leana's findings indicate that strong social capital can enable low-ability teachers (teachers with low human capital) to perform as well as teachers of average ability. Thus, this literature highlights the significant role collaboration plays in a teacher's effectiveness which, ultimately, impacts student learning.

While some scholars have questioned whether or not accountability-focused performance evaluations can produce the desired effect of improved teacher effectiveness (e.g. Darling-Hammond, 2015; Harris & Herrington, 2015; Hill & Grossman, 2013; Johnson, 2015; Lavigne, 2014; Leana, 2011) and, along a parallel line, other scholars have discussed the role of collaboration in professional growth (e.g. Galosy & Gillespie, 2013; Hargreaves & Fullan, 2012; Hawley & Valli, 1999; Kazemi & Hubbard, 2008), little research has been conducted to connect these two bodies of literature, and none in the context of high-stake evaluations. Bridging this gap, I look at the relationship between high-stakes evaluations and collaboration, utilizing a lens of professional capital—in which social capital that comes from collaboration plays a central role in overall effectiveness. Centering teachers' insights by utilizing data from surveys, interviews, and focus groups, I examine teachers' perceptions of collaboration in the context of high-stakes evaluations and the implications for improved teaching practices.

Methodology

In this study, I employed a qualitative approach (Maxwell, 2013; Merriam, 2009) that incorporated aspects of humanizing research (Paris, 2011; Paris & Winn, 2014) to explore

teachers' knowledge of and experiences with high-stakes evaluations, particularly in relation to collaboration. To gain a more complete understanding of the research questions, I collected data in three phases over the course of two years from teachers and principals in three public elementary schools in different suburban districts in Michigan. These schools were a convenience sample as I previously worked in two of the school districts and maintained contact with a former colleague who worked in the third school district. Collectively, the three schools employed three administrators and approximately 80 teachers. While the ethnicities of the teacher populations were similar across schools, the ethnicities and socioeconomic status of the student populations varied (see Table 1.1). As mandated by the state, all three districts employed evaluation systems that included classroom observations, measures of student growth, and the rating of each teacher on a four-point scale from highly effective to ineffective, with staffing decisions and performance pay tied to these ratings. I provide descriptions of the participants, data collection, and analysis in the sections that follow.

Table 1.1 Demographic Information of Participating Schools

	School A	**School B**	**School C**
Grade Levels	Kdg. – 5th	Kdg. – 5th	3rd – 6th
Evaluation Tool	5D+ Teacher Evaluation	Danielson Framework for Teaching	Danielson Framework for Teaching
Teachers (Total #)	23	27	31
Black			
Hispanic			
White	96%	100%	100%
Other	4% (Asian)		
Female	96%	96%	90.3% (28)
Administrator*	White Female	White Male	White Male
		White Male	White Female
Students (Total #)	461	438	587
Black	3%	12%	1.4%
Hispanic	4%	10%	3.7%
White	85%	76%	94.4%
Other	8% (Asian)	2% (American Indian)	0.1% (Asian)
Female	45%	48%	49%
ED	16.4%	85%	34.9%

* School A maintained the same principal over the course of the study. Schools B & C had a change in leadership over the course of the study.

Participants

Table 1.2 reflects the demographic information of the teachers who participated in each phase of data collection. As noted, all of the teachers who participated in the first phase of data collection (an online survey) identified themselves as White or Caucasian, and the vast majority identified themselves as female. These participants included twelve lower elementary (Kindergarten – 2nd grade) and 28 upper elementary (3rd-6th grade) teachers. This difference can be attributed to the fact that one of the participating schools is a 3rd through 6th grade building and, thus, has no lower elementary teachers. The remaining participants included Specials (Art, Music, P.E., and Spanish) teachers, Special Education teachers, and Instructional Coaches. The participants' years of experience ranged from one year to 36 years with 83.6% of the teachers

having six or more years of experience. This indicates the vast majority of participants were teaching prior to the implementation of the current evaluation ranking system and, thus, could speak to perceived changes. Lastly, 43.6% of the participants reported that they were rated "highly effective" on their most recent evaluation and 56.4% were rated "effective." Thus, no teacher who participated in this study was given a "minimally effective" or "ineffective" rating. This is not surprising as the principals indicated nearly all of their teachers do receive one of these two ratings. One of the principals did note that a teacher who had received a "minimally effective" rating was removed from her teaching position the prior year.

Table 1.2 Demographic Information of Survey Respondents, Interviewees, and Focus Group Participants

	Phase 1 Survey Respondents (n = 55)	Phase 2 Interviewees (n = 14)	Phase 3 Focus Group Participants* (n = 10)
Gender			
Female	51 (92.7%)	14 (100%)	10 (100%)
Race			
White/Caucasian	55 (100%)	14 (100%)	10 (100%)
Current Position			
Lower Elementary	12 (21.8%)	4 (28.6%)	2 (20%)
Upper Elementary	28 (50.9%)	8 (57.1%)	6 (60%)
Specials Teacher	8 (14.5%)	2 (14.3%)	2 (20%)
Special Education Teacher	3 (5.5%)		
Instructional Coach	2 (3.6%)		
"Elementary Teacher"	1 (1.8%)		
Degree Held			
Master's Degree	39 (70.9%)	8 (57.1%)	7 (70%)
School A (n = 9, 3, 2)	8 (88.9%)	3 (100%)	2 (100%)
School B (n = 21, 6, 3)	8 (38%)	1 (16.7%)	1 (33.3%)
School C (n = 25, 5, 5)	23 (92%)	4 (80%)	4 (80%)
Experience (in years)			
6 or more total	46 (83.6%)	8 (57.1%)	8 (80%)
School A (n = 9, 3, 2)	9 (100%)	3 (100%)	2 (100%)
School B (n = 21, 6, 3)	14 (66.7%)	1 (16.7%)	2 (66.7%)
School C (n = 25, 5, 5)	23 (92%)	4 (80%)	4 (80%)
6 or more at current school	35 (63.6%)	8 (57.1%)	6 (60%)
School A (n = 9, 3, 2)	8 (88.9%)	3 (100%)	2 (100%)
School B (n = 21, 6, 3)	6 (28.6%)	1 (16.7%)	0 (0%)
School C (n = 25, 5, 5)	21 (84%)	4 (80%)	4 (80%)
Evaluation Rating			
Highly Effective	24 (43.6%)	4 (28.6%)	6 (60%)
School A (n = 9, 3, 2)	2 (22.2%)	1 (33.3%)	1 (50%)
School B (n = 21, 6, 3)	4 (19%)	0 (0%)	1 (33.3%)
School C (n = 25, 5, 5)	18 (72%)	3 (60%)	4 (80%)
Effective	31 (56.4%)	10 (71.4%)	4 (40%)
School A (n = 9, 3, 2)	7 (77.8%)	2 (66.7%)	1 (50%)
School B (n = 21, 6, 3)	17 (81%)	6 (100%)	2 (66.7%)
School C (n = 25, 5, 5)	7 (28%)	2 (40%)	1 (20%)

*Focus groups occurred two years after the survey and interviews were conducted. Thus, the data in this column reflects this passage of time.

In the second phase of this study, I conducted interviews with 14 teachers. Eight of the teachers taught an upper elementary grade, four taught a lower elementary grade, and two were Specials teachers. Their years of experience range from one to 21, with eight having six or more years of experience in general and at their current school. Four of these teachers were rated "highly effective" on their evaluation and the other ten were rated "effective." Thus, while the interviewees were generally representative of the respondents who completed the survey in some demographic areas, they had fewer years of experience and lower evaluation ratings than the survey respondents overall. (See Table 1.2 for comparison.) However, this is not surprising since four of the interviewees were first year teachers who would not likely receive highly effective ratings as such.

In the third phase of this study, I invited the 14 teachers who were interviewed to participate in focus groups in the summer of 2018. Ten of the 14 teachers participated, with at least two representatives from each of the three schools. One of the participants had moved to a new district and, thus, shared insights from her experiences in two different schools. Two of the teachers were Spanish teachers, two taught lower elementary, and six taught upper elementary. At the time of the focus groups, they ranged in years of experience from three to 21. On their most recent evaluation, four of the teachers were rated "effective" and six were rated "highly effective." Thus, more of the teachers were rated "highly effective" by the time of the focus groups than when the interviews were conducted. (See Table 1.2 for additional demographic information.)

Data Collection

Beginning in the spring of 2016, I visited the participating schools, providing an overview of my project and answering questions. After gaining consent from the participants, I

began my data collection with an online survey to obtain general insights and a "numeric description of trends, attitudes, or opinions of a population by studying a sample of that population" (Creswell, 2014, p. 155). This survey (see Appendix A) consisted of demographic items, such as current teaching assignment, years of experience, and most recent evaluation rating, as well as items requiring responses on a four-point Likert scale that ranged from strongly agree to strongly disagree. These included items about the current performance evaluation system, plus more specific questions about evaluation's relationship to practice and collaboration. The survey concluded with two open-ended questions and a place for additional comments. While I did not provide a monetary incentive for completion, I did suggest to the teachers that the survey was an opportunity for their voices to be heard, a likely incentive to many of them.

After the surveys were completed, I began the second phase of my study in the summer of 2016. In this phase, I utilized a qualitative case study approach (Yin, 2013), interviewing at least three teachers and one principal per school. For these interviews, I utilized a semi-structured protocol that was informed by participants' survey responses. (See Appendix B for the instrument). Interviewing allowed me to gain a more in-depth understanding through which to explain the survey results (Creswell, 2014). While my intent was to interview a purposeful, representative sample (Creswell, 2014; Maxwell, 2013), I decided to interview all teachers who were willing as a way to honor their voices. In total, I conducted interviews with 14 teachers, three from School A, six from School B, and five from School C. Through these interviews, I was able to "elicit views and opinions from the participants" (Creswell, 2014, p. 190) regarding the current evaluation system and its influence on them, their working relationships, and their teaching practices. I conducted the interviews in-person at a location of the interviewee's

choosing at the end of the school year and throughout the summer. These interviews lasted 30-60 minutes and were audio-recorded. Interviewees were provided a $20 Amazon gift card for their participation.

In the third phase of my study, I facilitated focus groups (i.e. Byers & Wilcox, 1991; Creswell & Miller, 2000; Sagoe, 2012) in the summer of 2018. These focus groups consisted of two to four interviewees with most of the teachers participating in two sessions each. Each of the five focus group sessions lasted approximately two hours and explored the topic, among others, of collaboration in the context of teacher evaluation. For these focus groups, I drew upon the praxis of culture circles (Souto-Manning, 2010) and provided the teachers space for critical thinking, storying, and problem solving. I used a draft of my initial findings, descriptive statistics, and quotes from the first two phases as text-based think alouds as well as semi-structured questions (see Appendix C) to facilitate the focus group discussions. Within these discussions, participants spoke to my initial findings, shared additional insights about their lived—and often similar—experiences with high-stakes evaluation and problem-solved issues related to their evaluations. All of which provided further insights into my research questions. Participants in the focus groups were compensated with a $100 Amazon gift card for their time and insights.[1]

During all phases of the study, I kept a researcher's journal, recording my thoughts and ideas as I conducted my research, and wrote memos during my data analysis (Maxwell, 2013). Due to the sensitive nature of the information gathered, I treated all evaluation ratings, survey responses, interview and focus group transcripts, field notes, and memos as highly confidential. Furthermore, pseudonyms have been assigned to the schools and teachers to protect their identities.

---

[1] I received funding through a College of Education Research Enhancement Fellowship to help cover this cost.

Analysis

For this paper, my analysis focuses on the teachers' survey and interview responses and their focus group conversations, which are directly tied to my research questions. Thus, I began with an analysis of the teacher survey responses, noting patterns and trends based on descriptive statistics of frequency (Hoy & Adams, 2016; Nardi, 2006), to gain a general sense of teachers' thoughts regarding their evaluations and collaboration. Being most interested in the teachers' lived experiences, I used this quantitative data sparingly to illustrate overall trends explored in the qualitative data, giving primacy to the qualitative data. Additionally, because the response rate for School A was much lower than the response rates at School B and School C, I also utilized the descriptive statistics to complete a missing data analysis (see Appendix D).

For the second phase of analysis, I analyzed the participants' responses to the open-ended survey questions and the interview transcripts using analysis software (MaxQDA). I began by using open coding to look for emergent themes within and across the interview transcripts (Maxwell, 2013; Miles, Huberman, & Saldana, 2014). My focus was on participants' perceptions of: (1) collaboration in general, (2) evaluations in general, and (3) how their evaluation systems were influencing their collaboration (and, thus, social capital) with colleagues. After coding eight of the interviews in this manner, I created an initial list of codes that I used to analyze the remaining interview transcripts. After completing this cycle of coding, I revised and condensed similar codes into a final set of codes that I then used to code each transcript in a second round of coding. After using a similar process for the responses to the open-ended survey questions, I compared the codes and developed a coding scheme that I used with both sets of qualitative data for a third round of coding. I subsequently condensed these codes into themes. (See Appendix E for coding chart.) I also compared the data from the first two phases of data collection, looking

for trends, similarities, and contradictions. I then drafted initial findings based upon the

qualitative themes and descriptive statistics.

In the third phase of analysis, I utilized the same coding scheme with the transcripts of

the focus groups, while also looking for additional insights. I also looked for evidence that

supported and contradicted findings from the first two phases of my analysis. Additionally, I had

asked the focus group participants to review and address my initial findings; therefore, I had their

reflections on the findings as a source of triangulation. Ultimately, I compared the focus group

data to my initial findings, making adjustments as needed, and used it to create a more robust

representation of the influences of high-stakes evaluations on teachers' collaborative efforts.

Researcher's Positionality

As a former teacher and administrator, I am intimately familiar with the performance

evaluation system under investigation in this study. I have been evaluated as a teacher and

evaluated other teachers in the current system. I also have been present in conversations where

administrators discussed the rating of teachers as highly effective, effective, minimally effective,

or ineffective; in conversations where fellow teachers voiced their reactions to the ratings they

and their peers received; and in conversations where colleagues in other school districts shared

their related experiences. Thus, I also have a sense of issues related to high-stakes evaluations. I

used these notions and my experiences to build relationships with the teachers and principals and

encourage participation in this research. However, while I have existing knowledge of the

current performance evaluation system as well as conjectures about negative consequences, I

made a conscious effort to remain open to participants' interpretations, recognizing that it is their

lived experiences that were key to this study.

Secondly, having worked in two of the three schools in this study, some of the participants were former colleagues. While I intended to utilize my former relationships with them as an additional way to encourage participation and openness, I recognized that I returned to these schools with a different positionality as a researcher. This may have caused concern for some teachers, whether they knew me previously or not. Therefore, it was important for me to establish that the main intent of my study was to gather teacher voices amidst evaluation reform efforts because I thought (and the teachers seemingly agreed) that their voices were not being heard. I also emphasized the importance of confidentiality and the efforts I would (and did) take to maintain their confidentiality.

Lastly, hearing my teacher colleagues express disappointment, frustration, and resentment over the current evaluation system is very concerning, especially because these conversations almost always involve comparisons being made among teachers with negative connotations. As a teacher educator and a professional dedicated to public education, I have a keen interest in and working knowledge of professional development and the factors that constitute and influence successful teacher learning. As suggested by the literature, I know that collaboration is essential to the growth, development, and overall job satisfaction of teachers. Thus, I am compelled to investigate a system that could be eroding that collaboration. While some would call these aspects of my positionality limitations, I argue that I am uniquely positioned to investigate this system of evaluation and its potential effects on teachers' professional growth and effectiveness.

Limitations

The findings in this study are limited to the number of teachers who responded to the survey and the teachers who participated in the interviews and focus groups. Since this is a

convenience sample and the *n* is relatively small by quantitative standards, the ability to

generalize is limited. However, the findings are similar across different schools and different

evaluation tools and can be used to inform future studies. Furthermore, the study is informed by

existing theories and literature and the findings are compared to these theories and literature.

Another potential limitation is my close knowledge of two of the participating schools, as I

previously worked in both of these schools—in one as a teacher and in the other as a teacher

consultant. Therefore, I have prior connections to some of the participants and feel strongly

about their desire to teach in an environment that is conducive to their personal well-being and

growth as a teacher. However, this actually seemed to be a strength in this study as a large

majority of teachers who knew me agreed to participate.

Findings

Findings reveal teachers' perceptions of collaboration, as well as their perceptions of how

high-stakes evaluation systems influence their collaboration. Specifically, teachers value

collaboration and realize many benefits from collaboration for themselves and their students.

However, the data also illustrate that, while teachers themselves may value collaboration, they

also perceive that the current evaluation system both implicitly and explicitly discourages it by

promoting competitive and isolationist behaviors. Highlighting teachers' thoughts and

experiences, I elaborate on each of these themes and subthemes (see Figure 1.1) in the following

sections.
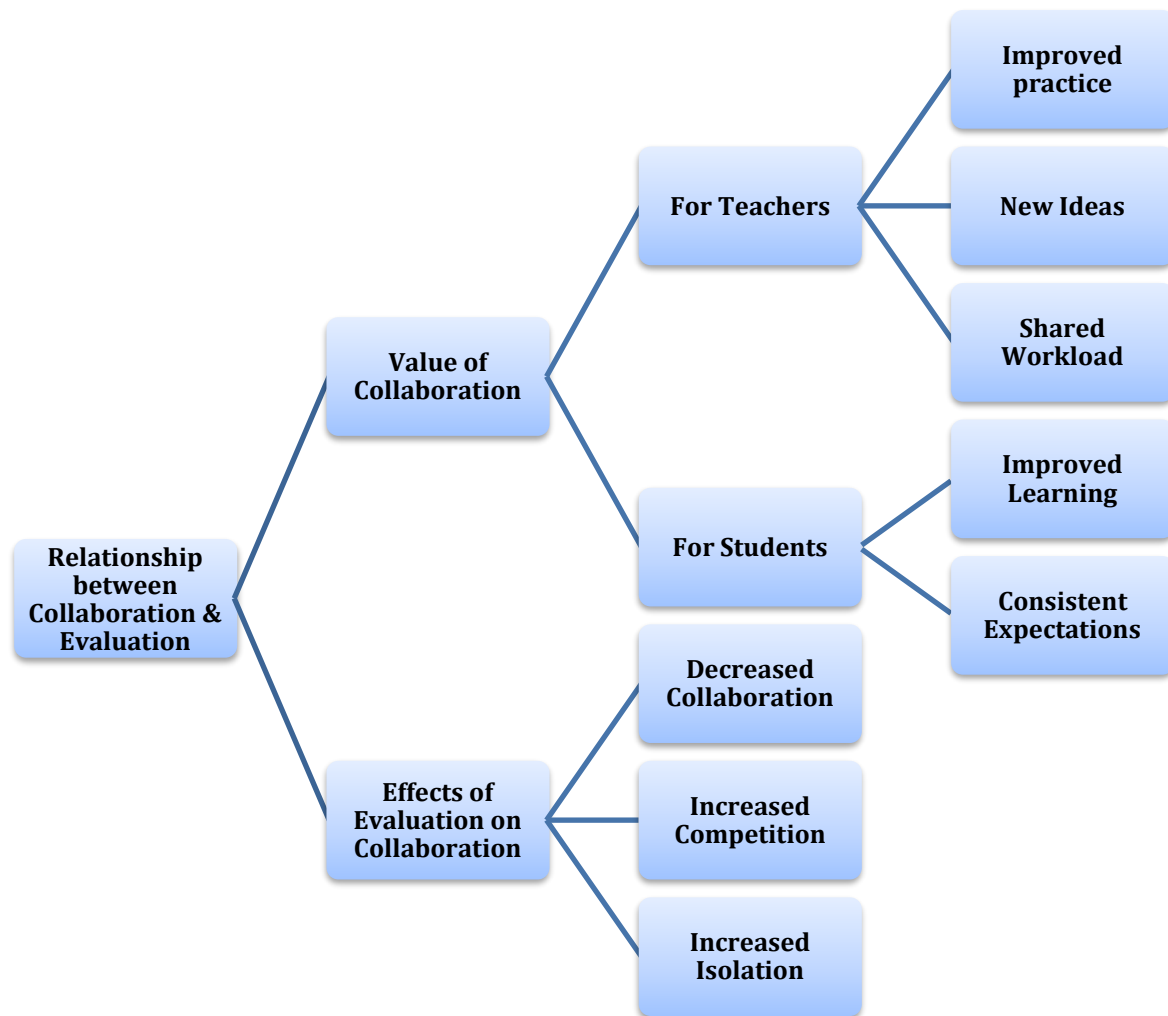
Figure 1.1 Overview of Findings.

"Two Heads are Better than One": The Value of Collaboration

All of the 55 survey respondents (see Table 1.3) indicated collaboration was beneficial to their teaching, and all except one participant (who indicated he did not have opportunities to collaborate with peers) viewed collaboration as an important part of their professional growth. Participants' interviews, focus group conversations, and responses to the open-ended survey prompt on collaboration reinforced the value these teachers place on collaboration as well as its positive influences on their work. As stated by a third grade teacher, their sentiments echoed an old adage: "Two heads are better than one" (Survey Response, May 6, 2016). In particular,

34

teachers noted benefits for themselves, including opportunities to improve their practice, access to their colleagues' ideas and expertise, and a shared workload. They also noted benefits for their students, such as a better and more consistent education.

Table 1.3 Likert-scale Responses to Selected[a] Survey Items Related to Collaboration, n = 55.

Item	Strongly Agree %	Agree %	Disagree %	Strongly Disagree %
Under the current teacher evaluation system, I am being compared to my colleagues.	10.9	47.3	40	1.8
I am concerned about how my evaluation rating compares to my colleagues' evaluation ratings.	9.1	61.8	25.5	3.6
The current evaluation system encourages me to work with my colleagues to improve my practice.	1.8	27.3	54.5	16.4
I see collaboration as an important part of my professional growth.	72.7	25.5	1.8	0
Collaborating with my colleagues is beneficial to my teaching.	81.8	18.2	0	0
My administrator values collaboration.	60	36.4	3.6	0
Since the new teacher rating system has been put in place, I am less likely to share my ideas with my colleagues.	0	5.5	69.1	25.5
Since the new teacher rating system has been put in place, I am less likely to share my resources/materials with my colleagues.	0	5.5	65.5	29.1

[a]Responses to all Likert-scale items can be found in Appendix F.

Benefiting themselves. Teachers in this study viewed collaboration as an opportunity to improve their practice. One teacher with 14 years of experience commented, "So much can be learned through discussion, questioning, analyzing, and sharing whatever we can" (Survey Response, April 26, 2016). A teacher of Spanish explained, "When we work together and pool our strengths, ideas, and resources, we become better teachers with better skills" (Survey Response, May 23, 2016). Many teachers shared similar sentiments. In fact, in their written

responses, nearly a third of the survey respondents specifically noted that one of the reasons they collaborated was to improve their teaching. Collaborating with their colleagues gave them opportunities to reflect on their own practices and encouraged them to try new instructional strategies. Collaboration was clearly a learning opportunity for these teachers.

Participants also saw each other as valuable resources. They noted that they collaborated because, as one first grade teacher stated, "I don't have all the answers," (Survey Response, May 18, 2016) and, thus, they wanted to hear their colleagues' ideas and perspectives. As one of the 38 teachers who explicitly identified this as a reason for collaboration expressed, "It is important to see other perspectives and often we end up with a better product then what we started with" (Survey Response, April 24, 2016).

Teachers also mentioned other benefits of collaboration. Pragmatically, these included a shared workload that "helps… to get things accomplished with more people to share all the tasks that are required of teachers today" (Survey Response, April 27, 2016). Collaboration was a system of support and an opportunity to access colleagues' expertise—a way to build their social capital and, thus, increase their overall effectiveness.

Benefiting students. Nearly half of the survey participants noted in their written responses that one of the reasons they collaborated was to provide a better education for their students. By working with their colleagues and utilizing their expertise, they were better able to meet the learning needs of their students. As a fifth grade teacher with 23 years of experience wrote, "It takes a team approach to reach all of our student[s'] needs" (Survey Response, May 23, 2016). They used their colleagues' ideas to improve and vary the lessons that they taught and "allow…students multiple ways of learning" (Survey Response, April 27, 2016). Collaboration also afforded teachers opportunities to support specific students and address behavioral issues.

"We're collaborating… on the styles, the systems, the strategies that we can put in place" for the

students they shared (Focus Group, June 22, 2018). These teachers saw collaboration as being in

the best interest of their students.

Another reason teachers gave for collaborating with their grade level colleagues was that

it made the learning experience more consistent for students across their grade level.

Collaboration afforded them opportunities to communicate, checking in with each other to see if

they were "on the same page" (Interview, June 7, 2016) with regard to content coverage and

expectations. This allowed them to provide similar content at a similar pace under similar

conditions. It also allowed them to brainstorm ideas around problems of practice and provide

more effective instruction to their students. As an upper elementary teacher explained,

> We work together and use the same pacing guide and curriculum so it is beneficial to
> compare strategies based on how our students are performing. If one of my colleagues is
> getting great success with their students through a strategy they are using, I would love to
> try it in my classroom with my students. Our students flourish when we all have the same
> expectations in our classrooms. (We switch for science, social studies, and writing.) We
> have different teaching styles, but we are consistent with our students. (Survey Response,
> May 1, 2016)

Through their collaboration, teachers increased consistency across their grade level classrooms,

which they saw as beneficial to their students, and also gave their students access to their

collective expertise. As a fourth grade teacher noted, "when teachers collaborate, students

perform better" (Survey Response, April 20, 2016). Indeed, by collaborating, teachers felt that

their instruction improved to the benefit of their students' learning. Collectively, these benefits of

collaboration reveal both the importance teachers' place on collaboration as well as the

importance of collaboration in increasing teachers' social capital and, thus, their human and

decisional capital and overall effectiveness.

"You Don't Tell the Opposing Teams your Plays": Collaboration Discouraged

While participants found great value in collaborating with their colleagues, their responses also suggested that their high-stakes evaluation system did not value or encourage collaboration. Furthermore, even though the vast majority of survey respondents indicated that they continued to share their resources and ideas with their colleagues (see Table 1.3), their written responses, interview comments, and focus group discussions revealed their high-stakes evaluation systems actually were influencing their collaboration in negative ways. Specifically, participants indicated these evaluation systems encouraged competition and isolation.

Discouraging collaboration. When asked if their current evaluation system encourages them to work with their colleagues to improve practice, 70.9% (39 out of 55) of the surveyed teachers indicated it does not. Thus, it would seem that these teachers did not see their evaluation system as valuing collaborative efforts. Furthermore, nine teachers (16.4%) marked "strongly disagree," which may suggest that it was actually discouraging them from working with their colleagues to improve their practice. Noting this very thing, a teacher with 15 years of experience stated, "Throughout teach[er] education and methods courses, collaboration is key. An essential piece of my master's courses was collaboration. The evaluation process, however, does the opposite and encourages competition and no collaboration" (Survey Response, April 24, 2016). Additionally, three teachers (5.5%) noted that they were less likely to share their ideas and resources since the new evaluation rating system was put in place, and two teachers (3.6%) indicated that it was not in their best interest to collaborate with their colleagues. For these teachers, the evaluation went beyond not valuing toward actually discouraging collaboration.

While these last two findings represent a small percentage of the respondents, they do show the potential damage that high-stakes evaluations can have on teachers' collaboration.

Furthermore, sentiments shared in the interviews and focus groups revealed other subtle ways high-stakes evaluations were undermining collaboration. Some of the teachers said they debated whether or not to share their ideas and materials because they knew they were being compared to their colleagues. They also noted questioning why their colleagues had not shared a particular activity or idea with them. As one sixth grade teacher explained,

> We've commented on it, and others have too, about, 'Well, how come so and so's teaching that? Or doing this fun thing and didn't tell any of us about it? Why is that?' And that conversation never came around until this grading scale, this evaluation scale, came about…. And then I think, too, you start to think about who's contributing equal parts? You know because you feel like... "Here I am giving them all this and they've given me nothing." It's that give and take and sometimes there's too much take and not enough give... It's those kinds of things that normally wouldn't have bothered you in the past, but all the sudden you're going, "...Are they going to use that STEM lesson for their evaluation? Are they putting that in their box to show how they're using technology in their classroom? ...Are they claiming that?" ...It's created this little bit of paranoia in us… Because it is, it's kind of intrinsically making you a little competitive… You're always kind of looking out for what the other people are doing. (Interview, June 3, 2016)

These sentiments reveal that some teachers were leery of sharing their ideas because of their high-stakes evaluation system. Yet getting ideas from each other was one of the benefits of collaboration they noted. If teachers are hesitant to share their ideas, it seems likely collaboration, and the benefits that come from it, will be diminished.

Some teachers also noted their hesitation to share because of instances when another colleague took credit for their ideas, as a focus group conversation revealed:

> Ainsley: So now we are very big on our social media platform… Posting what we need to on Facebook. And it's happened… where, you know, I've collaborated with a teacher and given them ideas and then they'll post it and then it gets shared with all these people and there's... no mention that I told...
> Becca: That it was your idea!
> Ainsley: Right!
> Becca: You had such a great idea!
> Ainsley: Right!
> Becca: Yeah.

Ainsley: Then them getting praise for it, like publicly, and no mention of "Well, Ainsley did it in her class last week and gave me the materials." You know what I mean?...

Sandy: Well, it's kind of competitive.

Ainsley: Right! But it's over things that shouldn't matter. You know? And I think that's a lot of what our building has kind of turned into is just petty competitiveness. (Focus group, June 19, 2018)

As explicated by these teacher and others, their high-stakes evaluation system caused them to be hesitant about sharing their ideas and to question their colleagues' motives in ways they had not previously. They began to see their colleagues as competition.

Encouraging competition. Teachers' responses suggest that high-stakes evaluations were indeed promoting varying levels of competition. Aware of the potential implications for their job status, 58.2% (32) of survey respondents felt they were being compared to their colleagues under the current evaluation system, and 70.9% (39) were concerned about their evaluation rating compared to that of their colleagues. As a fifth grade teacher explained, "When it comes to the evaluation… it's really hard to let go of the competition" (Focus Group, June 25, 2018). Relatedly, several of the interviewees explicitly noted that competition related to the evaluation system was negatively influencing working relationships. For example,

It makes teachers upset and defensive. It kind of pits teachers against each other, like tying merit pay to teacher evaluations, layoffs to teacher evaluations. It makes teachers less likely to share innovative ideas and to collaborate. And makes them a little bit more protective and like that they need to race to the top and not worry about other people around them. (Interview, June 21, 2016)

This damaging level of competition was often discussed in the context of the evaluation ratings. One teacher observed, "In the same building, it is a ranking system, and that's how it's seen by most people—even people who are rated highly effective because they want to keep that rating" (Interview, August 29, 2016). Another teacher elaborated on the influence of the rating system on collaboration.

While I still collaborate with my grade level colleagues, I do not feel all colleagues collaborate. With this evaluation process, teachers feel they are in competition to be the best. Therefore, when they have a really great idea, they are reluctant to share it because they want to ensure a higher evaluation. (Survey Response, May 15, 2016).

A sixth grade teacher echoed these sentiments that teachers may refrain from sharing their best ideas in order to protect their own ranking, demonstrating the negative influence of the evaluation ratings. She noted that some of her colleagues,

have kind of shut themselves off and don't share as much. And I think it's because they want to... rank higher than, say, me. They want to have that leg up on people…. And I think it's that little bit of competitive edge. You know, you don't tell all the opposing teams your plays. (Interview, June 3, 2016)

Her words suggest that some teachers saw their colleagues as their competitors and, as a result, had begun to isolate themselves from their colleagues.

Closing doors, promoting isolation. Consequently, another finding of this study is that the evaluation system promotes isolation. Teachers often spoke of this isolation as "the closing of doors," both literally and figuratively. Teachers not only identified this phenomenon, but also some of the consequences of it for themselves and their students. When asked how the high-stakes evaluation system had influenced her work environment, a teacher with 15 years of experience responded, "I think people aren't as collaborative. I think there's more of a closed door" (Interview, June 6, 2016). Providing an example of this and bringing to light a consequence to one's emotional well-being, a novice teacher noted the stress she experienced when attempting to plan a lesson for her formal evaluation observation:

Just planning the lesson I feel like is really stressful because you don't have anybody to help you do it. I mean you can't really go to anybody and say, 'Hey, what should I do for this?' because that person's getting evaluated like the next day or the next week or something…. You were on your own. Everybody else was on their own. It's like you're in a little bubble. (Interview, June 7, 2016)

Other teachers also noted these isolationist behaviors were most prevalent during evaluation events, such as when formal observations were taking place and toward the end of the year when teachers were completing their evaluation portfolios with evidence of their effectiveness based on the evaluation rubric. It seemed to them that some of their colleagues were more likely to work in isolation at these times of year, and they attributed this to their high-stakes evaluation system.

Finally, a second grade teacher's words highlight negative consequences for others beyond those of an individual teacher: "It [the evaluation system] definitely can create… some closing of doors which isn't what's best for our students… It can create a negative work environment with your colleagues" (Interview, June 15, 2016). Collectively, these teachers identified some of the by-products of high-stakes evaluation, namely competition and isolation, as well as the negative impacts these by-products had on both them and their students.

Discussion

Findings from this study reveal the significance of collaboration for building teachers' social capital and, thus, enhancing their overall effectiveness. Findings also provide evidence supporting the hypothesis that high-stakes evaluations that rate and rank teachers can have negative consequences for teachers and their practice. Specifically, the findings reveal that, at the very least, the current evaluation system does not encourage teachers to collaborate. Far more concerning, it appears to be encouraging teachers to teach in isolation, creating competition between teachers, and discouraging collaborative work among teachers. These conditions seemingly are diminishing teachers' access to the social capital that can help them increase their human and decisional capital. Thus, through the lens of the theory of professional capital (Hargreaves and Fullan, 2012), such evaluations are not likely to improve teachers' practice.

Teachers in this study valued collaboration, explicitly voicing the benefits of collaboration to themselves and their students. Collaboration afforded them opportunities to increase their human and decisional capital because it gave them access to their colleagues' ideas and expertise (their colleagues' human capital), which enabled them to provide a better education to their students. As Hawley and Valli (1999) and Feiman-Nemser (2001) suggest, collaboration provided these educators the opportunity to solve problems of practice, increase their knowledge and skills, improve their instruction, and better meet the needs of their students. As Hargreaves and Fullan (2012) contend in their theory of professional capital, this collaboration served as the social capital that the teachers needed to be more effective.

While the teachers clearly articulated the benefits of collaborating with their colleagues, they also spoke of ways the current high-stakes evaluation system was eroding this collaboration and, thus, their access to social capital. At the very least, their current high-stakes evaluation system did not encourage them to work together to improve their practice. In some cases, as Marcos, Machado, and Abelha (2015) also suggest, it was inhibiting collaborative efforts by promoting isolationism, the very "egg-crate model" that Johnson (2015) warned about. Furthermore, at its most consequential, these evaluation systems appeared to be creating competition, and thus adversarial relationships, among some teachers. While advocates of the rating/ranking system see high-stakes evaluations as a way to motivate teachers to improve (Goldhaber, 2015), do we really want teachers viewing each other as competition, the "opposing teams" due to an evaluation rating? Based upon the wealth of literature on the importance of collaboration to teacher learning, I argue this competition is counterproductive to improving the quality of teaching.

At a minimum, high-stakes performance evaluations that rate and rank teachers are "overselling the role" of the individual teacher and undervaluing the benefits that come from teacher collaboration (Leana, 2011, p. 30). By holding individual teachers accountable and rating and ranking them, high-stakes evaluations are fostering a competitive environment that discourages collaboration and, thus, negatively impacting a teacher's social capital. This, in turn, negatively impacts a teacher's human and decisional capital. As a result, the teacher's professional capital and overall effectiveness do not increase and may actually be diminished (see Figure 1.2). Thus, as indicated by the participants in this study, high-stakes evaluations that rate and rank teachers are not likely to improve the overall effectiveness of a teacher, at least not in the ways, or to the degree, that collegial opportunities for collaboration would. As Feiman-Nemser (2001) argues, "If we want schools to produce more powerful learning on the part of students, we have to offer more powerful learning opportunities to teachers" (p. 1013-1014). Collaboration, not accountability through evaluation, offers those powerful learning opportunities.

Such opportunities for learning are particularly important for novice teachers for whom "the first years of teaching are an intense and formative time in learning to teach, influencing not only whether people remain in teaching but what kind of teacher they become" (Feiman-Nemser, 2001, p. 1026). To grow in their practice, novice teachers need opportunities to access the human capital of their colleagues, "to talk with others about their teaching, to analyze their students' work, to examine problems, and to consider alternative explanations and actions" (Feiman-Nemser, 2001, p. 1030). Furthermore, collaboration with other teachers is one of the strongest factors associated with new teacher retention (Ingersoll & Smith, 2004). Considering the steep learning curve of new teachers and the high attrition rates among teachers within their first few

years of teaching (Ingersoll, 2003; Ingersoll, Merrill & Stuckey, 2014: Perda, 2013), opportunities for collaboration are critical. Thus, an evaluation system that discourages collaboration and makes teachers feel like they are "in a little bubble" is particularly detrimental to novice teachers.

In their theory of professional capital, Hargreaves and Fullan (2013) emphasize the importance of collaboration to teachers' overall effectiveness, stating, "The best way you can support and motivate teachers is to create the conditions where they can be effective day after day, *together* [emphasis added]" (p. 37). Hargreaves and Fullan's words are particularly significant when one considers the current high-stakes evaluation system is fostering isolationism and competition—conditions that discourage collaboration. According to their theory of professional capital, limiting collaboration diminishes one's social capital and, thus, one's effectiveness. Thus, by discouraging teachers from collaborating by rating and ranking them, highs-stakes evaluations are potentially restricting teacher growth and, thus, the benefits to students that result from this growth. Based upon the findings of this study, I argue that, by implicitly and explicitly discouraging collaboration, high-stakes evaluation systems such as the ones in my participants' school districts are decreasing teachers' access to the social capital that could help them be more effective in their practice. Thus, while intended to improve *teacher* quality, high-stakes performance evaluations actually are having an adverse effect on *teaching* quality to the detriment of students.
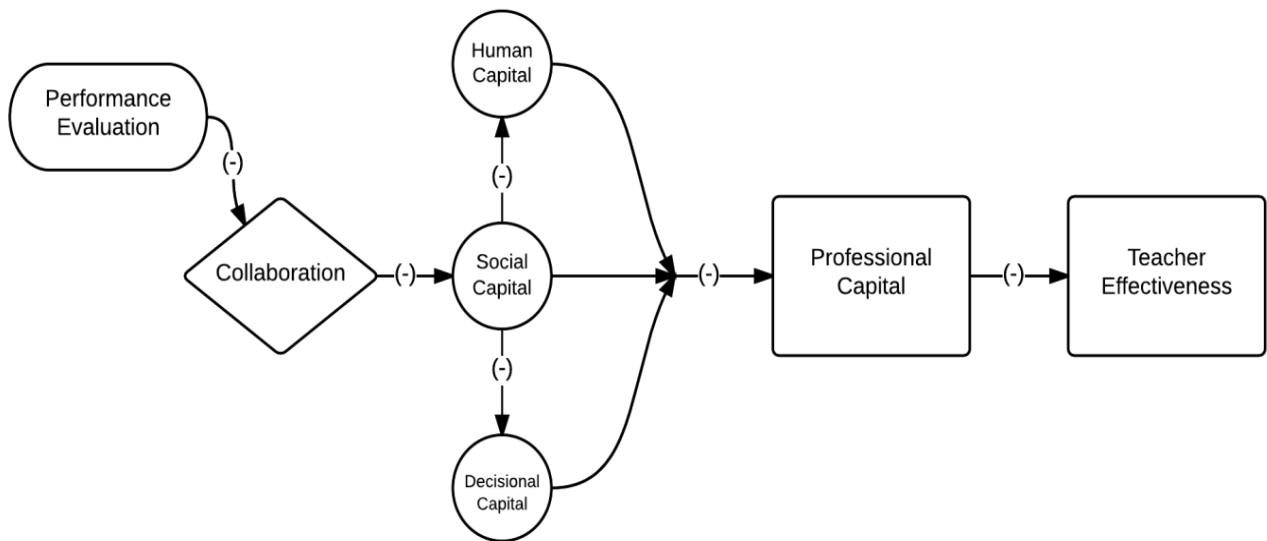
Figure 1.2 Conceptual Framework Showing the Potential Negative Influence of High-Stakes Evaluations on Teachers' Effectiveness. Based on Hargreaves and Fullan's (2012) Theory of Professional Capital.

Conclusion and Implications

While many scholars have questioned whether or not high-stakes evaluations are addressing their intended purpose of promoting teacher growth and student achievement (e.g. Darling-Hammond, 2015; Harris & Herrington, 2015) and some of these scholars have hypothesized negative consequences of high-stakes evaluations on teachers and their work (e.g. Johnson, 2015; Lavigne, 2014; Leana, 2011), little research has investigated whether or not these hypothesized consequences have come to fruition. This study begins to fill that gap by examining teachers' perspectives on collaboration in the context of high-stakes evaluation.

The current widespread implementation of high-stakes performance evaluations and their underexamined consequences, particularly on teachers' collaborative efforts and, thus, their practice, make this a relevant and timely study. Such scholarship is crucial because research has shown collaboration to be essential to teachers' professional development (Drago-Severson,

2012; Hawley & Valli, 1999; Kazemi & Hubbard, 2008). Yet, as Leana (2011) suggests, in accountability-focused evaluation systems the significance of collaboration in improving teacher performance is overshadowed by a focus on individual performance. As a result, collaboration among teachers is seemingly a casualty of high-stakes evaluation systems.

More research is needed to support the results of this study, identify other potential consequences of high-stakes evaluations, and explore factors that mitigate or compound the consequences of such a system. As the teachers in this study indicated, such factors could include the rewards tied to 'highly effective,' the stance of their grade level team members, and the influence of their principal. For example, how does tying monetary incentives to high-stakes evaluations affect collaboration? Or, as one teacher asked, "If there is a hierarchy of teaching tied to significant money, how many highly effective teachers will put their time and skills to use to help an ineffective teacher?" (Survey Response, April 24, 2016). It is also important to identify ways to counteract the negative influence of high-stakes evaluations on teachers and their working relationships. How can a collaborative environment be maintained in the context of high-stakes evaluations? What role might the school principal play in this? How might the evaluation system be used to capitalize on collaboration among teachers?

If, indeed, high-stakes evaluations are damaging collegial relationships and not promoting the collaboration that can lead to improved teaching quality, then these findings have important implications for school administrators and policymakers. We need to look closely at the ways in which evaluation systems are implemented and ratings are determined—or used at all. Most importantly, we need to identify teacher evaluation policies and practices that encourage rather than discourage collaboration among teachers. Ultimately, this is a vital policy

change because collaboration supports professional growth, which fosters improved teaching

and, thus, student learning.

APPENDICES

APPENDIX A:
Teacher Survey Questions

1. Gender: _____
2. Age: 21-25 26-30 31-35 36-40 41-45 46-50 51-55 56-60 61-65 66-70 over 70
3. Race: _____
4. Years of experience in a professional teaching position. (This is my _____ year in a professional teaching position.): _____
5. Current school: _____
6. Years in current school: _____
7. Current position: (i.e. lower elementary teacher, upper elementary teacher, specials teacher (Music, Art, P.E., Spanish, Media, etc.), Special Education teacher, etc.) _____
8. Years in current teaching position: _____
9. Highest Level of Education: Bachelor's Master's Ed. Specialist Ph.D./Ed.D.
10. Most recent evaluation rating: Highly Effective Effective Minimally Effective Ineffective

For the purposes of this survey, collaboration is defined as sharing knowledge and resources with the ultimate goal of improving teaching and student learning.

For each of the following statements, please indicate if you Strongly Agree, Agree, Disagree, or Strongly Disagree.

11. I am aware of the state requirements regarding teacher evaluation.
12. I agree with the state requirement of rating teachers through the evaluation process.
13. The rating system encourages teachers to improve their instruction.
14. I understand the current teacher evaluation system.
15. I think the current teacher evaluation system is fair.
16. I think it is important that teachers are rated on a scale from Highly Effective to Ineffective.
17. I think the current teacher evaluation system accurately assesses my teaching ability.
18. I think my current evaluation rating accurately reflects my teaching ability.
19. The current teacher evaluation system helps me grow professionally.
20. Under the current teacher evaluation system, I am being compared to my colleagues.
21. My administrator uses the current evaluation system to help me improve my practice.
22. I am concerned about how my evaluation rating compares to my colleagues' evaluation ratings.
23. The current evaluation system encourages me to work with my colleagues to improve my practice.
24. I see collaboration as an important part of my professional growth.
25. I respect my colleagues as professional educators.
26. Working with my colleagues is important regardless of what's required.
27. I enjoy collaborating with my colleagues.
28. Collaborating with my colleagues is beneficial to my teaching.
29. My administrator values collaboration.

30. Since the new teacher rating system has been put in place, I am less likely to share my ideas with my colleagues.
31. Since the new teacher rating system has been put in place, I am less likely to share my resources/materials with my colleagues.
32. It is in my best interest to collaborate with my colleagues.
33. I am proud to be a teacher.
34. I enjoy my job.

Open-ended
Please respond to the following prompts:

I collaborate with my colleagues because… I do not collaborate with my colleagues because…

I find the current evaluation system useful because… I do not find the current evaluation system useful because…

Other comments:

If you would be willing to be interviewed, please type your name here:
_____

Teacher Semi-Structured Interview Protocol

The participating teachers will be asked to respond to the following prompts individually during an interview.

1. Please tell me a little about your background, education, and teaching experiences.

2. What do you see as the role of teacher evaluation?

3. Do you think that evaluation plays a role in teacher growth? Why or why not?

4. Has the new evaluation system which requires teachers to be rated highly effective, effective, minimally effective, or ineffective affected how you perceive the evaluation process? If so, how?

5. What benefits do you see to this new system?

6. What challenges or constraints do you see with this system?

7. Do you have concerns about this new evaluation system? If so, what are they?

8. Has the new evaluation system influenced your work environment? If so, how? If not, why do you think it hasn't?

9. How do you define/describe collaboration among teachers?

10. Has the new evaluation system influenced your collaboration with your colleagues? If so, how? If not, why do you think it hasn't?

11. Do you think the new evaluation system has influenced collaboration among your colleagues? If so, how? If not, why do you think it hasn't?

12. What are some alternative evaluation methods that you think would foster collaboration among teachers?

13. Is there anything else you'd like to share with me?

14. Is there a question you thought I'd ask that I didn't?

In addition to particular prompts, the interviewer will follow-up on initial responses and ask pressing questions using questions such as:
- What do you mean by…?
- How did you do…?
- Tell me more about…
- Is there anything else you'd like to add that we have yet to discuss?

APPENDIX C:
Focus Group Questions

The following questions were provided to the focus group participants prior to their focus group sessions and used to varying degrees, along with follow-up questions, during the focus group sessions to facilitate discussion:

1. What are your thoughts as you read the findings/excerpts regarding collaboration/teacher well-being/professionalism? Do you agree? Disagree?

2. What, if anything, has changed regarding your evaluation system/tool/process/etc. over the last two years?

3. How have you made/do you make sense of the current evaluation system? Your rating?

4. What dilemmas, if any, have you faced as a result of/in relation to your evaluation/the evaluation system?

5. How have high-stakes evaluations influenced you personally? Your well-being? Your sense of professionalism?

6. How have high-stakes evaluations influenced you professionally? How have they influenced your work environment? Your work with colleagues? Your instruction?

7. What else is important for me, administrators, policymakers, etc. to know about the influences of high-stakes evaluations on collaboration/well-being/professionalism/ instruction?

8. How could the evaluation system at your school be improved? What actions do you think teachers could take to improve the current evaluation system?

9. What role does politics play in teacher evaluation?

10. In what ways, if any, does the current teacher evaluation system reflect the state of education in the United States?

For the initial survey, my goal was a response rate of 75% at each school, which I exceeded at schools B and C. At School B, 84% of the eligible teachers (21 of 25) completed the survey. (Two long-term substitute teachers were not invited to participate as they were not being assessed under the evaluation system.) At School C, 81% of the teachers (25 of 31) completed the survey. However, at School A, only 39% of the teachers (9 of 23) completed the survey. Three additional teachers at School A had agreed to participate, but technical difficulties as well as additional factors likely played a role in their and others' lack of participation. One factor may have been that I was not able to meet with all of the teachers to explain my project and answer questions because I presented at a voluntary staff meeting where not all teachers were present. A second factor that likely influenced participation was that the survey was given to the teachers late in the school year, nearly a month later than the other two schools, due to the timeline established by the principal in that building. As a result of the low participation rate at School A, the overall participation rate was 69.6% (55 of 79 teachers). Because I did not reach the stated goal of 75%, I completed a missing data analysis for School A as described below.

Because School A had a much smaller percentage of participants (39%) than Schools B (84%) and C (81%), I looked at the descriptive statistics of the demographic information (see Table 1.2) and survey responses across the three schools and noted any apparent differences between the three participant groups. Specifically, I found that Schools A and C had a considerable higher percentage of participants with 6 or more years of experience than School B (100%, 92%, and 66.7%, respectively). Relatedly, Schools A and C had a notably higher percentage of participants with Master's degree than School B (88.9%, 92%, and 38%, respectively). It is also noteworthy that a much larger percentage of participants in school C received "highly effective" ratings than in schools A or B (72%, 22.2%, and 19%, respectively.) While the respondents from School A had considerably more years of experience and Master's degrees on average than School B, they were similar to School C respondents in both regards. Also, while considerably fewer participants in School A received a "highly effective" rating than in School C, the percentage was similar to that of School B. Also notable, a larger proportion of respondents in School A (55.5%) taught lower elementary grades than School B (33.3%), which has the same grade levels. However, this difference can likely be attributed to the small n at School A.

The Likert-item responses of participants from School A were similar to the other schools with two exceptions. A considerably larger proportion of participants from School A, 4 of the 9 teachers (44.4%), indicated that their current evaluation rating does not accurately reflect their teaching ability as compared to 2 teachers (9.5%) at School B and no teachers (0%) at School C. All four of these respondents indicated that their most recent evaluation rating was "effective," which may mean they believed their rating should have been "highly effective" rather than merely "effective." However, as previously noted, the ratings at School A generally aligned with the ratings at School B. Secondly, all nine of the teachers (100%) from School A strongly agreed that their administrator values collaboration, a higher percentage than the other schools (76.2% and 32%, respectively). However, collectively 95.7% of the teachers from Schools B and C did respond with agree or strongly agree. Other than these instances, survey responses from School A aligned with responses from the other two schools and are reflected in the collective,

descriptive statistics that are shared and expanded upon with qualitative data in the findings section.

APPENDIX E:
Code Funneling[2]

Initial Codes	Condensed Codes	Themes/Subthemes

• Desire to collaborate
• Sees value in collaboration
• Self-improvement
• Improved practice
• Teacher learning
• Access to others' perspectives & ideas
• Access to other's strengths/expertise
• Support
• Shared responsibilities
• Shared workload/help
• Enjoys working with colleagues
• Required
	• Improved Practice
	• Valuable resources
	  o Ideas
	  o Expertise
	• Shared workload
	• Intrinsic Rewards
		• Value of Collaboration
		• Benefitting Teachers

• Benefits school
• Benefits grade level team
• Meeting student needs
• Variation in lessons
• Consistency in lessons
• Consistency in communication
	• Better education for students
	  o Meet student needs
	  o Improved & varied lessons
	  o Consistency in communication
	  o Consistency in content coverage
		• Value of Collaboration
		• Benefitting Students

• Evaluation doesn't encourage collaboration • Discourages collaboration	• Evaluation doesn't encourage collaboration • Discourages collaboration	• Collaboration Discouraged • Devaluing Collaboration (This theme emerged largely from the survey responses.)
• Comparison to others • Competition • Reasons for not collaborating	• Comparison • Competition	• Collaboration Discouraged • Encouraging Competition
• Isolation • Shut out • Closed door • Lack of sharing • "Nobody talks about it" • Lack of someone to collaborate with • Effects of not collaborating • Emotion	• Isolation • Decreased collaboration	• Collaboration Discouraged • Closing doors, promoting isolation

---

[2] Additional codes were identified in the initial coding cycle that related to the evaluation system more broadly. The codes displayed are those specific to the research questions.

n=55	Strongly Agree %	Agree %	Disagree %	Strongly Disagree %
I am aware of the state requirements regarding teacher evaluation.	20	69.1	9.1	1.8
I agree with the state requirement of rating teachers through the evaluation process.	0	40	58.2	3.6
The rating system encourages teachers to improve their instruction.	1.8	45.5	41.8	10.9
I understand the current teacher evaluation system.	9.1	65.5	23.6	1.8
I think the current teacher evaluation system is fair.	0	23.6	65.5	10.9
I think it is important that teachers are rated on a scale from Highly Effective to Ineffective.	1.8	50.9	43.6	3.6
I think the current teacher evaluation system accurately assesses my teaching ability.	0	30.9	61.8	7.3
I think my current evaluation rating accurately reflects my teaching ability.	18.2	70.9	10.9	0
The current teacher evaluation system helps me grow professionally.	3.6	27.3	56.4	12.7
My administrator uses the current evaluation system to help me improve my practice.	7.3	40	49.1	3.6
Under the current teacher evaluation system, I am being compared to my colleagues.	10.9	47.3	40	1.8
I am concerned about how my evaluation rating compares to my colleagues' evaluation ratings.	9.1	61.8	25.5	3.6
The current evaluation system encourages me to work with my colleagues to improve my practice.	1.8	27.3	54.5	16.4
I see collaboration as an important part of my	72.7	25.5	1.8	0

professional growth.				
I respect my colleagues as professional educators.	74.5	23.6	1.8	0
Working with my colleagues is important regardless of what's required.	81.8	18.2	0	0
I enjoy collaborating with my colleagues.	76.4	21.8	1.8	0
Collaborating with my colleagues is beneficial to my teaching.	81.8	18.2	0	0
My administrator values collaboration.	60	36.4	3.6	0
Since the new teacher rating system has been put in place, I am less likely to share my ideas with my colleagues.	0	5.5	69.1	25.5
Since the new teacher rating system has been put in place, I am less likely to share my resources/materials with my colleagues.	0	5.5	65.5	29.1
It is in my best interest to collaborate with my colleagues.	63.6	32.7	3.6	0
I am proud to be a teacher.	65.5	34.5	0	0
I enjoy my job.	38.2	61.8	0	0

REFERENCES

REFERENCES

Alexander, P. A., & Murphy, P. K. (1998). The research base for APA's learner-centered psychological principles. In N. M. Lambert & B. L. McCombs (Eds.), *Issues* in s*chool reform:* A *sampler of psychological perspectives on learner-centered schools* (pp. 25-60)*.* Washington, D.C.: The American Psychological Association.

American Educational Research Association. (2015). AERA statement on use of value-added models (VAM) for the evaluation of educators and educator preparation programs. *Educational Researcher*, *44*(8), 448-452.

American Statistical Association. (2014). ASA statement on using value-added models for educational assessment. Retrieved from http://www.scribd.com/doc/217916454/ASA-VAM-Statement-1

Amrein-Beardsley, A. (2014). *Rethinking value-added models in education: Critical perspectives on tests and assessment-based accountability.* New York, NY: Routledge.

Borko, H. (2009). Professional development and teacher learning: Mapping the terrain. *Educational Researcher, 33*(8), 3-15.

Byers, P. & Wilcox, J. (1991). Focus groups: A qualitative opportunity for researchers. *The Journal of Business Communication, 28*(1), 63-78.

Creswell, J. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches*. Thousand Oaks, CA: SAGE Publications.

Creswell, J. & Miller, D. (2000). Determining validity in qualitative inquiry. *Theory Into Practice, 39*(3), 124-130.

Danielson, Charlotte. (2014). The Framework for Teaching Evaluation Instrument. Retrieved from http://danielsongroup.org/framework/

Darling-Hammond, L. (2015). Can value added add value to teacher evaluation? *Educational Researcher, 44*(2), 132-137.

Drago-Severson, E. (2012). New opportunities for principal leadership: Shaping school climates for enhanced teacher development. *Teachers College Record*, *114*(3), 1-44.

Feiman-Nemser, S. (2001). From preparation to practice: Designing a continuum to strengthen and sustain teaching. *Teachers College Record, 103*(6), 1013-1055.

Galosy, J. A., & Gillespie, N. M. (2013). Community, inquiry, leadership: Exploring early career opportunities that support stem teacher growth and sustainability. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas, 86*(6), 207-215.

Goldhaber, D. (2015). Exploring the potential of value-added performance measures to affect the quality of the teacher workforce. *Educational Researcher, 44*(2), 87-95.

Goldstein, D. (2014). *The teacher wars: A history of America's most embattled profession.* New York, NY: Doubleday.

Hargreaves, A., & Fullan, M. (2012). *Professional capital: Transforming teaching in every school.* New York, NY: Teachers College Press.

Hargreaves, A., & Fullan, M. (2013). The power of professional capital. *Learning Forward*, *34*(3), 36-39.

Harris, D. & Herrington, C. (Eds.). (2015). Value added meets the schools: The effects of using test-based teacher evaluation on the work of teachers and leaders [Special issue]. *Educational Researcher*, *44*(2), 71-76.

Hawley, W. D., & Valli, L. (1999). The essentials of effective professional development. In L. Darling-Hammond & G. Sykes (Eds.), *Teaching as the learning profession: Handbook of policy and practice.* (pp. 127-145). San Francisco, CA: Jossey-Bass.

Hill, H.C. & Grossman, P. (2013). Learning from teacher observations: Challenges and opportunities posed by new teacher evaluation systems. *Harvard Educational Review, 83*(2), 371-384.

Hoy, W.K. & Adams, C.M. (2016). *Quantitative research in education: A primer*. (2nd Edition) Thousand Oaks, CA: SAGE Publications.

Ingersoll, R. (2003). *Is there really a teacher shortage?* (CPRE Research Report # R-03-4). Consortium for Policy Research in Education, University of Pennsylvania.

Ingersoll, R., Merrill, L., & Stuckey, D. (2014). *Seven trends: The transformation of the teaching force, updated April 2014*. (CPRE Report #RR-80). Consortium for Policy Research in Education, University of Pennsylvania.

Ingersoll, R. & Smith, T. (2004). Do teacher induction and mentoring matter? *NASSP Bulletin, 88*(638), 28-40.

Johnson, S. M. (2015). Will VAMS reinforce the walls of the egg-crate school? *Educational Researcher, 44*(2), 117-126.

Kazemi, E., & Hubbard, A. (2008). New directions for the design and study of professional development. *Journal of Teacher Education, 59*(5), 428-441.

Lavigne, A. (2014). Exploring the intended and unintended consequences of high-stakes teacher evaluation on schools, teachers, and students. *Teachers College Record*, *116*(1), 1-29.

Leana, C. R. (2011). The missing link in school reform. *Stanford Social Innovation Review*, *9*(4), 1-11.

Marcos, A., Machado, E., & Abelha, M. (2015). Effect(s) of teacher evaluation on collaborative practices: Induction or inhibition?. *Procedia - Social and Behavioral Sciences*, *174*, 3674-3680.

Maxwell, J. A. (2013). *Qualitative research design: An interactive approach* (3rd ed.). Thousand Oaks, CA: SAGE Publications.

Merriam, S. B. (2009). *Qualitative research: A guide to design and implementation*. San Francisco, CA: John Wiley & Sons.

Michigan Act 451 of 1976, Michigan Revised School Code §380.1249 (2010). Retrieved from http://www.legislature.mi.gov/(S(ksfrxop0jjx4urdvmqrydwtc))/mileg.aspx?page=GetOb ect&objectname=mcl-380-1249

Michigan Department of Education (*n.d.*). *Michigan educator evaluations at-a-glance*. Retrieved from https://www.michigan.gov/documents/mde/Educator_Evaluations_At-A-Glance_522133_7.pdf

Michigan Governor's Office. (2011, July 19). Teacher tenure reform signed into law [Press release]. Retrieved from http://www.michigan.gov/snyder/0,1607,7-277-57577-259445--,00.html

Miles, M. B., Huberman, A. M., & Saldana, J. (2014). *Qualitative data analysis: A method sourcebook.* Thousand Oaks: SAGE Publications.

Murnane, R., & Cohen, D. (1986). Merit pay and the evaluation problem: Why most merit pay plans fail and few survive. *Harvard Educational Review, 56*(1), 1–17.

Nardi, P.M. (2006). *Interpreting data: A guide to understanding research.* Boston, MA: Pearson.

Papay, J. (2012). Refocusing the debate: Assessing the purposes and tools of teacher evaluation. *Harvard Educational Review, 82*(1), 123-167.

Paris, D. (2011). 'A friend who understand fully': Notes on humanizing research in a multiethnic youth community. *International Journal of Qualitative Studies in Education*, *24*(2), 137-149.

Paris, D., & Winn, M. T. (2014). *Humanizing research: Decolonizing qualitative inquiry with youth and communities*. Thousand Oaks, CA: SAGE Publications.

Perda, D. (2013). *Transitions into and out of teaching: A longitudinal analysis of early career teacher turnover* (Unpublished doctoral dissertation). University of Pennsylvania, Philadelphia.

Souto-Manning, M. (2010). *Freire, teaching, and learning: Culture circles across contexts*. New York, NY: Peter Lang.

Steinberg, M. P., & Donaldson, M. L. (2016). The new educational accountability: Understanding the landscape of teacher evaluation in the post-NCLB era. *Education Finance and Policy, 11*(3), 340-359.

U.S. Department of Education. (2014). *Setting the Pace: Expanding opportunity for America's students under Race to the Top*. Retrieved from http://www.whitehouse.gov/sites/ default/files/docs/settingthepacerttreport_3-2414_b.pdf

University of Washington, Center for Educational Leadership. (2012). 5 Dimensions of Teaching and Learning. http://info.k-12leadership.org/5-dimensions-of-teaching-and-learning?_ga=1.57564131.87949747.1473787149 .

Weisberg, D., Sexton, S., Mulhern, J., & Keeling, D. (2009). *The widget effect: Our national failure to acknowledge and act on differences in teacher effectiveness*. (J. Schunck, A. Palcisco, & K. Morgan, Contributing authors.) Retrieved from The New Teacher Project website: https://tntp.org/assets/documents/TheWidgetEffect_2nd_ed.pdf

Wilson, S.M., & Berne, J. (1999). Teacher learning and the acquisition of professional knowledge: An examination of research on contemporary professional development. *Review of Research in Education, 24,* 173-209.

ARTICLE TWO: "HOW IS THIS MAKING MY INSTRUCTION BETTER AT ALL?":

TEACHERS' PERCEPTIONS OF HIGH-STAKES EVALUATION AND ITS

INFLUENCE ON THEIR PRACTICE AND IDENTITY

Introduction

The evaluation of teacher performance has been a long-standing practice in U.S. public schools, but recent education reform policy has placed a much greater emphasis on teacher evaluation (Cohen & Goldhaber, 2016). Motivated by Race to the Top incentives and NCLB waivers, the vast majority of states have implemented performance-based teacher evaluation systems with many attaching high-stakes (rewards and sanctions) to these evaluations in an effort to improve teacher quality (Amrein-Beardsley, 2014; Goldstein, 2014; Lavigne, 2014). Ideally, evaluation is meant to promote individual teacher growth and, ultimately, student learning (Feeney, 2007; Michigan Council for Educator Effectiveness, 2013). While evaluations can provide teachers with feedback to enhance their practice, recent evaluation policies largely have focused on another purpose: assessing performance to hold teachers accountable (Papay, 2012). As a result, teachers are being held individually responsible for student achievement now more than ever before (Steinberg & Donaldson, 2016)—with rewards and punishments, such as performance pay and dismissal, often tied to that achievement.

Some scholars have begun to look at the effectiveness of these high-stakes evaluations in improving teacher quality as measured by student achievement on standardized tests (e.g. Dee & Wyckoff, 2015; Prado Tuma, Hamilton, & Tsai, 2018), finding contradictory results. However, few have considered why these high-stakes evaluations may not be improving the quality of teaching or the consequences teachers may be experiencing as a result of these high-stakes evaluations, whether test scores have improved or not. While several scholars have questioned

the current implementation of high-stakes evaluations and whether or not they are addressing

their intended purpose of promoting teacher growth and student achievement (e.g. Darling-

Hammond, 2015; Harris & Herrington, 2015), few have sought the perspectives of teachers in

this current era of high-stakes evaluation. As Paufler (2018) notes, "how… teachers experience

and perceive these systems and related highstakes consequences remains largely unexamined,

and consequently ignored, at multiple policy levels" (p. 2). With teacher evaluation policies and

practices frequently shifting, it is critically important for policymakers and researchers to include

teachers' voices and perspectives when considering changes to teacher evaluation.

Thus, in this study, I examine teachers' lived experiences with high-stakes evaluations.

Specifically, I highlight teachers' perceptions of their current evaluation system and its influence

on their teaching by exploring the following research question:

- What are teachers' perceptions of how high-stakes evaluations influence their

  practice?

Distinctively, my study occurs in Michigan where the implementation of high-stakes evaluations

for accountability purposes has been in use for an extended time. Building on and adding to the

limited empirical work, I utilize a lens of identity to reveal teachers' perceptions of the

influences of high-stakes evaluation on their teaching. Through an analysis of survey, interview,

and focus group data, I identify themes that have implications for the redesign of these

evaluations.

Theoretical Framework

I draw on theories of identity to examine teachers' perceptions of high-stakes evaluations

because of identity's connections to agency, well-being, commitment, and job satisfaction (e.g.,

see Day & Kington, 2008). In a basic sense, identity is "being recognized as a certain 'kind of

person' in a given context. [Thus] …people have multiple identities connected… to their performances in society" (Gee, 2001, p. 99). In other words, one's identity is not singular and set, but rather multifaceted and responsive to one's setting. In addition to contextual factors, Beauchamp & Thomas (2009) maintain that identity is shaped by emotion, discourse and stories, reflection, an understanding of the self, and agency, which are all interlinked. Highlighting the multidimensional nature of identity, Day and Kington (2008) suggest identity is "a composite consisting of interactions between personal, professional and situational factors" (p. 11) that can include competing or conflicting elements. While there are many interacting factors that influence identity, these factors are not always congruent.

Looking more specifically at teachers and their professional identity, Beijaard, Meijer, and Verloop (2004) identify four features of teachers' professional identity from their review of research: (1) Professional identity is an ongoing, dynamic "process of interpretation and re-interpretation of experiences," whereby teacher development is a process of lifelong learning (p. 122). (2) Professional identity includes both the person and the context, wherein teachers adopt the expectations of their profession in unique ways based upon the value they attach to those expectations. (3) Professional identity is composed of sub-identities, which are tied to different relationships and contexts and vary in degrees of influence on a teacher's overall identity. (4) Professional identity includes the notion of agency, a teacher's active attempts toward professional development and learning based upon their goals and the resources available to them.

Within their discussion of teacher professional identity, Beijaard et al. (2004) note that congruence among sub-identities is essential for teachers. Yet Cooper and Olson (1996) argue contextual factors, in particular historical, sociological, psychological, and cultural factors, are

continually influencing teacher identity and can promote conflicting teacher identities. Indeed, Beijaard et al. identify education change as a factor that can be a source of conflict for teachers' sub-identities. Day and Kington (2008) argue, "Change affects not only teachers' work, but also how teachers feel about their work. There is an unavoidable interrelationship between cognitive and emotional identities, if only because the overwhelming evidence is that teaching demands significant personal investment of these" (p. 8). Simply put, due to the nature of teaching, changes made to teachers' work can create conflicting sub-identities which, in turn, influence teachers' views of their work.

Furthermore, "the more central a sub-identity is [to a teacher's overall identity], the more costly it is to change or lose that identity" (Beijaard et al., 2004, p. 122). In other words, the degree to which education reforms challenge or put into conflict teachers' main sub-identities, the more likely these reforms are to have negative influences on teachers' overall identities. Day and Kington (2008) argue that this influence on identity can affect teachers' commitments to their work and, thus, contend that "research into teacher identities is important as a means of furthering understandings of the job of teaching and what it means to be a teacher in different policy and personal contexts and different times" (p. 9). I argue high-stakes evaluations are a change that has the potential to challenge teachers' identities and, thus, worthy of study.

In her study of teacher identity, Bukor (2015) notes the importance of examining teachers' perceptions because "what an individual perceives may *not* exist, but *his perception* does, and for an individual, his perception is real" (p. 309, emphasis in the original). While there appears to be no empirical research that specifically investigates teacher identity in the context of high-stakes evaluations, there does exists empirical work about teachers' perceptions of high-stakes evaluation. While limited, this work reveals issues related to teacher identity. For

example, Jiang, Sporte, and Luppescu (2015) found that teachers were concerned about student growth measures being used in their evaluations as they saw the measures as narrow and/or unfair. Relatedly, the vast majority of the teachers in their study indicated experiencing increased levels of stress and anxiety related to their evaluation. Similarly, other scholars have found that the use of student growth measures as part of evaluation has a negative influence on teachers' sense of pedagogical and curricular autonomy (Dunn, in press; Hult & Edstrom, 2016; Wright, Shields, Black, Banerjee, & Waxman, 2018) and job satisfaction (Dunn, in press; Wright et al., 2018) and increases mistrust (Hult & Edstrom, 2016).

Scholars suggest that increased stress and loss of self-efficacy due to high-stakes evaluations are likely to negatively impact teachers' performance (Lavigne, 2014). Indeed, when examining teachers' motivation for improvement and continued commitment to the teaching profession during the implementation of a high-stakes teacher evaluation system, researchers discovered that "many teachers experienced… significant negative arousal events and profound losses of satisfaction and commitment to the profession—this despite most being rated as 'highly effective'" (Ford, Van Sickle, Clark, Fazio-Brunson, & Schween, 2017, pp. 202-203). Noting teachers' diminished sense of self-efficacy, autonomy, and satisfaction due in large part to feelings of a lack of support and loss of control, Ford et al. conclude that this evaluation system is unlikely to promote teachers' improvement of practice. Thus, such consequences of high-stakes evaluations could outweigh any gains in student achievement because of their negative effects on teachers (Lavigne, 2014). Learning about teachers' perceptions of unintended consequences is particularly salient to this study as such consequences suggest that a system meant to improve performance is actually creating conditions that work against it. Thus, through

the current study, I add to the theoretical and empirical literature on teachers' perceptions of high-stakes evaluations, which has implications for their identities.

Related Literature

Teachers' perceptions of an education reform influence the ways they take up that reform (or not) (e.g. Coburn, 2001; Moran, 2017); yet this is an under-researched area regarding evaluation reform (Paufler, 2018; Pizmony-Levy & Woolsey, 2017). In the following sections, I highlight some factors that can influence teachers' perceptions of high-stakes evaluations and, thus, the potential for high-stake evaluations to improve the quality of teaching. These factors include their underlying purposes and uses, as well as the feedback provided (or not) as part of the evaluation process.

Underlying Purposes

While performance evaluations can serve various purposes, current teacher evaluations are mainly used to hold teachers accountable for their work and, to a lesser extent, provide feedback to support professional development (Papay, 2012: Steinberg & Donaldson, 2016). Some scholars (e.g. Firestone, 2014; Glickman, Gordon, & Ross-Gordon, 2014; Popham, 1988) argue that evaluations cannot concurrently serve both of these purposes. Indeed, recent research on the Intensive Partnership for Effective Teaching initiative (Stecher et al., 2018) cited the conflicting goals of accountability and improvement as a potential factor in the initiative's overall lack of impact on student achievement. One reason why evaluations may not be able to serve both purposes is because the two purposes rely on conflicting theories of motivation. Firestone (2014) contends that, within high-stakes evaluation systems, the theories of intrinsic and extrinsic motivation "are difficult to reconcile because the [extrinsic] incentive that comes with the threat of losing one's job and the promise of extra pay for high performance can

undermine intrinsic incentives" to improve one's practice (p. 100). In other words, the high-stakes tied to accountability can be counterproductive to improved practice. Furthermore, "the quality of experience and performance can be very different when one is behaving for intrinsic versus extrinsic reasons" (Ryan & Deci, 2000, p. 55). Thus, it seems that performance related to evaluation likely is influenced by what is motivating that performance.

Other scholars (e.g. Derrington & Kirk, 2017; Goe, Biggers, & Croft, 2012) suggest that evaluation can be used concurrently for accountability and professional development purposes, particularly if the evaluation system includes:

1. High-quality standards for instruction;

2. Multiple standards-based measures of teacher effectiveness;

3. High-quality training on standards, tools, and measures;

4. Trained individuals to interpret results and make professional development recommendations;

5. High-quality professional growth opportunities for individuals and groups of teachers;

6. High-quality standards for professional learning. (Goe et al., 2012, p. 2)

In fact, Derrington and Kirk (2017) found that, in the context of accountability, principals can and do use teacher evaluation data for professional development purposes, integrating the professional development into existing school structures. In both cases, evaluation data is used in formative ways to promote professional development. Indeed, the literature suggests that the likelihood of improving practice is related to whether and how the evaluation system functions in summative and/or formative ways.

As a summative function, evaluation data is a summary of past performance used to judge teaching for accountability purposes (Glickman et al., 2014). As Mielke and Frontier (2012) suggest, such evaluation places "teachers in a passive role as recipients of external judgment" with little information to improve (p. 10). Furthermore, Conley and Glasman (2008) contend that the fear experienced by teachers related to the sanctions often tied to summative evaluations may actually discourage instructional improvements. As such, summative evaluations likely have limited usefulness for informing and improving practice.

When used as a formative function, the focus of evaluation is on professional development. In this case, evaluation data is used to provide feedback that assists and supports teachers' professional growth and improvement (Glickman et al., 2014). Indeed, Taylor and Tyler (2012) found evaluation that provides formative feedback can positively influence and have lasting effects on teacher instruction. The key to evaluation improving practice appears to be the type feedback teachers receive.

Formative Feedback

Formative feedback serves an instructional purpose in a learning context (Hattie & Timperley, 2007) and is "information communicated to the learner that is intended to modify his or her thinking or behavior for the purpose of improving learning" (Shute, 2008, p. 154). While there is some variation based upon the learner and their particular phase of learning as well as aspects of the task (Hattie & Timperley, 2007; Shute, 2008), scholars have identified traits of feedback that are necessary to make it instructional, or formative. In general, these traits include feedback that is timely (Hattie & Timperley, 2007; Shute, 2008), specific (Scheeler, Ruhl, & McAfee, 2004; Shute, 2008), focused on the task rather than the learner (Hattie & Timperley,

2007; Shute, 2008), and supportive or positive--but not simply general praise (Hattie &

Timperley, 2007; Scheeler et al., 2004; Shute, 2008).

More specifically in the context of teacher evaluation, to be formative the feedback needs

to focus "on individual performance and on aspects of classroom… practice" (Little, 2006, p.

22), offer suggestions for improved practice (Danielson, 2010), and be perceived as useful by the

teacher (Delvaux, Vanhoof, Tuytens, Vekeman, Devos, & Petegem, 2013). Furthermore, the

feedback needs to include characteristics of effective teaching (Danielson, 1996; Marzano,

Pickering, and Pollock, 2001), be based on observable data (Danielson & McGreal, 2000), be

supported by evidence of student learning (Glickman, 2002), and encourage reflection on and

inquiry into one's practice (Glickman, 2002). Taken together, this type of feedback can promote

teacher growth and increased effectiveness (Feeney, 2007).

While formative feedback is important for learning (Hattie and Timperly, 2007;

Voerman, Meijer, Korthagen, and Simons, 2012), it can also be a motivator when it is given

frequently, allowing the learner to see progress (Bransford, Derry, Berliner, and Hammerness,

2005). Indeed, in a study by Anast-May, Penick, Schroyer, and Howell (2011), teachers

indicated that frequent observations over time accompanied by systematic feedback was vital to

improving their performance, their motivation, and their satisfaction. While a recent study

indicates teachers do receive regular feedback as part of the evaluation process (Prado Tuma et

al., 2018), other studies suggest formative feedback is often absent from the evaluation process

(Donaldson, 2012; Donaldson, 2016; Frase & Streshly, 1994; Weisberg et al., 2009). Without

formative feedback, evaluations are unlikely to improve instructional practices (Goe et al., 2012;

Papay, 2012). Based upon this collective literature, I conjecture that teachers' perceptions of

high-stakes evaluation and its purpose(s) are likely to influence not only how they take up these

evaluations, but also how these evaluations affect their identity and practice. Thus, I utilized this existing literature as part of my data analysis process when examining teachers' perceptions of evaluations.

Methodology

In this paper, I draw upon data from a larger study in which I employed qualitative methods and focused on teachers' perspectives of and experiences with high-stakes evaluations.

Context

Michigan's current teacher evaluation system requires all public school teachers to be evaluated annually based on measures of student growth and classroom observations that utilize one of the four state-approved observation tools: Danielson's Framework for Teaching (FTF), the Marzano Teacher Evaluation Model, the Thoughtful Classroom, and the 5 Dimensions of Teaching and Learning (Michigan Department of Education, *n.d.* b). As part of the evaluation process, teachers must receive feedback within 30 days of their classroom observations and be assigned a summative rating of highly effective, effective, minimally effective, or ineffective at the end of the year (MDE, *n.d.* b). These ratings must be considered in staffing and performance pay decisions. According to the Michigan Department of Education (MDE), the intent of this new system is to "evaluate the teacher's… job performance… while providing timely and constructive feedback" (*n.d.* a, p. 7) to improve their practice. However, with the ratings tied to staffing and seniority no longer guaranteeing a position, these high-stakes evaluations also now have potentially life-changing implications that make a teacher's performance rating significantly more important.

I conducted this research in three phases over the course of two years with teachers in three public elementary schools from different suburban districts in Michigan. While the

ethnicities and socioeconomic status of the student populations varied, the ethnicities of the

teacher populations were similar across schools (see Table 2.1), similarly reflecting the

demographics of the national teaching population (U.S. Department of Education, 2016). As

mandated by the state, all three districts employed evaluation systems that included classroom

observations and measures of student growth. Two of the schools utilized the Danielson

Framework for Teaching (2014) as their observation tool and the other school utilized 5

Dimensions of Teaching and Learning (University of Washington, 2012).

Table 2.1 Demographic Information of Participating Schools

	School A	School B	School C
Grade Levels	Kdg. – 5th	Kdg. – 5th	3rd – 6th
Evaluation Tool	5D+ Teacher Evaluation	Danielson Framework	Danielson Framework
Teachers (Total #) Black Hispanic White Other	23 96% 4% (Asian)	27 100%	31 100%
Female	96%	96%	90.3% (28)
Administrator*	White Female	White Male White Male	White Male White Female
Students (Total #) Black Hispanic White Other	461 3% 4% 85% 8% (Asian)	438 12% 10% 76% 2% (American Indian)	587 1.4% 3.7% 94.4% 0.1% (Asian)
Female	45%	48%	49%
ED	16.4%	85%	34.9%

* School A maintained the same principal over the course of the study. Schools B & C had a change in leadership over the course of the study.

Data Collection

In the spring of 2016, after gaining consent, I collected data from 55 teachers through an

online survey (see Appendix A) in the first of three phases of data collection. After the surveys

were completed, I utilized a qualitative case study approach (Yin, 2013), interviewing teachers with a semi-structured protocol (see Appendix B) that was informed by participants' survey responses. Through these interviews, I was able to "elicit views and opinions from the participants" (Creswell, 2014, p. 190) regarding the current evaluation system and its influence on their practice. I interviewed 14 teachers (all who indicated interest) at a location of their choosing throughout the summer of 2016. I audio-recorded and transcribed verbatim each of these interviews, which lasted 30-60 minutes each, and gave the interviewees a $20 Amazon gift card.

Then, using data from the first two phases, I conducted focus groups (i.e. Byers & Wilcox, 1991; Creswell & Miller, 2000; Sagoe, 2012), drawing on the praxis of culture circles (Freire, 2000; Souto-Manning, 2010) to gain further insights into my research questions and afford my participants an opportunity to speak to my initial findings. Ten of the 14 teachers who were interviewed participated in the focus groups in the summer of 2018. In these small groups of two to four teachers, I used survey data and quotes from the interviews as text-based think alouds, as well as semi-structured interview questions (see Appendix C), to facilitate the focus group discussions. For their time and insights, I compensated the focus group participants with a $100 Amazon gift card.[3]

Throughout this research, I recorded my thoughts and ideas and wrote memos during my data analysis (Maxwell, 2013). Due to the sensitive nature of the information gathered, I treated all evaluation ratings, survey responses, interview and focus group transcripts, field notes, and memos as highly confidential. Furthermore, I assigned pseudonyms to the schools and teachers to protect their identities.

---

[3] I received funding through a College of Education Research Enhancement Fellowship to help cover this cost.

Participants

In this paper, I draw upon the survey, interview, and focus group data of the teachers who chose to participate in all three phases of data collection. These ten teachers self-identified as White or Caucasian and female. As shown in Table 2.2, they taught Kindergarten through 6th grade and included Specials (Art, Music, P.E., Spanish, etc.) teachers. At the beginning of the study, their years of experience ranged from one year to 21 years, with a majority having six or more years of experience and a Master's Degree. At the time of the focus groups, half of the teachers had been rated "highly effective" and the other half as "effective" on their most recent evaluation.

Table 2.2 Demographic Information of Participants*

Teacher	School	Years of Experience	Teaching Position	Highest Degree Held	Evaluation Rating
Myra[4]	A	21	Spanish	Master's	Highly Effective
Lindsay	A	18	Kindergarten	Master's	Effective
Ainsley	B	3	4th Grade	Bachelor's	Effective
Jordan	B	6	5th Grade	Bachelor's	Effective
Sandy	B	8	2nd Grade	Master's	Highly Effective
Avery	C	15	6th Grade	Master's	Highly Effective
Heather	C	17	Spanish	Master's	Highly Effective
Becca	C	3	6th Grade	Bachelor's	Effective
Erin	C	10	6th Grade	Master's	Effective
Christa	C	21	3rd Grade	Master's	Highly Effective

*Focus groups occurred two years after the survey and interviews were conducted. The data in this table reflects the teachers' demographic information at the time of the focus groups.

---

[4] All names are pseudonyms.

Analysis

For my analysis, I examined the transcripts of the participants' interviews and focus group conversations, as well as their responses to the open-ended survey prompts. Utilizing MaxQDA software, I completed an iterative open-coding process looking for emergent themes within and across the data sources (Maxwell, 2013; Miles, Huberman, & Saldana, 2014). I created an initial list of open codes as I analyzed the interviews. I then used the complete list of codes to analyze the open-ended survey responses and focus group conversations, as well as the interviews for a second time. My initial focus was on participants' perceptions of: (1) teacher evaluation in general and (2) how their current high-stakes evaluation systems were influencing their practice. However, as I analyzed my open codes (which included "frustration," "deprofessionalization," and "can't be highly effective"), it became evident that these evaluation systems were also influencing the teachers' identities. Thus, through this coding process, themes related to teachers' perceptions of evaluations as well as the current high-stakes systems' influence on their practice and identity became apparent and are discussed in conjunction with the aforementioned literature in the findings and discussion sections below.

Researcher's Positionality

As a former teacher and administrator in the state where this research takes place, I am intimately familiar with the high-stakes evaluation system under investigation in this study. I have been evaluated as a teacher and evaluated other teachers in this current system. Thus, I have a sense of issues related to high-stakes evaluations. I used these notions and my experiences to build camaraderie with the participants. At the same time, I made a conscious effort to remain open to participants' interpretations, recognizing that it is their lived experiences that are key to this study. Secondly, having worked in two of the three schools in this study, some of the

participants were former colleagues. While I intended to utilize my former relationships with them as an additional way to encourage participation and openness, I recognized that I returned to these schools with a different positionality as a researcher. This may have caused concern for some teachers, whether they knew me previously or not. Therefore, it was important for me to establish that the main intent of my study was to gather teacher voices amidst evaluation reform efforts because I thought (and the teachers seemingly agreed) that their voices are not being heard. I also emphasized the importance of confidentiality and the efforts I would (and did) take to maintain their confidentiality. While some would call these aspects of my positionality limitations, I argue that I am uniquely positioned to investigate this system of evaluation and its influences on teachers.

Findings

In my analysis, some common themes emerged across participants and schools regarding teachers' perceptions of high-stakes evaluation (see Figure 2.1). Generally, the participants indicated a need for evaluation, in particular for the purposes of professional growth/improved practice, accountability, and validation. While the teachers felt there were some aspects of the evaluation process that were helpful, overall they indicated that the current system does little to support their improvement—largely due to a lack of useful feedback. Furthermore, they viewed their current evaluation system as a formality that did not accurately reflect their teaching. While discussing these evaluation-related topics, the teachers also revealed some potential harmful consequences on their identities that highlight issues with systems that try to serve accountability purposes in conjunction with—or to the exclusion of—professional development purposes. I discuss each of these themes in more detail in the following sections.

```
┌─────────────────────┐        ┌─────────────────────┐
│  Desired Purposes   │        │   Limitations of    │
│   for Evaluation    │        │   Current System    │
└─────────────────────┘        └─────────────────────┘
          │                               │
          │   ┌──────────────┐            │   ┌──────────────────────────┐
          ├───│ Professional │            ├───│ Inaccurate Reflection of │
          │   │    Growth    │            │   │        Teaching          │
          │   └──────────────┘            │   │ • Limited data           │
          │                               │   │ • Not all-encompassing   │
          │                               │   │ • Just a show            │
          │   ┌──────────────┐            │   └──────────────────────────┘
          ├───│Accountability│            │   ┌──────────────────────────┐
          │   │              │            ├───│   Doesn't support        │
          │   └──────────────┘            │   │     improvement          │
          │                               │   │ • Formality              │
          │                               │   │ • Missing feedback       │
          │   ┌──────────────┐            │   │ • Can't be highly effective│
          └───│  Validation  │            │   └──────────────────────────┘
              │              │            │   ┌──────────────────────────┐
              └──────────────┘            └───│  Negative Consequences   │
                                              │                          │
                                              │ • Stress                 │
                                              │ • Misues of time         │
                                              │ • Harming professionalism│
                                              └──────────────────────────┘
```
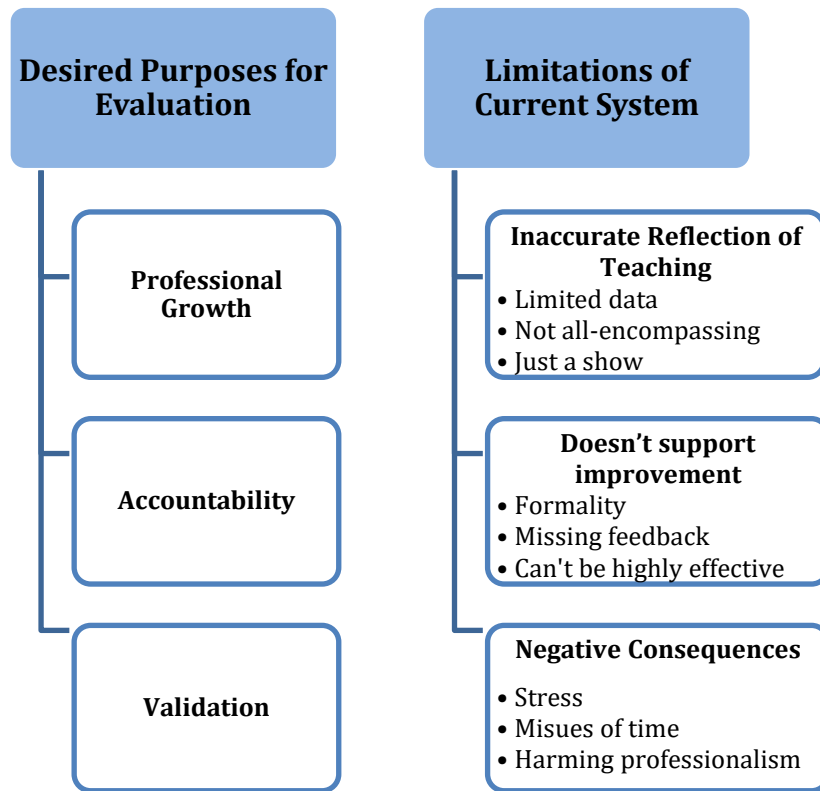
Figure 2.1 Overview of Findings.

We Need an Evaluation System

Across surveys, interviews, and focus groups, participants saw the need for a teacher evaluation system. As Christa, a 3rd grade teacher, voiced, "I do think teachers need to be evaluated. As in any job, it is important to know where your administrator feels that you are strong and areas that you can improve upon." Myra, a Spanish teacher, noted that teachers should be "evaluated based on our skills and our performance, rather than being safe with tenure." While indicating different purposes for evaluation, these teachers—as well as their fellow participants—viewed an evaluation system as necessary and important. Christa's and Myra's responses also reveal two of the three purposes that participants felt evaluations should serve: professional growth and accountability. Some of the teachers also indicated a third purpose: validation of their efforts.

For professional growth. All of the teachers in this study indicated that the role of evaluation should be to help them improve their practice. As Heather, a Spanish teacher, stated,

> There's always new things you can learn, new ways to tweak your instruction so that you're giving your students your best instruction. [Evaluation should] help foster teacher growth and facilitate that they're doing the best job in the classroom that they can.

Similarly, Becca, a novice teacher, explained that teacher evaluation "should be a way to evaluate how well teachers teach and what their strengths are, what their weaknesses are. Identify those areas to grow in and identify ways they're effective." Like those of many other participants, these teachers' comments reveal a desire to enhance their practice and their belief that the evaluation process should serve that purpose.

Notably, a few teachers did identify one avenue their current evaluation system offered as an opportunity for growth: using the observation rubric to identify aspects of effective teaching and for self-assessment. Some of the teachers noted the rubric helped them to see what they should focus on to be considered "highly effective." Echoing other teachers' descriptions of how they use the rubric as a means of self-evaluation, Erin, a 6th grade teacher, explained,

> It's forced me to step back and look at what I do on a daily basis. Like, I've always thought that I was a good teacher. You know, that I do my job. My kids learn. They grow. They enjoy themselves... But I think it's forced me to step back and look and see what specific things am I doing or not doing well.

Interestingly, these teachers seemed to inherently accept that the rubrics identify practices of a highly effective teacher and, thus, could tell them if they are a good teacher. However, other teachers questioned the validity of the rubric itself, as discussed below.

For accountability. Several of the teachers also indicated that evaluation is needed for accountability purposes. The participants noted that all teachers, including themselves, should be held accountable for the instruction that is happening in their classrooms. As Jordan plainly

stated, "I think we need to be held accountable for what we're doing in the classroom." Erin offered some additional insights into teachers' desired purposes for evaluation:

> I think that teacher evaluation should be helping me become a better teacher and kind of holding me accountable or holding us accountable for what we do in the classroom, but mostly it should be a constructive thing that's helping me improve what I do daily in my classroom. Not telling me I'm a bad teacher, comparing me to others, but making me better from year to year to year.

Here, Erin suggests the emphasis of evaluation should be individual professional growth over accountability. Her response also reveals that the current evaluation system may be more focused on blame and the comparison of teachers, rather than on professional growth.

For validation. A third purpose that some of the teachers in this study felt evaluation should serve is to track teachers' growth and validate their efforts. As Sandy, a 2nd grade teacher explained,

> I feel like it should be used as a way to track your growth as a teacher, to help identify those areas that maybe you are struggling in a bit or could use some improvement on. And then have those conversations and like an evaluation piece at the beginning of the year. And then see how you've grown, talk about the ways that you've grown and improved in those areas at the end, by the end of the year.

Sandy sees it as important to monitor and acknowledge the improvements teachers make over the course of a school year. As Heather explained, while the other members of her focus group nodded in agreement,

> I think that, truly reflective teachers who care about doing well, like they just want that pat on the back. Like we've talked about the gold star, the blue ribbon of being highly effective. Like you just want to be validated that what you're doing is noticed.

The teachers wanted to be recognized for doing a good job/being a highly effective teacher and felt that evaluation could validate their efforts.

In summary, the teachers indicated that there should be an evaluation system for the purposes of improved instruction, accountability, and validation. However, as revealed in the

next section, the teachers generally perceived their high-stakes evaluation systems as having little impact on their actual instruction and failing to foster the kind of growth they desired. Furthermore, as hinted above, the teachers suggested such systems had some negative consequences, which were harmful for their identities.

Limitations of the Current System

While acknowledging the need for an evaluation system, the teachers also indicated that the current system did not meet the needs they identified. The teachers noted several reasons for this failure and vocalized many concerns about their high-stakes evaluation systems. Most prevalent were issues related to its inaccurate reflections of teaching, its ineffectiveness in improving practice, and its harmful consequences for teachers. As described in the following subsections, many of these limitations of the current system also negatively influenced the teachers' identities.

Not an accurate reflection. Across all data sources teachers overwhelmingly indicated that their current evaluation system did not capture an accurate reflection of their (or their colleagues') teaching. They noted three distinct reasons for this: (1) It was based on limited, and often flawed, data; (2) It did not encompass all of, or even close to, what they do as a teacher; and (3) It did not reflect what happens on a regular basis, but rather a performance for the observation. In all three situations, the teachers' (sub)identities either were challenged or conflicted.

Limited data. Often describing their evaluations as a "snapshot" of their teaching, the teachers felt their evaluations were based on limited data that did not accurately capture their effectiveness as teachers. As Jordan explained, "I do not understand how my effectiveness can be measured on just a few items: student growth based on test scores, two full lessons out of the

entire year, and a few 'walk throughs' where my administrator may see me in action for 10 minutes a few times throughout the year." The teachers noted there are many things that could not be observed in just a couple of visits to their classroom. Reflecting the sentiments of others, Avery, a sixth grade teacher, shared, "Having one lesson observed does not allow someone to see genuine happenings in the classroom or give enough information to evaluate me." In other words, the teachers felt the limited data points could not truly assess who they were as teachers because they were only snippets of their true teacher identities.

In addition to the limited information gained through observations, the teachers also noted that the measures of student growth used as part of their evaluation were also flawed. Much like the observations, the teachers felt the tests used for measures of student growth were, at best, a snapshot of student learning and, at worst, irrelevant, frustrating, and unfair for teachers and students, as reflected in Christa and Lindsay's comments:

> I think there needs to be some measure. I do. But, to have it tied to one test, one day, I don't think that's fair. (Christa)

> But that's the one thing that I should say scares me in evaluations is when it's going to student growth…. What if you have just that group or certain number of students that's just enough? It can even be one student that knocks your percentage to a point where you don't get your goal, your growth goal. Well, that's not fair. If this kid came in from next to nothing and they came so far. They're not where they should be, but boy they came far… [It] isn't fair just looking at the little bit of data. (Lindsay)

Because it was based upon limited data that easily could be flawed, the teachers questioned the validity of their evaluations. They also did not want their identity as a "highly effective" teacher to be determined by this flawed data and, thus were highly exasperated when it was.

Not all-encompassing. The teachers often expressed frustration that their evaluation system does not encompass all, or even close, to what they do on a daily basis as a teacher or consider factors outside of their classroom that influence what they must do for students. As

Jordan opined, "It's just that the whole idea, the idea of the whole thing, is just so disconnected from what we're actually doing in our classrooms." Thus, the teachers felt they were being measured by incomplete or unreasonable criteria and/or criteria that were disassociated from their actual jobs—and, thus, their teacher identities. Underscoring issues with the evaluation criteria, Becca stated,

> In theory, the structure is useful to score teachers on the same set of standards and expectations. However, I do not find it useful in effectively evaluating all teachers. Since different evaluators have different expectations and classrooms vary in style, I don't think every classroom will look exactly as outlined for a 'highly effective' teacher.

Here, Becca suggested that 'highly effective' teaching can look different in various classrooms but not necessarily align with 'highly effective' on the evaluation rubric. Heather expounded on the issues of using the same criteria for all teachers:

> There are so many areas of the evaluation system that do not apply to a Specials class. The administrators evaluating must either make something up that sounds good or simply give us an 'effective' for lack of understanding how a specific area applies to our teaching assignments.

As a Specials teacher, Heather felt the rubric used to measure her effectiveness did not accurately encompass what a 'highly effective' Specials teacher does, but rather created conflicting identities that her administrators were unable to rectify, often to the detriment of Heather and the other Specials teachers.

Other teachers noted how the requirements for earning 'highly effective' were unrealistic, in that it would "not ever be the case probably" (Sandy) that all students would achieve the same academic benchmark. Not only unrealistic, the teachers felt the evaluation (and those who created it) did not consider factors outside of their control, such as student behavior and students' personal lives, that influence student learning as revealed in the following focus group conversation:

<pre>
Erin: We had a 6th grader pee on the playground this year. Pulled down his pants and
          peed on the playground.
Jordan: That doesn't surprise me…
Erin: But I've got to grow that kid a year... I'm supposed to grow him a year. But they
          don't think about that when they're, they can't. They cannot think about that stuff
          when they're building these evaluation tools. They can't.
Myra: No!
Erin: And they can't think about the kids' home lives… It can't be.
</pre>

These teachers expressed how their evaluations do not capture what they are asked to do as teachers and what they must attend to while also trying to promote student learning. Rather, their evaluations reflected only a small portion of their work as teachers, diminishing their identities by narrowly defining what the teachers' perceived as a multi-faceted, robust role.

*"**Just a show**.*" Calling into question the validity of evaluations and highlighting some potential negative consequences of high-stakes evaluation, the teachers in this study also suggested that evaluations do not accurately reflect teachers' effectiveness because the observed lessons are not representative of daily lessons and self-reported data or evidence can be enhanced. As Ainsley shared, "I believe many teachers are using their [observations] to perform lessons that they usually don't teach. The evaluation may not be a true depiction of that teacher's everyday effectiveness." While noting the "performance" of others, the teachers also readily admitted that they also enacted certain practices during their observations just because those practices were being assessed on the rubric. This was evident in a focus group conversation:

<pre>
Becca: This year I haven't gotten as much feedback and so anything that I've done was
          like language from the rubric. Like one example is 'Students are creating a
          rubric to assess their own work.' Ok, personally, I don't think that the amount of
          time you could spend on students creating a quality rubric is worth it.
Christa: Right. For sure not.
Becca: But you want me to do that? The students are going to make a rubric!
Christa: I can check that box! Yeah. But that did not improve your instruction, right?
Becca: Do I think that improved my?
Christa: Don't you feel like that?
Becca: No. I don't feel that improved my instruction.
Christa: But you're going to do that.
</pre>

> Becca: But, that's what a highly effective teacher's doing!
> Christa: And so you're going to do that for one lesson.
> Becca: Yep.
> Christa: And go back, like you said, to what really works.
> Becca: Yep.

Even though these teachers did not see the value in a particular practice, they felt the compulsion of earning 'highly effective' status as measured by the evaluation rubric and, thus, enacted the practice anyway.

Teachers felt pressure to "put on a show" in other ways as well. In Ainsley's case, she was strongly encouraged by her principal to teach a lesson that would score well on the rubric.

> [He said], 'Just do a lesson you know is going to be successful. Read up on all the domains. Circle all the things that are highly effective. And just try to do that.' So I felt like a fraud doing my second [observation] because... I mean I had to print out all of the domains. Like one of them… is recognizing cultural differences within your students. Well, we had just done a cultural dinner that previous week. So I'm dismissing the students back to their [seats], 'Everyone who found out their ancestors were Polish!' Because I'm sitting there trying to cover every little thing. And it just wasn't authentic, but I was scored really well on that lesson. And it's just, it felt like such a waste and like a show.

Not only did Ainsley acknowledge she was performing to the rubric, but felt that these inauthentic actions had a negative impact on her identity—she felt like a fraud.

The teachers also suggested that performativity could also be found in the self-assessments teachers completed as part of their evaluations. As Erin explained,

> I can make myself look really good on my evaluation, if I provide the right kind of BS evidence. And I can check all those little attribute boxes... I can make myself look good on paper... You can look like a highly effective teacher, but I don't think that makes you a highly effective teacher in reality.

Adding some nuance to the issue, Becca noted,

> It gets tricky because people can provide evidence for just about anything... People want to be effective or highly effective. Like I don't think there's anyone who wouldn't want to be considered that. And not to say that teachers aren't honest about it. But, you might beef up an area.

Becca's statement suggests that the desire to be perceived, and presumably rated, as 'highly effective' encourages teachers to enhance the evidence they provide in their self-assessment.

Teachers feeling the need to 'put on a show', whether during observations or on self-assessments, suggests that they perceive their evaluation system as being used for accountability rather than professional growth. Performing for accountability's sake is counterproductive to actual growth as Ainsley's comment, during her first year of teaching, suggests.

> I know that a lot of people will just give each other tips on like what you should do. But, again, that's not a true reflection on what you do in the classroom day to day. It's what's going to make you 'highly effective'.... But, I keep trying to think of this as a growing experience. And, if it's something I do with my kids and they don't see that, then how are they going to tell me if what I'm doing is actually helpful?

As a new teacher, Ainsley saw her evaluation as an opportunity to grow and, thus, did not feel the need to put on a performance as her colleagues suggested. However, as described above, by the time she was in her third year of teaching, Ainsley succumbed to the pressure of getting a satisfactory rating and performed to the rubric, even though she still felt the importance of being authentic during her observations. Like her colleagues who were providing tips for observations, Ainsley found herself putting on a show for accountability versus professional development purposes, which resulted in conflicting identities for Ainsley.

Does not support improvement. While wanting to do well on their evaluations and improve their instruction, the majority of the teachers indicated their current teacher evaluation system generally does not help them grow professionally. They described the evaluation process as a formality that did little to improve their instruction. Participants repeatedly noted that they received minimal to no formative feedback as part of their evaluation, even though they felt such feedback should be part of the evaluation process. Finally, many of the teachers noted the system

was not motivating because they were told they could not and would not be rated as highly effective.

"Jumping through hoops." Repeatedly, teachers depicted the evaluation process as a formality, "a box that you check." As Jordan noted, "the evaluation process is more of a compliance thing… It's just something that's always going to be there, and here's what you're going to be expected to do." For these teachers, compliance meant their evaluation served someone's purpose other than their own. As Avery explained, "I do not find the current evaluation system useful because it feels like we are jumping through hoops to give the state what they want." Similarly, Erin reflected, "I don't feel like [the current evaluation system] is being utilized to help improve the education that's provided to the students. It's just a 'hoop' to jump through for our administration." Instead of helping them improve their instruction, their evaluation was an exercise in meeting a requirement established by someone else. Responding to a comment made by a fellow focus group member about checking off boxes, Heather reflected:

> When you were talking about, 'How is this making it better for my students? How is this making my instruction better?' It's not. Like checking off those little attributes on the Danielson rubric. How? …How is this making my instruction better at all?

To these teachers, the evaluation process was often just another task to complete, rather than a valuable learning experience worth their time and effort. Thus, the teachers gave little credence to, and often resented, their evaluation systems, which seemingly served someone else's purpose(s) but not their own.

Missing feedback. Reinforcing the idea that their current evaluation was just a formality, the teachers noted the lack of feedback they received. As articulated by Jordan, "I feel that it is treated as a formality with very little useful feedback." The teachers expressed a desire for conversations that included formative feedback to enhance their teaching, but such conversations

were rarely, if at all, a part of their evaluation system. As Christa explained, "There is very little or no conversation about a teacher's strengths and weaknesses. It is definitely not something that helps a teacher grow." Yet all of the teachers desired feedback because, as Avery noted, "I always feel like there's room to grow. And I feel like there's always room for all of us to grow." Jordan echoed these sentiments: "I want authentic feedback. I want to grow and, if I am inadequate, I need to know where I'm inadequate and I need to know what I can do to make it better." Sandy's comment sums up her fellow participants' overarching view of evaluation: "It would be better if it was actually used as a tool to help us improve and find areas of growth." While these teachers felt evaluation should and could provide them with valuable feedback to improve their practice, they did not feel the current system did or could do that.

"Can't be 'highly effective'." The teachers in this study also specifically spoke about their effectiveness ratings, which are a component of their high-stakes evaluation systems. While the teachers wanted to be rated as 'highly effective,' many of them found the ratings arbitrary and frustrating, as Avery's comment reveals:

> [My principal] said this year because of the new system... those that were highly [effective] might not be highly [effective]. And so how? Like, the teaching didn't change. The scores really didn't change. Then how can you go from one to the other just because somebody arbitrarily decided to change the rules?

From the onset, these teachers questioned the purpose and value of the rating system.

Furthermore, the teachers felt the limits placed on the number of teachers who could be rated as 'highly effective' did not support improvement. Noting the illogical and counterintuitive nature of limiting 'highly effective' ratings, Christa contended,

> How can an evaluation system be useful if administrators can make a statement like that? ….Because you should want a whole school of 'highly effectives.' I mean that should be your goal as an administrator is to get everybody there, right?

Thus, some teachers felt efforts to earn 'highly effective' status were fruitless because the way the system was administered made it impossible to achieve that status. For a few teachers, this resulted in indifference: "It made me not care one bit what my rating was. I was like so whoever makes up the best stuff in their evidence... and was the best BSer was going to get highly effective." Christa's comment reflects the performative nature of the evaluation system and her resistance to playing a role in such a system.

For other teachers, this approach to ratings was discouraging and had a negative impact on their well-being. As Sandy described,

> We [were] told that our district is not supposed to rate many, if any, teachers as highly effective. This bothers me, as many of the teachers I know are working SO hard and doing everything they can to help the children... We care so deeply and to be told we can't achieve Highly Effective as a status, no matter how great we are doing, is sad.

Reflecting such sadness, Lindsay shared, "I want to get highly effective, but that's so rarely given out…. So it kind of can make you feel down if you don't get 'highly effective' and you only just get 'effective.'" Also noting the deflating nature of getting an 'effective' rather than 'highly effective' rating, Erin stated, "You feel like you're putting in all that time, effort, money, resource usage... and then to get that 'effective' rating, it's kind of crushing… to your psyche, to your confidence." As Sandy suggested, "Maybe there's a problem with even just having those labels." Wanting and working toward the 'highly effective' rating made an 'effective' rating damaging for these teachers and their identities. Additional negative influences of such a system to teachers and their identities were also revealed.

Consequences of a High-Stakes System

Frustration, indifference, and discouragement were just some of the negative consequences of their current high-stakes evaluations that teachers identified. Among others, they also noted the stress they experienced related to their evaluations, the cumbersome and

90

time-consuming nature of the evaluation process, and the damage it was doing to their sense of professionalism.

Causing stress. While a few of the teachers generally felt indifference toward their evaluation, most of the teachers repeatedly described the evaluation process as stressful and identified several reasons why their evaluation systems caused them stress. For example, Sandy stated,

> In one way it's not a big deal… because I feel like no matter what I'm going to be effective. I'm going to fall in that range and so why stress out about it? But then, what I was going to say in another breath, it's very stressful. And, when those kids, you see them making growth in the classroom and in your reading groups. And then they take a test and they bomb it, it's so frustrating. And it's very stressful to think, 'Oh, this [child]... I know they've made so much growth, but they didn't do well on that test.'

Sandy reveals that one source of stress was the measure of student growth. Another component of the evaluation that caused stress was the effectiveness ratings. While the teachers noted that they felt pressure to achieve 'highly effective,' the stress experienced by the teachers who had achieved that status is particularly noteworthy. For example, Myra shared, "Actually, it's kind of scary to be highly effective because you freak that all you can do is go down." Jordan described the aftermath when this very thing happened to a teacher at her school:

> We had a kindergarten teacher this year that was like very, very stressed out because she had been highly effective and then she wasn't this year…. And she's like so sensitive, too, about it. She was like taking it so personally…She cried a lot this year.... She had a very stressful year.

The stress of failing to maintain her 'highly effective' status caused this Kindergarten teacher a great deal of emotional distress.

The teachers also noted that their principals played a role in the stress they experience. In Erin's case, it was the fear of the unknown:

> It was really stressful the first couple of years because we didn't know how our principal was going to react to it (the new high-stakes evaluation) and what he was going to do

91

with it. And we didn't know what was going to get us the certain rankings. And we didn't know how much we had to justify or explain.

Not knowing how her principal was going to interpret the new evaluation was a source of concern for Erin. In another school, the principal's overzealous focus on evaluation caused stress. As Ainsley explained,

> I think everyone was pretty stressed about evals this year... because it would come up at every single staff meeting. And like just casual conversations with our principal. I mean for Christmas I had to dress up as the elf, like Santa's helper. And so he came up to me in the copy room, 'Hey, I heard you're going to be the elf!' And I'm thinking, 'Yes! Like just an actual conversation with him.' I said, "Yeah." And he goes, 'Well you know that's going to go really well on your Domain 4!' And I'm like, 'Come on!' So it was everything, was kind of focused on this evaluation.

As evidenced in these teachers' comments, their evaluation was a stress-producing experience that was taking a toll on their personal well-being and their professional identities.

Taking up time that could be better spent. In addition to causing stress, the evaluation process took up an inordinate amount of the teachers' time, time they felt could be better spent elsewhere. Describing the time and effort some teachers spent on completing their self-assessment, Myra shared,

> There are all kinds of arguments about how much are you supposed to really write. 'What did you do? Did you write a lot? Did you write like six paragraphs? Did you put up like evidence or what?' I'm like, 'I don't have time to do that!'

Sandy also felt she didn't have the time, nor was it worth her time, to provide evidence of her effectiveness:

> [It's] overwhelming because then you have like ten things that you could upload evidence for. I didn't upload one thing this year...They give us like a date if you want to upload evidence. Otherwise you can wait and see...what your score is, and then if you want to provide evidence like, 'Oh, I think I did better in this area,' you can provide evidence. I was like, 'I've got two toddlers at home. I'm finishing my Master's...Whatever you're going to give me, [give me].'

Also noting how time-consuming completing the self-assessment was, Erin suggested the time

could be better spent on something more meaningful:

> [It's] taking time away from what else you could be doing for your kids. I mean, Lord knows how long it takes all of us to fill out those evaluation rubrics! The amount of time that I spent doing that could have much better been spent prepping for different activities or lessons or differentiating for my kids, but instead I'm sitting up until 11:30 at night trying to figure out what the heck this component means and how I can say that I kind of achieved it.

Erin and the rest of these teachers felt providing evidence as part of their self-assessment was not

a good investment of time—time needed for other roles and responsibilities. Thus, the evaluation

system once again created conflicting identities: the teachers were doing what they felt they had

to for accountability purposes rather than what they knew to be a better use of their time for

themselves and their students.

Harming teachers' sense of professionalism. Among other factors, the stress of the

evaluation process, the inordinate amount of time it required, and its focus on accountability

versus growth negatively influenced the teachers' sense of professionalism. For Myra, the

evaluation was "so much work. And… after a while you start to feel like, 'Am I being

micromanaged? Or can somebody trust me to be a professional?' Clearly if I was highly

effective this year, you can leave me alone for a couple of years." As a highly effective teacher,

Myra felt the evaluation process was an attack on her professionalism. She went on to express, "I

don't think it makes people work harder to try to be effective or highly effective... I think it

makes people bitter." Perhaps, as Myra suggests, it was the micromanaging aspect of the

evaluation that caused such negative, deprofessionalizing feelings.

In focusing on accountability, the damage to professionalism caused by the evaluation

process was particularly noteworthy with the novice teachers. As Becca explained,

As a first year teacher, I feel that I am being compared to other teachers with experience through the evaluation system. I did not expect to be highly effective in my first year of teaching since I still have more to learn, yet there is external pressure to be there already.

Becca clearly saw herself as a novice who was still learning to teach, but the evaluation did not account for her stage of teaching, putting undue pressure on her to perform at an unreasonable level. Perhaps even more concerning is Ainsley's feelings in response to her evaluation experience:

After that conversation I had with my principal (about my observation)... I went home that night and I'm like, 'I shouldn't be a teacher!' Because I thought that lesson went so well, and like obviously I just don't know what I'm doing. And I'm looking up jobs elsewhere, and I'm like, 'This isn't for me.'

Not only was Ainsley's sense of professionalism compromised, her identity as a teacher had been devastated.

Discussion

The findings above indicate that the age-old issue of evaluation's ineffectiveness at improving teachers' practice persists. More importantly, the findings reveal negative consequences associated with an evaluation system that focuses on accountability rather than professional development. Specifically, these evaluation systems constitute a change in teachers' work that is damaging their professional identities. Thus, I contend that an accountability-focused system cannot truly improve teaching because it is based on an inherently flawed theory of action that both maintains the status quo and causes additional harm.

Like teachers in other studies (e.g. Donaldson & Papay, 2015; Moran, 2017), the participants in this study generally viewed evaluation as important and necessary. Furthermore, they felt evaluation should be a mechanism for professional growth that improves their instruction. However, the teachers viewed their current evaluation system as a formality and not

94

a true depiction of their teaching. As suggested elsewhere (e.g. Donaldson, 2016), they also

found the 'paperwork' involved time-consuming and a poor use of their time.

Importantly, the teachers felt their evaluations did little to enhance their practice, largely

due to the lack of formative feedback. Their evaluation process did not provide the specific,

timely, and useful feedback that the literature suggests is needed for teachers to improve their

practice. Any changes to practice the teachers did make were largely superficial, as teachers

performed to the rubric for their observations due to the pressure they felt to meet the 'highly

effective' criteria. While such performance is counterproductive to receiving formative feedback

and, thus, improving practice, it reveals that the teachers perceived their evaluation system as

being used for accountability, rather than professional growth purposes. Perhaps these teachers

would have been more likely to go about their daily instruction during observations if they knew

their evaluation was going to provide actual feedback on that instruction to help them improve

rather than informing their rating, which was tied to their performance pay and job status. In its

current form, the evaluation process provided little to no support for improvement that would

encourage teachers to see the evaluation process as a "growing experience" rather than a

performance or formality for accountability purposes.

The lack of substantive, constructive feedback has been cited repeatedly as a key reason

for the failure of teacher evaluations to effect change (Donaldson, 2012; Donaldson, 2016; Frase

& Streshly, 1994; Weisberg et al., 2009). As the literature suggests, formative feedback is

necessary for improved performance. Thus, it is not likely that teacher evaluations that

emphasize summative over formative purposes will result in real changes in practice that

promote student learning. While some studies (i.e. Dee & Wyckoff, 2015) suggest that high-

stakes evaluation systems improve teacher performance, I question whether they actually

promote improved practice without formative feedback. The teachers in this study described

performing to the rubric for their observations to earn a 'highly effective' checkmark, but

reverted back to what they felt was best for their students when the observations were over.

Additionally, while the participants in this study did not mention it, teachers could improve their

students' test scores by teaching to the test. Thus, an increase in student test scores does not

necessarily reflect that instruction was improved in ambitious ways.

Moreover, while the teachers perceived the current accountability-focused system as

doing little, if anything, to enhance their practice, their perceptions reveal this system is

influencing their professional identities in harmful ways. First, the evaluation created conflicting

identities for the teachers, which had negative implications. Beijaard et al. (2004) argue,

> What is found relevant to the profession, especially in light of the many educational
> changes currently taking place, may conflict with what teachers personally desire and
> experience as good. Such a conflict can lead to friction in teachers' professional identity
> in cases in which the 'personal' and the 'professional' are too far removed from each
> other. (Beijaard & Co, 2004, p. 109)

Indeed, the way the evaluation system defined a 'highly effective' teacher did not align with

what the teachers deemed to be highly effective teaching because the evaluation criteria painted a

narrow and, in some cases, unrealistic view of teaching and encouraged practices that

contradicted what they knew to be best for their students. As evidenced in their comments above,

this misalignment was very frustrating for the teachers. Second, because there were

consequences attached to their evaluations, the teachers felt they had to perform the teacher

identity encompassed in the evaluation criteria, even though they did not agree with and/or saw

no value in the criteria. This performativity resulted in inauthentic teaching that was

counterproductive to the learning opportunity the teachers sought through evaluation.

Furthermore, it suggests that the teachers did not feel they could remain true to their teaching

identities in those moments but, rather, had to conform to the teacher identity outlined in the rubric in order to earn a highly effective rating. For at least one of the teachers, the gap between the teacher identity she wanted and the one described on the evaluation was so great that performing to the evaluation resulted in her identifying as a fraud.

Evaluation systems focused on accountability also affected teachers' identities in other ways. They caused the teachers to feel micromanaged and deskilled, both of which were deprofessionalizing. The ratings attached to evaluations also had implications for the teachers' identities. All of the teachers saw themselves as good teachers who worked hard and did what was best for their students. They identified as highly effective teachers. Because of this, being rated 'effective' versus 'highly effective' was demoralizing or, as Erin noted, crushing to the psyche. Lastly, for some, the process made them question whether they should be teachers at all. In other words, their entire identity as a teacher was put to question.

This study, like others (e.g. Mausethagen, 2013; Wise, Darling-Hammond, McLaughlin, & Bernstein, 1985), suggests that teacher evaluation does not necessarily produce its intended outcomes. One reason for this may be a flawed theory of action underlying the evaluation system. The current system of attaching high-stakes to evaluations reflects a Measure and Punish Theory of Change (Amrein-Beardsley, 2014), which reasons that "change in performance can be evoked by a series of rewards and punishments linked to measured outcomes (Paufler, 2018). Researchers have found that evaluations based on this theory can have negative impacts on teachers such as increased pressure (Collins, 2014) and reduced morale (Collins, 2014; Paufler, 2018). The teachers in this study clearly felt extrinsic pressure to perform in certain—often inauthentic—ways in order to earn highly effective ratings for accountability purposes and, as a

result, experienced reduced morale. However, as suggested elsewhere (i.e. Hill & Grossman, 2013), being held accountable did not improve their teaching to an ambitious level.

Furthermore, the findings in this study point to additional negative impacts of an evaluation system based upon a Measure and Punish Theory of Change, namely the ways it negatively influences teachers' identities. Their high-stakes evaluation system created competing and conflicting identities for the teachers. This finding is particularly significant when one considers that teachers' identities have been linked to their commitment, well-being, sense of agency, and effectiveness (Day & Kington, 2008). Thus, I contend that when teachers perceive evaluation as serving accountability versus professional development purposes such a system will not produce the desired effect of improved teaching and may actually be counterproductive because it negatively influencing teachers' identities.

Implications and Conclusion

The current implementation of high-stakes performance evaluations and the related debate over whether or not they are addressing their intended purpose of promoting teacher growth and student achievement (e.g. Harris & Herrington, 2015) make mine a relevant and timely study. If, as many policymakers claim, the ultimate goal of teacher evaluation truly is to increase student learning, then the emphasis of teacher evaluation should be on teachers' instructional development versus teachers' accountability. As Hill and Grossman (2013) suggest,

> Policy makers must resist the urge to think that simply holding teachers accountable through evaluation systems will result in the changes in teaching that are required for students to meet more ambitious standards. Instead, policy makers must engage in the kind of high-demand, high-support policies that both help teachers learn more about the kinds of instruction envisioned by new standards and to receive the feedback and professional development required to develop new knowledge and skills. (p. 382)

In other words, teachers need formative feedback and support to enhance their practice and increase student achievement.

The findings presented in this paper suggest the current high-stakes evaluation system is not providing the type of feedback that can help teachers improve. Thus, the teachers in this study perceived their evaluation as serving accountability, not professional development, purposes. Moreover, these teachers identified several negative consequences associated with such an evaluation system that negatively influenced their teacher identities. Specifically, the evaluation system challenged their teacher identities, causing stress, frustration, and deprofessionalization. Thus, it is doubtful their current evaluation system and others like it are producing the desired effect and may actually be counterproductive.

More research is needed to confirm the results of this study and explore other potential negative consequences of an evaluation system that focuses on accountability over professional development. Considering the principal's central role in the evaluation process, it would also be beneficial to know more about how principals can promote evaluation for professional growth and mitigate some of the negative consequences of a system focused on accountability. Finally, it is important to critically examine current evaluation systems and identify evaluation policies and practices teachers find meaningful and useful to improve teaching and enhance their identities. ESSA provides an opportunity for state policymakers to make changes to the current teacher evaluation system. This study and others like it suggest that teachers' perceptions of evaluation should inform that change.

APPENDICES

Teacher Survey Questions

1. Gender: _____
2. Age: 21-25 26-30 31-35 36-40 41-45 46-50 51-55 56-60 61-65 66-70 over 70
3. Race: _____
4. Years of experience in a professional teaching position. (This is my _____ year in a professional teaching position.): _____
5. Current school: _____
6. Years in current school: _____
7. Current position: (i.e. lower elementary teacher, upper elementary teacher, specials teacher (Music, Art, P.E., Spanish, Media, etc.), Special Education teacher, etc.) _____
8. Years in current teaching position: _____
9. Highest Level of Education: Bachelor's Master's Ed. Specialist Ph.D./Ed.D.
10. Most recent evaluation rating: Highly Effective Effective Minimally Effective Ineffective

For the purposes of this survey, collaboration is defined as sharing knowledge and resources with the ultimate goal of improving teaching and student learning.

For each of the following statements, please indicate if you Strongly Agree, Agree, Disagree, or Strongly Disagree.

11. I am aware of the state requirements regarding teacher evaluation.
12. I agree with the state requirement of rating teachers through the evaluation process.
13. The rating system encourages teachers to improve their instruction.
14. I understand the current teacher evaluation system.
15. I think the current teacher evaluation system is fair.
16. I think it is important that teachers are rated on a scale from Highly Effective to Ineffective.
17. I think the current teacher evaluation system accurately assesses my teaching ability.
18. I think my current evaluation rating accurately reflects my teaching ability.
19. The current teacher evaluation system helps me grow professionally.
20. Under the current teacher evaluation system, I am being compared to my colleagues.
21. My administrator uses the current evaluation system to help me improve my practice.
22. I am concerned about how my evaluation rating compares to my colleagues' evaluation ratings.
23. The current evaluation system encourages me to work with my colleagues to improve my practice.
24. I see collaboration as an important part of my professional growth.
25. I respect my colleagues as professional educators.
26. Working with my colleagues is important regardless of what's required.
27. I enjoy collaborating with my colleagues.
28. Collaborating with my colleagues is beneficial to my teaching.
29. My administrator values collaboration.

30. Since the new teacher rating system has been put in place, I am less likely to share my ideas with my colleagues.
31. Since the new teacher rating system has been put in place, I am less likely to share my resources/materials with my colleagues.
32. It is in my best interest to collaborate with my colleagues.
33. I am proud to be a teacher.
34. I enjoy my job.

Open-ended
Please respond to the following prompts:

I collaborate with my colleagues because… I do not collaborate with my colleagues because…

I find the current evaluation system useful because… I do not find the current evaluation system useful because…

Other comments:

If you would be willing to be interviewed, please type your name here:
_____

APPENDIX B:
Teacher Semi-Structured Interview Protocol

The participating teachers will be asked to respond to the following prompts individually during an interview.

1. Please tell me a little about your background, education, and teaching experiences.

2. What do you see as the role of teacher evaluation?

3. Do you think that evaluation plays a role in teacher growth? Why or why not?

4. Has the new evaluation system which requires teachers to be rated highly effective, effective, minimally effective, or ineffective affected how you perceive the evaluation process? If so, how?

5. What benefits do you see to this new system?

6. What challenges or constraints do you see with this system?

7. Do you have concerns about this new evaluation system? If so, what are they?

8. Has the new evaluation system influenced your work environment? If so, how? If not, why do you think it hasn't?

9. How do you define/describe collaboration among teachers?

10. Has the new evaluation system influenced your collaboration with your colleagues? If so, how? If not, why do you think it hasn't?

11. Do you think the new evaluation system has influenced collaboration among your colleagues? If so, how? If not, why do you think it hasn't?

12. What are some alternative evaluation methods that you think would foster collaboration among teachers?

13. Is there anything else you'd like to share with me?

14. Is there a question you thought I'd ask that I didn't?

In addition to particular prompts, the interviewer will follow-up on initial responses and ask pressing questions using questions such as:
- What do you mean by…?

- How did you do…?

- Tell me more about…

- Is there anything else you'd like to add that we have yet to discuss?

APPENDIX C:
Focus Group Questions

The following questions were provided to the focus group participants prior to their focus group sessions and used to varying degrees, along with follow-up questions, during the focus group sessions to facilitate discussion:

1. What are your thoughts as you read the findings/excerpts regarding collaboration/teacher well-being/professionalism? Do you agree? Disagree?

2. What, if anything, has changed regarding your evaluation system/tool/process/etc. over the last two years?

3. How have you made/do you make sense of the current evaluation system? Your rating?

4. What dilemmas, if any, have you faced as a result of/in relation to your evaluation/the evaluation system?

5. How have high-stakes evaluations influenced you personally? Your well-being? Your sense of professionalism?

6. How have high-stakes evaluations influenced you professionally? How have they influenced your work environment? Your work with colleagues? Your instruction?

7. What else is important for me, administrators, policymakers, etc. to know about the influences of high-stakes evaluations on collaboration/well-being/professionalism/ instruction?

8. How could the evaluation system at your school be improved? What actions do you think teachers could take to improve the current evaluation system?

9. What role does politics play in teacher evaluation?

10. In what ways, if any, does the current teacher evaluation system reflect the state of education in the United States?

REFERENCES

REFERENCES

Amrein-Beardsley, A. (2014). *Rethinking value-added models in education: Critical perspectives on tests and assessment-based accountability.* New York, NY: Routledge.

Anast-May, L., Penick, D., Schroyer, R., & Howell, A. (2011). Teacher conferencing and feedback: Necessary but missing!. *International Journal of Educational Leadership Preparation*, *6*(2), ISSN 2155-9635.

Beauchamp, C., & Thomas, L. (2009). Understanding teacher identity: An overview of issues in the literature and implications for teacher education. *Cambridge Journal of Education, 39*(2), 175-189.

Beijaard, D., Meijer, P. C., & Verloop, N. (2004). Reconsidering research on teachers' professional identity. *Teaching and Teacher Education*, *20*(2), 107-128.

Bransford, J., Derry, S., Berliner, D., & Hammerness, K. (2005). Theories of learning and their roles in teaching. In L. Darling-Hammond & J. Bransford (Eds.), *Preparing teachers for a changing world: What teachers should know and be able to do* (pp. 40-87). San Francisco, CA: Jossey-Bass.

Bukor, E. (2015). Exploring teacher identity from a holistic perspective: Reconstructing and reconnecting personal and professional selves. *Teachers and Teaching*, *21*(3), 305-327.

Byers, P. & Wilcox, J. (1991). Focus groups: A qualitative opportunity for researchers. *The Journal of Business Communication, 28*(1), 63-78.

Coburn, C. E. (2001). Collective sensemaking about reading: How teachers mediate reading policy in their professional communities. *Educational Evaluation and Policy Analysis*, *23*(2), 145-170.

Cohen, J., & Goldhaber, D. (2016). Building a more complete understanding of teacher evaluation using classroom observations. *Educational Researcher*, *45*(6), 378-387.

Collins, C. (2014). Houston, we have a problem: Teachers find no value in the SAS Education Value-Added Assessment System (EVAAS®). *Education Policy Analysis Archives*, *22*(98), 1-42.

Conley, S. & Glasman, N. (2008). Fear, the school organization, and teacher evaluation. *Education Policy, 22*(1), 63-85.

Cooper, K., & Olson, M. R. (1996). The multiple Ts' of teacher identity. In M. Kompf, W. R. Bond, D. Dworet, & R. T. Boak (Eds.), *Changing research and practice: Teachers' professionalism, identities, and knowledge* (pp.78-89). Washington, D.C.: Falmer Press.

Creswell, J. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches*. Thousand Oaks, CA: SAGE Publications.

Creswell, J. & Miller, D. (2000). Determining validity in qualitative inquiry. *Theory Into Practice, 39*(3), 124-130.

Danielson, C. (1996). *Teacher evaluation: Enhancing professional practice: A framework for teaching*. Alexandria, VA: Association of Supervision and Curriculum Development.

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

Danielson, C. (2014). The Framework for Teaching Evaluation Instrument. Retrieved from http://danielsongroup.org/framework/

Danielson, C., & McGreal, T. L. (2000). *Teacher evaluation to enhance professional practice*. Alexandria, VA: Association for Supervision and Curriculum Development.

Darling-Hammond, L. (2015). Can value added add value to teacher evaluation? *Educational Researcher, 44*(2), 132-137.

Day, C., & Kington, A. (2008). Identity, well-being and effectiveness: The emotional contexts of teaching. *Pedagogy, Culture & Society*, *16*(1), 7-23.

Dee, T. S., & Wyckoff, J. (2015). Incentives, selection, and teacher performance: Evidence from IMPACT. *Journal of Policy Analysis and Management*, *34*(2), 267-297.

Derrington, M. L., & Kirk, J. (2017). Linking job-embedded professional development and mandated teacher evaluation: Teacher as learner. *Professional Development in Education*, *43*(4), 630-644.

Delvaux, E., Vanhoof, J., Tuytens, M., Vekeman, E., Devos, G., & Van Petegem, P. (2013). How may teacher evaluation have an impact on professional development? A multilevel analysis. *Teaching and Teacher Education, 36*, 1–11.

Donaldson, M. L. (2012). *Teachers' perspectives on evaluation reform*. Retrieved from Center for American Progress website: https://cdn.americanprogress.org/wp-content/uploads/2012/12/TeacherPerspectives.pdf

Donaldson, M. L. (2016). Teacher evaluation reform: Focus, feedback, and fear. *Educational Leadership*, *73*(8), 72-76.

Donaldson, M. L., & Papay, J. P. (2015). An idea whose time had come: Negotiating teacher evaluation reform in New Haven, Connecticut. *American Journal of Education*, *122*(1), 39-70.

Dunn, A.H. (in press). "A vicious cycle of disempowerment:" The relationship between teacher morale, pedagogy, and agency in an urban high school. *Teachers College Record*.

Feeney, E.J. (2007). Quality feedback: The essential ingredient for teacher success. *The Clearing House, 80*(4), 191-197.

Firestone, W. (2014). Teacher evaluation policy and conflicting theories of motivation. *Educational Researcher, 43*(2), 100-107.

Ford, T., Van Sickle, M., Clark, L., Fazio-Brunson, M. & Schween, D. (2017). Teacher self-efficacy, professional commitment, and high-stakes teacher evaluation policy in Louisiana. *Educational Policy, 31*(2), 202-248.

Frase, L. E., & Streshly, W. (1994). Lack of accuracy, feedback, and commitment in teacher evaluation. *Journal of Personnel Evaluation in Education*, *8*(1), 47-57.

Freire, P. (2000). *Pedagogy of the oppressed* (M. B. Ramos, Trans.) (30th anniversary edition). New York, NY: Continuum.

Gee, J.P. (2001). Identity as an analytic lens for research in education. *Review of Research in Education*, *25*, 99–125.

Glickman, C. (2002). *Leadership for learning: How to help teachers succeed*. Alexandria, VA: Association of Supervision and Curriculum Development.

Glickman, C.D., Gordon, S.P., and Ross-Gordon, J.M. (2014). *Supervision and instructional Leadership: a developmental approach* (9th ed.). Boston, MA: Allyn & Bacon.

Goe, L., Biggers, K., & Croft, A. (2012). *Linking teacher evaluation to professional development: Focusing on improving teaching and learning*. Research & Policy Brief: National Comprehensive Center for Teacher Quality.

Goldstein, D. (2014). *The teacher wars: A history of America's most embattled profession*. New York, NY: Doubleday.

Harris, D. & Herrington, C. (Eds.). (2015). Value added meets the schools: The effects of using test-based teacher evaluation on the work of teachers and leaders [Special issue]. *Educational Researcher*, *44*(2), 71-76.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81-112.

Hill, H.C. & Grossman, P. (2013). Learning from teacher observations: Challenges and opportunities posed by new teacher evaluation systems. *Harvard Educational Review, 83*(2), 371-384.

Hult, A., & Edström, C. (2016). Teacher ambivalence towards school evaluation: Promoting and ruining teacher professionalism. *Education Inquiry*, *7*(3), 305-325.

Jiang, J., Sporte, S., & Luppescu, S. (2015). Teacher perspectives on evaluation reform: Chicago's REACH students. *Educational Researcher, 44*(2), 105-116.

Lavigne, A. (2014). Exploring the intended and unintended consequences of high-stakes teacher evaluation on schools, teachers, and students. *Teachers College Record*, *116*(1), 1-29.

Little, J. W. (2006). *Professional community and professional development in the learning-centered school* (NEA Best Practices Working Paper Series). Berkeley, CA: University of California, Berkeley.

Mausethagen, S. (2013). A research review of the impact of accountability policies on teachers' workplace relations. *Educational Research Review*, *9*, 16-33.

Marzano, R. J., Pickering, D.J. & Pollock, J.E. (2001). *Classroom instruction that works: Research-based strategies for increasing student achievement*. Alexandria, VA: Association for Supervision and Curriculum Development.

Maxwell, J. A. (2013). *Qualitative research design: An interactive approach* (3rd ed.). Thousand Oaks, CA: SAGE Publications.

Michigan Council for Educator Effectiveness. (2013, July). *Building an improvement-focused system of educator evaluation in Michigan: Final recommendations*. Retrieved from http://www.mcede.org/

Michigan Department of Education. (*n.d.* a). *2013-14 Educator evaluations & effectiveness in Michigan*. Retrieved from https://www.michigan.gov/documents/mde/2013-14_Educator_Evaluations_and_Effectiveness_485909_7.pdf

Michigan Department of Education (*n.d.* b). *Michigan educator evaluations at-a-glance*. Retrieved from https://www.michigan.gov/documents/mde/Educator_Evaluations_At-A-Glance_522133_7.pdf

Mielke, P., & Frontier, T. (2012). Keeping improvement in mind. *Educational Leadership*, *70*(3), 10-13.

Miles, M. B., Huberman, A. M., & Saldana, J. (2014). *Qualitative data analysis: A method sourcebook.* Thousand Oaks, CA: SAGE Publications.

Moran, R. M. R. (2017). The impact of a high stakes teacher evaluation system: Educator perspectives on accountability. *Educational Studies, 53*(2), 178-193.

Papay, J. (2012). Refocusing the debate: Assessing the purposes and tools of teacher evaluation. *Harvard Educational Review, 82*(1), 123-167.

Paufler, N. A. (2018). Declining morale, diminishing autonomy, and decreasing value: Principal reflections on a high-stakes teacher evaluation system. *International Journal of Education Policy and Leadership*, *13*(8), 1-15.

Pizmony-Levy, O., & Woolsey, A. (2017). Politics of education and teachers' support for high-stakes teacher accountability policies. *Education Policy Analysis Archives, 25*(87), 1-31.

Popham, W. (1988). The dysfunctional marriage of formative and summative teacher evaluation. *Journal of Personnel Evaluation in Education, 1*(3), 269–273.

Prado Tuma, A., Hamilton, L.S., & Tsai, T. (2018). *A nationwide look at teacher perceptions of feedback and evaluation systems: Findings from the American Teacher Panel*. Santa Monica, CA: RAND Corporation.

Ryan, R. M., & Deci, E. L. (2006). Self-regulation and the problem of human autonomy: Does psychology need choice, self-determination, and will?. *Journal of Personality*, *74*(6), 1557-1586.

Sagoe, D. (2012). Precincts and prospects in the use of focus groups in social and behavioral science research. *The Qualitative Report, 17*(29), 1-16.

Scheeler, M. C., Ruhl, K. L., & McAfee, J. K. (2004). Providing performance feedback to teachers: A review. *Teacher Education and Special Education, 27*(4), 396-407.

Shute, V. (2008). Focus on formative feedback. *Review of Educational Research, 78*(1) 153-189.

Souto-Manning, M. (2010). *Freire, teaching, and learning: Culture circles across contexts*. New York: P. Lang.

Stecher, B. M., Holtzman, D. J., Garet, M.S., Hamilton, L.S., Engberg, J., Steiner, E.D., Robyn, A., Baird, M.D., Gutierrez, I.A., Peet, E.D., Brodziak de los Reyes, I., Fronberg, K., Weinberger, G., Hunter, G.P., & Chambers, J. (2018). *Improving teaching effectiveness: Final Report: The Intensive Partnerships for Effective Teaching through 2015–2016*. Retrieved from https://www.rand.org/pubs/research_reports/RR2242.html

Steinberg, M. P., & Donaldson, M. L. (2016). The new educational accountability: Understanding the landscape of teacher evaluation in the post-NCLB era. *Education Finance and Policy, 11*(3), 340-359.

Taylor, E. S. & Tyler, J. H. (2012). Can teacher evaluation improve teaching? Evidence of systematic growth in the effectiveness of midcareer teachers. *Education Next, 12*(4), 78-84.

U.S. Department of Education, Office of Planning, Evaluation and Policy Development, Policy and Program Studies Service. (2016). *The state of racial diversity in the educator*

workforce. Retrieved from https://www2.ed.gov/rschstat/eval/highered/racial-diversity/state-racial-diversity-workforce.pdf

University of Washington, Center for Educational Leadership. (2012). 5 Dimensions of Teaching and Learning. Retrieved from http://info.k-12leadership.org/5-dimensions-of-teaching-and-learning?_ga=1.57564131.87949747.1473787149

Voerman, L., Meijer, P. C., Korthagen, F. A., & Simons, R. J. (2012). Types and frequencies of feedback interventions in classroom interaction in secondary education. *Teaching and Teacher Education*, *28*(8), 1107-1115.

Weisberg, D., Sexton, S., Mulhern, J., & Keeling, D. (2009). *The widget effect: Our national failure to acknowledge and act on differences in teacher effectiveness*. (J. Schunck, A. Palcisco, & K. Morgan, Contributing authors.) Retrieved from The New Teacher Project website: https://tntp.org/assets/documents/TheWidgetEffect_2nd_ed.pdf

Wise, A. E., Darling-Hammond, L., McLaughlin, M. W., & Bernstein, H. T. (1985). Teacher evaluation: A study of effective practices. *The Elementary School Journal*, *86*(1), 61-121.

Wright, K. B., Shields, S. M., Black, K., Banerjee, M., & Waxman, H. C. (2018). Teacher perceptions of influence, autonomy, and satisfaction in the early Race to the Top era. *Education Policy Analysis Archives*, *26*(62), 1-29.

Yin, R. K. (2013). *Case study research: Design and methods*. Thousand Oaks, CA: SAGE Publications.

ARTICLE THREE: MOVING TOWARD HUMANIZING RESEARCH IN THE

CONTEXT OF A DEHUMANIZING POLICY[5]: CULTURE CIRCLE-INFORMED

FOCUS GROUPS

Introduction

...to foreground the worth of such processes of humanization for inquiry and for society.
Paris & Winn (2014)

Sandy: I don't feel like we're always very well... respected…. You hear a lot of comments about, or people sharing things on social media, about how teachers are glorified babysitters or Common Core this or things that aren't even our choices or we're just told we have to do them. And so I feel like... it used to be that teachers—or maybe we just didn't have social media so I didn't, we didn't, really know how people felt. I don't know. But, you had this sort of system where your principal [was] in your class more and you [didn't] have the anxiety of him coming in.

Christa: Right.

Sandy: He's just coming in and it's just okay because it's not something that could possibly count against you.

Christa: Right.

Sandy: It's just...

Christa: It's just a coworker.

Sandy: Everybody's just professional and working together.

Christa: Yes. Yes.

Sandy: And I feel like… teachers were seen in a better light then. And now we have this system that is more, I feel like, sometimes negative.

Christa: Like punitive.

Sandy: And now we have this view of teachers that we're not doing our jobs.

(Focus Group, June 2018)

Sandy and Christa[6], two elementary teachers, are reflecting on their high-stakes

evaluation systems in a focus group with other educators. Even though they teach in different

school districts and have only recently met, they easily relate to and affirm each other's

experiences with—and within—a system that evaluates and rates them based on students' test

scores and a random classroom observation from a school administrator. As the teachers note,

these systems have a punitive element that their former systems did not, purporting a broader,

---

[5] A riff on Irizarry and Brown's (2014) title.
[6] All names are pseudonyms.

negative discourse about teachers. The disconnect between this discourse and the work they do is palpable, both in their classrooms and in our discussion.

Referencing two neoliberal reforms (high-stakes evaluations and the Common Core State Standards), Sandy's and Christa's comments underscore that they, like other public school teachers across the United States, are teaching in an era when neoliberalism is intensely influencing their work and their identities. Neoliberalism is an economic ideology that calls for improving education efficiency through reduced public expenditures. Supporters of neoliberalism aim to use education systems for economic purposes and have introduced deregulation, privatization, and competition into education through, among other efforts, charter schools, high-stakes testing, and accountability (Hursh, 2004; Robertson, 2008).

Among other effects, neoliberal policies have narrowed the focus and flexibility of curriculum, emphasized achievement on standardized tests as evidence of learning, and placed intense scrutiny on teachers, tying their performance to consequential sanctions (Irizarry & Brown, 2014). Scholars (e.g. Apple, 2001; Dunn, in press) suggest that such neoliberal policies are *deprofessionalizing* for teachers, which has been linked to teacher dissatisfaction, frustration, and attrition (Dunn, in press; Ingersoll, Merrill, & Stuckey, 2014). Furthermore, according to Blackburn's (2014) definition of dehumanizing, it is likely these policies also are *dehumanizing* for teachers in that they actively take away teachers' individuality, creativity, and humanity and treat teachers like a number or an object. High-stakes evaluation, which rests on the rationale that rewarding and firing teachers based on individual performance will improve their practice, is one such neoliberal policy and is the focus of this research.

My Responsibilities: Part One

Jordan: And I feel like it's abuse, too. Like one of our teachers, she just quit. She left. She… doesn't even have another job. This was her first job... This is her third year teaching… This year, like it (the evaluation process), she just, she's done.

<div align="right">(Focus Group, June 2018)</div>

Because research has a history of dehumanizing individuals and groups (Freire, 2000; hooks, 1994; Tuck & Yang, 2014), researching a dehumanizing policy in traditionally dehumanizing ways would likely cause additional harm to participants. In my research of high-stakes evaluations, I most certainly did not want to cause additional harm to my participants. Furthermore, because of my respect and concern for teachers broadly, I wanted my participants to feel that their participation in my research was worthy of their time and that they were getting something in return. Seeking to move away from positivist notions of research and to make the research experience more humanizing for my participants, I invoked Paris and Winn's (2014) humanizing research methodology in which researchers and participants engage in critical conversation and build relationships of respect and care. In particular, I facilitated focus groups where I drew upon the praxis of culture circles (Freire, 2000; Souto-Manning, 2007, 2010).

In the following sections, I "provide a roadmap to foreground the worth of such processes of humanization for inquiry and society" (Paris & Winn, 2014, p. xiv) for scholars who are researching teachers and their work in the context of a neoliberal policy. I begin by providing an overview of humanizing research, focus groups, and culture circles. Then, I describe my enactment of culture circle-informed focus groups and share vignettes from and participants' reflections on these focus groups. I highlight what this method afforded them and, thus, establish a case for using such a method with teacher participants. Ultimately, I assert that humanizing research can be applied and should be considered necessary in social science research more broadly, particularly in the context of neoliberalism in K-12 schools.

Humanizing Research

I just felt like I grew as I took time to think about different aspects of my profession and evaluation, and as I found ways to articulate my experience with other professionals.
(Focus Group Participant reflection, July 2018)

Grounded in ethnographic and decolonizing methodologies and informed by Freire's (2000) notions on the necessity of relationships in the dialogic process of achieving critical consciousness, humanizing research is a methodological stance for enacting social justice research that centers participants and their social and political contexts (Paris, 2011; Paris & Winn, 2014). As a methodology, humanizing research "seeks to decolonize and thus humanize the research process" (Paris & Winn, 2014, p. xiii). Researchers engaging in humanizing research aim to learn with their participants through critical thinking, where both gain understandings "to push against inequities not only through the findings of research but also through the research act itself" (Paris, 2011, p. 140). Thus, not only is the research process equally, if not more, important than the findings, the two are inextricably linked.

Key attributes of humanizing research are a genuine respect for the participants as individuals from and with whom much can be learned and the building of trusting, reciprocal relationships (Irizarry & Brown, 2014; Paris, 2011; Paris & Winn, 2014). To position participants in this way and facilitate these relationships, humanizing researchers reflect upon their own positionalities and the roles they play in the work, as well as whom the research is for (Green; 2014; Paris, 2011; Souto-Manning, 2014). They also consider their entry into the work and how participants will become involved (Paris, 2011), in addition to their fulfillment of commitments and departure from the work (Mangual Figueroa, 2014). Furthermore, in enacting humanizing research, researchers share of themselves as they ask their participants to do, rather than maintaining the notion of a neutral researcher (Paris, 2011). Ultimately, as Paris and Winn

115

(2014) assert, researchers need to always be "mindful of how critically important it is to respect the humanity of the people who invite us into their worlds and help us answer questions about educational, social, and cultural justice" (p. xv). Humanizing methods provide an approach to research that both respects and sustains the humanity of participants.

Scholars (e.g. Blackburn, 2014; Green, 2014; Irizzary & Brown, 2014; Paris, 2011) mainly have enacted humanizing research with youth "who are oppressed and marginalized by systems of inequality based on race, ethnicity, class, gender, and other social and cultural categories" (Paris, 2011, p. 140) in concerted efforts to disrupt that oppression and marginalization. Establishing the critical nature of humanizing research for these participants and communities, Paris (2011) also opens the door for this type of research in other spaces, noting that humanizing research is a methodological stance that "is important in all research" (p. 140). Thus, drawing on the humanizing work of others, I utilize aspects of humanizing research with my participants—who are not considered oppressed or marginalized in the ways described above because of the privileges that come with their race and socioeconomic status, but do feel the deprofessionalizing and dehumanizing effects of neoliberalism. By centering teachers and the political contexts in which they teach and participating in critical thinking with them, I attempt to engage in a type of research that "is not only ethically necessary but also increases the validity of the truths we gain through research" (Paris, 2011, p. 137).

Importantly, as a White researcher working with White participants, I seek not to appropriate humanizing research, which was and is grounded in the work of scholars of Color working in and with communities of Color, but rather to understand the ways I can make the research experience more humanizing for my participants. Furthermore, by incorporating aspects of humanizing research into my study, I attempt to implicate and counter the deprofessionalizing

116

and dehumanizing effects of neoliberal policies on teachers. Thus, in this paper, I describe what humanizing research can look like for me, a White researcher, doing particular research with White teachers in the context of a neoliberal policy. I also argue that enacting focus groups based on the praxis of culture circles can afford participants within this context opportunities for storying, finding solidarity, and reading and questioning their political context, and, thus, engage them in a humanizing experience.

Centering Participants and Their Lived Experiences

Thank you for asking 'real' teachers about their thoughts and opinions.
I think we had so much to say because, while we are important stakeholders, very few decision-makers seek our input/advice/opinions.
(Focus Group Participant reflection, July 2018)

If research is to be humanizing, it is "important to look closely and listen carefully in order to understand the perspectives and experiences of participants in their own terms rather than superimposing our own perspectives of what is problematic and needs to be transformed" (Souto-Manning, 2014, p. 201). Wanting to learn how high-stakes evaluations were influencing teachers' practice, their work with colleagues, and their personal well-being, I deliberately sought out and listened to the voices of teachers, those who felt and could speak to the deep impact of these evaluations. The teachers in my study taught Kindergarten through 6th grade. Some of them were novice teachers, while others had over 20 years of teaching experience at the time of the study. All of the teachers self-identified as female and White or Caucasian. Some of the teachers had earned their Master's Degree, while others were working on it. Some were rated "effective" in their most recent evaluations, and others were rated "highly effective." These teachers, ten in total, participated in all phases of my data collection—an initial survey, interviews, and focus groups—sharing their time, insights, and something of themselves.

At the time of the study, these participants taught in three suburban districts in Michigan where high-stakes evaluation systems were being implemented. These evaluations systems were just one of numerous neoliberal education reforms that have been implemented in Michigan. For example, Michigan leads the nation in for-profit charter schools (Miron & Gulosino, 2013), has prescribed K-12 academic standards, and requires state standardized testing annually beginning in the third grade. Due to the preponderance of neoliberal education policies enacted, it seems imperative to consider them an integral, and influential, part of the context in which public school teachers in Michigan teach. Indeed, the teachers in my study often referenced increased stress and decreased professionalism (Guenther, 2019), which often could be linked back to a neoliberal policy.

To center my participants and their lived experiences and avoid (as best as I or any researcher can) superimposing my own perspectives on those of my participants, I must critically examine my own intersecting identities within the work (Green, 2014). As a White novice researcher who recently left the classroom after many years in K-12 schools and who has a spouse and many friends still teaching in these spaces, I possess many intersecting identities that I need to examine within my study on high-stakes teacher evaluations. Having been evaluated as a teacher and evaluated other teachers as an administrator, I am intimately familiar with the current high-stakes evaluation system. I also have listened to my spouse, teacher colleagues, and friends express disappointment, frustration, and resentment over the current evaluation system. Thus, I have a sense of the challenges related to high-stakes evaluations.

As a White female, I am the demographic of the majority of teachers in the U.S. (U.S. Department of Education, 2016) and all of the teachers who participated in my study. This likely made me relatable to my participants. It also means our collective perspective is influenced by

privileges associated with being White, and, as a result, are limited. Furthermore, having worked

in two of the three schools in my study, I am a former colleague to some of my participants.

While I hoped the relationships I had with these former colleagues would encourage

participation and openness, I returned to these schools with a different positionality as a

researcher. This may have caused concern for some teachers, whether they knew me previously

or not. Lastly, as a White researcher, I am afforded certain privileges, including, among others,

the ability to enter and depart the work largely at my discretion. By examining and

acknowledging my intersecting identities throughout the research process, I am better able to

resist the urge to impose my "own understandings, assumptions, and experiences upon" my

participants and instead actively seek my participants' understandings (Souto-Manning, 2014, p.

201), for it is their lived experiences that are key to this study.

Humanizing Focus Groups

I learned a lot and was able to see how others view the evaluation process. Some of my ideas/feelings were validated and echoed in my peers, and other times I was able to see things from different viewpoints, which is equally beneficial.
(Focus Group Participant reflection, July 2018)

I began my study of teachers' perceptions of high-stakes evaluations in the spring and

summer of 2016, asking teachers to complete an online survey to gain general insights and then

interviewing some of them to gain more in-depth insights. As I analyzed and wrote up this data, I

began to feel that my work was unfinished—there was more to the story I was attempting to tell

and more that I owed my participants. This sense of unfinished business found a home in my

doctoral humanizing research class. Through this class and the course instructor, Dr. Django

Paris, persistently compelling me to broaden my view regarding research, I began more closely

exploring my responsibilities as a researcher and the ways I could make my research more

humanizing for my participants. This led me to add a third phase of data collection to my research: focus groups informed by the praxis of culture circles.

For this third phase, I invited the teachers I had previously interviewed because I had drawn largely from their interviews in drafting my initial findings. In an effort to engage in critical thinking with these participants, I facilitated what traditionally might be called focus groups in which I embedded elements of culture circles. In the following sections, I describe focus groups and culture circles and then share how I enacted a combination of the two.

Focus Groups

Typically, a focus group is a discussion group, facilitated by a moderator, consisting of approximately six to twelve individuals who discuss and answer questions about a specific topic from their perspectives and personal experiences (Byers & Wilcox, 1991; Cyr, 2016; Powell & Single, 1996). Originating from the "focused interview" in sociology (Byers & Wilcox, 1991; Merton, Fiske, & Kendall, 1956), focus groups were mainly used in market research before becoming widespread in social and behavioral science research in the 1990s and 2000s (Byers & Wilcox, 1991; Cyr, 2016; Sagoe, 2012).

The purposes of using focus groups vary, but a primary goal is to "generate conversations that uncover individual opinions regarding a particular issue" and the reasons behind these opinions (Cyr, 2016, p. 233). Among other uses, focus groups can be utilized for triangulating data, pretesting surveys and other measurements, and exploring ideas (Cyr, 2016). Reflecting the multiple functions of focus groups, Creswell and Miller (2000) suggest that a focus group can be a form of member checking where research participants review and comment on the themes and accuracy of the findings (the researcher's interpretations of the data). In turn, the researcher can "incorporate participants' comments into the final narrative" (p. 127). Used in this way, focus

groups add credibility to a study and provide additional data for consideration (Creswell & Miller, 2000). Indeed, focus groups can be an excellent source of data in and of themselves because they afford discussions that are not possible with surveys or individual interviews (Byers & Wilcox, 1991; Cyr, 2016; Fife, 2005; Sagoe, 2012).

While I indeed gained additional data on the topic of my overarching study on teacher evaluations from the focus groups and used this data in conjunction with survey and interview data, these were not the only nor primary purposes for this third phase of my research. Because I believe that teachers' voices are often not elicited (and the teachers I interviewed seemingly agreed) and high-stakes evaluations are dehumanizing, I wanted to provide an opportunity for teachers to come together to discuss an issue that is affecting all of them, where they could share their insights and potentially find solidarity and problem solve. Thus, I drew upon the praxis of culture circles to create a more humanizing form of focus groups.

Culture Circles

Culture circles originate in the dialogic teaching practices of Paulo Freire, an educator, philosopher, and advocate for critical pedagogy, who wanted his students to learn to read both the written word and the world (Souto-Manning, 2007). Accordingly, culture circles are based on two tenets: the political nature of education and dialogue in the process of educating (Souto-Manning, 2010). A culture circle is "a group of individuals involved in learning… [and] in the political analysis of their immediate reality and national interests" (Giroux, 1985, p. viii as cited by Souto-Manning, 2010, p. 207). Culture circles start from issues of participants' everyday lives, where themes from prior experiences and community concerns serve as starting points to problem posing, and dialogue facilitates multiple perspectives (Souto-Manning, 2010). The goal of culture circles is problem solving that leads to action (Souto-Manning, 2010).

Culture circles honor participants' knowledge and experiences and have been used in various educational contexts, including with practicing teachers (Souto-Manning 2010). Furthermore, culture circles provide openings for storying (Kinloch & San Pedro, 2014) through which "narrators commence questioning their realities and problem solve" (Souto-Manning, 2014, p. 206). Considering the affordances of culture circles, it seemed particularly appropriate to draw on this method in my study as I sought to provide an opportunity for teachers to collectively read and question the political context of teaching and problem solve issues related to their realities. Certainly, I also sought to listen and learn from teachers about the influences of a neoliberal policy that is part of a larger institutional discourse.

Humanizing Focus Groups

For my culture circle-informed focus groups, I organized the teachers who agreed to participate into three focus groups of three to four teachers each. I purposefully limited the size of each group to allow for more individual voices to be heard. Additionally, I formed the groups, when possible, to include a representative from each of the three participating schools because I thought various perspectives could prove thought-provoking for the participants and myself. I also attempted to pair each teacher with at least one other teacher with whom they had something in common, such as subject area, grade level, or years of experience, as a way for the teachers to be able to relate to each other beyond the topic of evaluation. Because I thought it possible that some teachers may not feel comfortable participating in a focus group with a colleague from their school, I also offered the option of meeting in a group with only teachers from the other schools.

Having established relationships with my participants and wanting to create a more humanizing experience for them, I facilitated the focus groups in my home and provided lunch

and snacks. Food somehow often seems to make conversation easier, and my home provided a private and, seemingly relaxed, atmosphere. I conducted the focus groups in mid to late June, after the teachers had completed their school years, but before their summer breaks were in full swing and when their evaluations would be fresh in the teachers' minds because they get their evaluation results at the end of the school year.

I scheduled two 2-hour sessions with each small group around the teachers' preferences, including days of the week and time of day. However, not every teacher was able to attend their assigned focus group due to scheduling conflicts and health issues that arose (a reality of doing research with people). In the end, I conducted five focus groups with two to four teachers in each with the majority of the teachers participating in two sessions each. Notably, even though I scheduled two sessions with each small group, I initially thought one two-hour session for each small group would be ample time. However, the teachers had a lot to say about high-stakes evaluations and the second session was needed. For two of the groups, we actually could have used a third session, but I wanted to be respectful of the time commitment I had originally established.

At the beginning of each focus group, I reminded the teachers that I was seeking their honest thoughts, that there were no right or wrong answers, and that what they shared would be kept confidential. To elicit the teachers' insights on the influences of high-stakes teacher evaluations, I had prepared semi-structured, open-ended questions (see Appendix A), which I provided to the teachers prior to the focus groups for transparency purposes and to attempt to alleviate any apprehension the teachers might feel. I planned to, and indeed did, use these questions sparingly as I wanted to facilitate in a way that encouraged the teachers to lead the discussion, speak freely, and ask their own questions, which they did. During the focus groups, I

also provided the teachers an opportunity to review and comment upon both the raw data and the findings I had drafted from the first two phases of my data collection. These text-based think alouds served both as openings for discussion and as a form of member checking (Creswell & Miller, 2000), which was particularly important to me because I wanted to ensure I was, as accurately as possible, representing what the teachers previously had shared in their survey responses and interviews. The focus group conversations generally consisted of free-flowing dialogue with the teachers building off of each other's ideas.

As with the interviews I conducted, I did not position myself as an objective observer or neutral researcher but rather engaged in conversation with the teachers. I worked to build and maintain relationships of dignity and care by refraining from being the "neutral" researcher and resisting "the notion that sharing about ourselves during interviews attains less genuine and valid responses" (Paris, 2011, p. 142). As Paris (2011) suggests, in many contexts it is the sharing of ourselves that elicits participants' truest thoughts and feelings. Thus, while still being mindful of my purpose in centering the teachers' voices, I answered the questions the teachers posed to me and shared something of myself—my experiences as a teacher and an administrator—which indeed seemingly encouraged more genuine and valid responses. I say this based upon what the teachers shared with me, which was often raw and unapologetic. In the sharing of their stories, these teachers were trusting me.[7]

Engaging in a Humanizing Focus Group Amidst a Neoliberal Policy

As the following vignettes from the culture-circle informed focus groups and responses to a follow-up survey (see Appendix B) reveal, the teachers partook in much more than the communicating of opinions that occurs in traditional focus groups. Rather, because the focus groups provided a safe space, they dialogued, sharing personal experiences from their everyday

---

[7] I continue to question the trust they have placed in me as I reflect on my responsibilities at the end of this piece.

lives. Through this sharing, they engaged in storying (Kinloch & San Pedro, 2014) where their

stories became "woven into the stories of others, …construct[ing] a new reality based on a set of

relationships" (San Pedro, Carlos, & Mburu, 2017, p. 670). While most of the teachers had been

strangers prior to the focus groups, through dialogue, they found solidarity and became

colleagues. The teachers gained insights and built on each other's ideas, giving each other

(much-needed) validation and support. Together, they read and questioned the political context

surrounding high-stakes evaluations and teaching in a neoliberal era more broadly and, as a

result, felt a sense of empowerment.

A Safe Space

Myra:	I had a kindergartener with a social story he had to read every day about how putting his hands in his pants made people uncomfortable...
Erin:	Yeah… They don't think about that. They don't think about that kind of stuff...We had a 6th grader pee on the playground this year. Pulled down his pants and peed on the playground.
Jordan:	That doesn't surprise me…
Erin:	But I've got to grow that kid a year... I'm supposed to grow him a year. But they don't think about that when they're, they can't. They cannot think about that stuff when they're building these evaluation tools. They can't.
Myra:	No!
Erin:	And they can't think about the kids' home lives… It can't be.

Storying the unrealistic, and unreasonable, expectations of the current evaluation system

with very candid examples, Myra, Erin and Jordan's conversation reflects what the teachers

repeatedly stated in their survey responses: the focus groups were "a safe space to respond and

talk truthfully." Thus, the teachers were willing to be vulnerable and found it "liberating to

discuss openly with no fear of judgement or repercussion." This sense of safety encouraged the

teachers to share thoughts and feelings they would not share in other places, as encapsulated in

one teacher's comment: "I spoke freely and knew it was confidential… I would not have shared

all of what I did if it was in front of my administrator, or even colleagues." The focus groups

provided a safe space for the teachers to share personal experiences and frank thoughts.

Affirmation and Solidarity

> Ainsley: Well and that's what's frustrating too about when we have… evaluations linked
> to M-STEP. There's so many students like, "Well, dad hit mom today. I'm
> pissed off." And that's just...
> Sandy: Those kids… have so many outside factors.
> Ainsley: Yes!
> Sandy: They do not care about that test.
> Ainsley: No, they don't.
> Sandy: That test does nothing for them.
> Ainsley: And it's just like you get one moment and let's hope it's a good day. You can't...
> I don't know. It's so frustrating.

In addition to a safe space to share their frustrations, the focus groups provided the

teachers with a sense of affirmation and solidarity. For example, as Ainsley talked about the

Michigan Student Test of Educational Progress (M-STEP) and how teachers' evaluations were

tied to students' test scores, the other teachers in the group nodded vigorously, concurring with

her story. Sandy specifically validated Ainsley's experiences, agreeing and then repeating

Ainsley's statements in her own words. Through the focus group discussions, the teachers came

to realize that they had many shared experiences and, thus, were "not alone". Their comments

reveal some of the benefits of identifying with other teachers. Identifying evaluations as a

"systemic problem," one teacher reflected, "The focus groups were beneficial for affirmations

that we are all in this together. Many issues are not 'just me' or 'just my school', but issues

across the board." Another teacher shared, "I was less stressed and felt less alone in the world of

education. I felt like I was on the right path." For another teacher, the shared conversations

"reminded me what an important job we have." The opportunity to share their thoughts and

experiences with colleagues validated their stance as knowledgeable individuals and affirmed the

importance of the collective work they do as teachers.

Opportunities for Critical Examination

     Becca: This year I haven't gotten as much feedback and so anything that I've done was like language from the rubric. Like one example is 'Students are creating a rubric to assess their own work.' Ok, personally, I don't think that the amount of time you could spend on students creating a quality rubric is worth it.

     Christa: Right. For sure not.

     Becca: But you want me to do that? The students are going to make a rubric!

     Christa: I can check that box! Yeah. But that did not improve your instruction, right?

     Becca: Do I think that improved my [instruction]?

     Christa: Don't you feel like that?

     Becca: No. I don't feel that improved my instruction.

     Christa: But you're going to do that.

     Becca: But, that's what a highly effective teacher's doing!

     Christa: And so you're going to do that for one lesson.

     Becca: Yep.

     Christa: And go back, like you said, to what really works.

     Becca: Yep.

In questioning both the evaluation criteria and the effectiveness of their evaluations to improve their instruction, Becca and Christa's conversation reveals that the focus groups were an opportunity for the teachers to critically examine the topic of high-stakes evaluations. Echoing the thoughts of others, one teacher reflected: "Participating in the focus group gave me the opportunity to think about the evaluation practice differently than I have before." Listening to each other's perspectives and processing these perspectives with colleagues provided new insights that encouraged the teachers to think more deeply about, and question, their current evaluation system.

Empowerment

     Myra: Sometimes it's disillusioning. It is... to be in education, right?

     Heather: And it's so vilified right now with the government.

     Jordan: Yeah.

     Heather: Who signs up for it?

     Erin: But that's why I think some people don't want to encourage their children or their nieces or whatever to go into education because we do see the inside, the dark, the dark side of it. So dramatic. But like...

     Myra: But it's true!

Erin: But you see the injustices… It's sad. And as much as you know you're proud
            of [your daughter] and you want her to go into [teaching] and you know she's
            going to be great, you know she's going to go through those same struggles.
Myra: Yes.
Erin: And who knows in six years or four years when she gets out, what is it going
            to be like then? You know? It's kind of that scary.
Jordan: Well and I think there's like such a disconnect… Even like the people that are
            not in politics, we just [are] looked down on. Like… "You're just a teacher."
            That like drives me nuts!
Myra: I hate that…
Jordan: It's like you're not educated enough so you settled at teaching... It's so
            offensive, and I feel like there's just this disconnect… [And] we're not well
            equipped… to be there politically for our jobs. But I think that that needs to
            change because, if you look at the people that are the policy makers and the
            ones that are making the decisions, maybe a handful of them are actually
            educators.
Myra: A handful.

The ultimate goal of culture circles is problem solving that leads to action (Souto-

Manning, 2010). In this final vignette, the teachers identify the problem Sandy and Christa

alluded to in the opening vignette: teachers are being vilified by people who do not understand

what they do. Importantly, within this conversation, Jordan provides a solution. She contends

that teachers need to become equipped to be politically active. The teachers referenced such

problem solving in their survey responses and that they were empowered by the focus group

discussions. As shared by one teacher, "It was encouraging to hear from others and to come up

with viable solutions to current problems with the evaluation system." Reflecting on how the

learning that occurred in the focus group empowered her, one teacher shared, "Participating in

the focus group made me more comfortable and equipped to discuss these issues with other

teachers. I'm now more likely to engage in discussions about evaluations and believe I have

valuable input." Yet another teacher noted steps she was taking as a result of the focus group

conversations: "[It] encouraged me to learn more about educational policy and become more

equipped to defend myself and my students and to be involved whenever I can when it comes to

making decisions for the field." In other words, as stated by one participant, it made the teachers "want to be more of an advocate for the teaching profession!" In addition to, and because of, new insights, affirmation, and solidarity, the teachers found paths forward amidst a dehumanizing policy.

Because I identified the topic, asked the teachers to respond to my writing, and facilitated the discussion, my focus groups were not culture circles in their truest form. However, like culture circles, they did provide an opportunity for storying and for teachers to collectively read and question the political context of teaching, find solidarity, and problem solve issues related to their realities. Most importantly, the dialogue within the focus groups empowered the teachers to take action. The vignettes from, and teachers' reflections on, the focus groups allude to these aspects of culture circles and reveal attributes of a humanizing experience.

My Responsibilities: Part Two (and Three and Four and…)

I like knowing that something will come from the discussions that we have had.
That our voices will be heard.
(Focus Group Participant reflection, July 2018)

I enjoyed discussions with other like-minded individuals whom are so passionate about what they do and, in their participation, desire to have their voices heard just as I do.
(Focus Group Participant reflection, July 2018)

One has agency to resist and rebut dehumanizing forces, to reassert one's humanity,
and to play a part in work that humanizes others.
Blackburn, 2014

In light of her efforts to engage in humanizing ethnographic research, Mangual Figueroa (2014) argues that accounting for how a researcher departs from their research is just as important as accounting for how they entered. Furthermore, she suggests accounting for departure includes consideration of responsibilities to participants and the fulfillment of those responsibilities—both from the researcher's perspective and the participants' perspectives. While

mine was not ethnographic research, I recognize the importance of considering my responsibilities to my participants and how I may have, may have not, and may still fulfill my commitments to them as part of humanizing research.

I entered into this work with care and concern for my participants and a sense of responsibility to teachers, to sharing their words, to making a difference in whatever way I could, which, at the time, seemed limited to articles and presentations. However, I had not considered what the teachers might think my responsibilities were. I had not thought about who I was to them or who I was for them (Hill, 2009). Through my interactions with the participants, I came to realize that my positionalities—and, thus, my responsibilities—were informed not only by the roles I had assigned myself, but also by the ones my participants attributed to me. Importantly, this caused me to consider how my participants positioned me—and how I may have positioned myself—as someone who can do something about teacher evaluations.

In my initial meetings with the teachers, I shared that the intent of my study was to gather teacher perspectives on high-stakes evaluations because I thought (and still think) their voices were sorely lacking in the research about education policies and in the policymaking itself. Hearing this, it would be fair for participants to assume that their participation in my research would result in their voices being heard. Certainly, if I was going to ask teachers for their insights, then I was going to share them with a larger audience, right? It is what I believed to be true and what I think the teachers believed to be true as well.

Reflecting on what my participants shared with me, I realized that I also represent(ed) hope to some of them. Yes, they appreciated someone listening to them and caring about their situations, but, more importantly, they wanted me, and presumably others, to know how they were being negatively affected by a policy that is counterintuitive to them, and to do something

to address their concerns. Indeed, some of the teachers explicitly asked me to do something, anything. They believed, and hoped, I could do something.

But, what is that something? What is even possible for me to do? And will that something make a difference to the teachers? I believe facilitating humanizing focus groups that honored the participants' knowledge and experiences was one way I may have begun to address my responsibilities. Through these focus groups, the teachers had the opportunity to share their stories and found affirmation and empowerment. Importantly, they found their own ways 'to do something.'

As the teachers' quotations above articulate, having their voices heard by others outside of the focus groups was important and expected. Relatedly, the work of humanizing researchers broadly suggests that humanizing research does not culminate at the end of data collection or analysis but continues through and beyond the writing of manuscripts and development of presentations. Scholars who enact humanizing methodologies write about these experiences and their participants in humanizing ways (e.g. see Paris & Winn, 2014). They make rhetorical and philosophical moves that are humanizing, including the stances they take, the words they choose, and the perspectives they share. They reveal the complexity of individuals and their communities and provide insights that are humanizing and challenge the dominant discourse. Thus, another way I have attempted to address my responsibilities to my participants is to not only write about what I learned from them, but to do so in ways that honor their voices and reflect the complexity of their experiences.

In my attempts to conduct humanizing research and attend to my responsibilities in the eyes of my participants, I also have contemplated whether I was a "worthy witness," someone who made my participants feel valued and was seen by my participants as more than a researcher

gathering data (Winn & Ubiles, 2011). I think my enactment of the focus groups lends itself to worthy witnessing and, yet, I recognize there is still more work for me to do in this regard. For example, I am exploring ways that my research (i.e. the teachers' voices) can reach audiences outside of academia, specifically those who directly influence evaluation policy and implementation, such as state politicians and administrators. It also is important to consider ways that I can provide opportunities for my participants' continued involvement, particularly because several of the teachers expressed a desire for this. In this regard, I have shared drafts of the manuscripts I have written thus far for their review and invited them to attend, and participate in, my presentations of this work.

Lastly, as Blackburn (2014) suggests in the quotation at the beginning of this section, I have the "agency to resist and rebut dehumanizing forces" in my privileged position as a researcher, and, as such, I believe it is incumbent for me to do so. By seeking out and sharing teachers' perspectives in humanizing ways, I can resist and rebut the neoliberal narrative. This resisting and rebutting is necessary to reassert my own humanity and the humanity of the teachers in my study who continue to experience the damaging effects of neoliberal policies. It can give back to the teachers (and myself) some of the dignity that neoliberal policies, such as high-stakes evaluations, have taken away from us (as explicated by Sandy and Christa in the opening vignette).

Departing from the Work

I do not believe that I will ever truly depart from this work. I am not sure that someone who engages in humanizing research ever can. My participants became friends. Their stories reverberate in my heart and mind as I continue to wrestle with whether I have fulfilled my

responsibilities to them. And perhaps I am not meant to depart from this work, but rather engage in it in different ways.

I think that Blackburn's quotation above is also suggesting that such agency to humanize extends beyond a researcher and a particular study. Her words challenge me to think beyond myself and my study and consider ways I can "play a part in work that humanizes others." As a scholar and teacher educator, I believe that, with an informed and concerted effort, the possibilities of such work are endless. As one such effort toward meeting this challenge, I offer this paper to other researchers who will be conducting research with teachers amidst neoliberal policies. Through my description of the enactment of culture circle-informed focus groups, the vignettes from the focus groups, the reflections from the teachers, and my own reflections, I aim to identify a "path forward" for utilizing humanizing research amidst a dehumanizing policy and rationale that will encourage other researchers to take up research in humanizing ways.

In particular, I encourage researchers to examine their positionalities within their research both from their own perspectives as well as the perspectives of their participants and to think carefully about whom the research is for and what it represents. I also suggest making the effort to build relationships of dignity and care and creating safe spaces for teachers to engage in dialogue where critical thinking and storying can occur. Throughout the research process and particularly when departing from the work, researchers should consider their responsibilities to their participants and the ways in which they are fulfilling, or will fulfill, those responsibilities. It is also important for researchers to think critically about the ways they write about teachers amidst neoliberal policies so as to reflect the complexities of teachers' experiences and counteract the neoliberal narrative. Lastly, if we are to create humanizing experiences, it is

imperative to keep in mind that, not only is research regarding the effects of neoliberal

educational policies critical, so, too, are the ways in which this research is conducted.

APPENDICES

APPENDIX A:
Focus Group Questions

1. What are your thoughts as you read the findings/excerpts regarding collaboration/teacher well-being/professionalism? Do you agree? Disagree?

2. What, if anything, has changed regarding your evaluation system/tool/process/etc. over the last two years?

3. How have you made/do you make sense of the current evaluation system? Your rating?

4. What dilemmas, if any, have you faced as a result of/in relation to your evaluation/the evaluation system?

5. How have high-stakes evaluations influenced you personally? Your well-being? Your sense of professionalism?

6. How have high-stakes evaluations influenced you professionally? How have they influenced your work environment? Your work with colleagues? Your instruction?

7. What else is important for me, administrators, policymakers, etc. to know about the influences of high-stakes evaluations on collaboration/well-being/professionalism/ instruction?

8. How could the evaluation system at your school be improved? What actions do you think teachers could take to improve the current evaluation system?

9. What role does politics play in teacher evaluation?

10. In what ways, if any, does the current teacher evaluation system reflect the state of education in the United States?

Focus Group Survey

1. How would you describe the focus group experience to others? What is it like to participate in a focus group?

2. How, if at all, did participating in the focus groups influence you professionally/your sense of professionalism?

3. How, if at all, did participating in the focus groups influence you personally?

4. What benefits, if any, do you feel you gained from participating in the focus groups?

5. If I were to conduct focus groups in the future, what should I do the same? What should I change?

6. Would you encourage others to participate in focus groups? Why or why not?

7. What benefits, if any, do you feel you gained from participating in this study overall?

8. Would you encourage other teachers to participate in research studies? Why or why not?

9. What, if anything, did you learn from participating in this study?

10. What else would you like me to know about your experience participating in this study?

11. What else, if anything, would you like me to know about your experiences with/thoughts about teacher evaluation? (This could include something you thought, but didn't share during our focus groups and/or something you've thought about since we met.)

12. How, if at all, might you like to continue with this work?

REFERENCES

REFERENCES

Apple, M. (2006). *Educating the "right" way: Markets, standards, God, and inequality*. New York, NY: Routledge.

Blackburn, M. V. (2014). Humanizing research with LGBTQ youth through dialogic communication, consciousness raising, and action. In D. Paris & M. T. Winn (Eds.), *Humanizing research: Decolonizing qualitative inquiry with youth and communities* (pp. 43-56). Thousand Oaks, CA: SAGE Publications.

Byers, P. & Wilcox, J. (1991). Focus groups: A qualitative opportunity for researchers. *The Journal of Business Communication, 28*(1), 63-78.

Creswell, J. & Miller, D. (2000). Determining validity in qualitative inquiry. *Theory Into Practice, 39*(3), 124-130.

Cyr, J. (2016). The pitfalls and promise of focus groups as a data collection method. *Sociological Methods & Research, 45*(2), 231-259.

Dunn, A. H. (in press). Leaving a profession after it's left you: Teachers' public resignation letters as resistance amidst neoliberalism. *Teachers College Record*.

Fife, E. (2005). A focus group activity for the research methods class. *Communication Teacher, 19*(1), 9-12.

Freire, P. (2000). *Pedagogy of the oppressed* (M. B. Ramos, Trans.) (30th anniversary edition). New York, NY: Continuum.

Green, K. (2014). Doing double dutch methodology: Playing with the practice of participant observer. In D. Paris & M. T. Winn (Eds.), *Humanizing research: Decolonizing qualitative inquiry with youth and communities* (pp. 147-160). Thousand Oaks, CA: SAGE Publications.

Guenther, A. (2019). *"How is this making my instruction better at all?": Teachers' perceptions of high-stakes evaluation and its influence on their practice and identity.* Manuscript in preparation.

Hill, M. L. (2009). *Beats, rhymes and classroom life: Hip-hop pedagogy and the politics of identity*. New York, NY: Teachers College Press.

hooks, b. (1994). *Teaching to transgress: Education as the practice of freedom.* New York, NY: Routledge.

Hursh, D. (2004). Undermining democratic education in the USA: the consequences of global capitalism and neo-liberal policies for education policies at the local, state and

federal levels. *Policy Futures in Education, 2*(3), 607-620.

Ingersoll, R., Merrill, L., & Stuckey, D. (2014). *Seven trends: The transformation of the teaching force, updated April 2014*. (CPRE Report #RR-80). Consortium for Policy Research in Education, University of Pennsylvania.

Irizzary, J. & Brown, T. (2014). Humanizing research in dehumanizing spaces: The challenges and opportunities of conducting Participatory Action Research with youth in schools. In D. Paris & M. T. Winn (Eds.), *Humanizing research: Decolonizing qualitative inquiry with youth and communities* (pp. 63-80). Thousand Oaks, CA: SAGE Publications.

Kinloch, V., & San Pedro, T. (2014). The space between listening and storying: Foundations for projects in humanization. In D. Paris & M. T. Winn (Eds.), *Humanizing research: Decolonizing qualitative inquiry with youth and communities* (pp. 21-42). Thousand Oaks, CA: SAGE Publications.

Mangual Figueroa, A. (2014). La carta de responsabilidad: The problem of departure. In D. Paris & M. T. Winn (Eds.), *Humanizing research: Decolonizing qualitative inquiry with youth and communities* (pp. 129-146). Thousand Oaks, CA: SAGE Publications.

Merton, R. & Kendall, P.L. (1946). The focused interview. *American Journal of Sociology, 51*(6), 541-557.

Miron, G., & Gulosino, C. (2013). *Profiles of for-profit and nonprofit education management organizations: Fourteenth Edition—2011-2012*. National Education Policy Center, University of Colorado. Retrieved from http://nepc.colorado.edu/publication/EMO-profiles-11-12

Paris, D. (2011). 'A friend who understand fully': Notes on humanizing research in a multiethnic youth community. *International Journal of Qualitative Studies in Education*, *24*(2), 137-149.

Paris, D., & Winn, M. T. (2014). *Humanizing research: Decolonizing qualitative inquiry with youth and communities*. Thousand Oaks, CA: SAGE Publications.

Powell, R. A., & Single, H. M. (1996). Focus groups. *International Journal of Quality in Health Care, 5*(5), 499-504.

Robertson, S. (2007). Remaking the world: Neoliberalism and the transformation of education and teachers' labor. In M. Compton & L. Weiner (Eds.), *The global assault on teaching, teachers and their unions* (pp. 11-36). New York, NY: Palgrave.

Sagoe, D. (2012). Precincts and prospects in the use of focus groups in social and behavioral science research. *The Qualitative Report, 17*(29), 1-16.

San Pedro, T., Carlos, E., & Mburu, J. (2017). Critical listening and storying: Fostering respect for difference and action within and beyond a Native American literature classroom. *Urban Education*, *52*(5), 667-693.

Souto-Manning, M. (2007). Education for democracy: The text and context of Freirean culture circles in Brazil. In B. Levinson & D. Stevick (Eds.), *Reimagining civic education: How diverse societies form democratic citizens* (pp. 121-146). Lanham, MD: Rowman & Littlefield.

Souto-Manning, M. (2010). *Freire, teaching, and learning: Culture circles across contexts*. New York, NY: Peter Lang.

Souto-Manning, M. (2014). Critical for whom? Theoretical and methodological dilemmas in critical approaches to language research. In D. Paris & M. T. Winn (Eds.), *Humanizing research: Decolonizing qualitative inquiry with youth and communities* (pp. 201-222). Thousand Oaks, CA: SAGE Publications.

Tuck, E., & Yang, K. W. (2014). R-words: Refusing research. In D. Paris & M. T. Winn (Eds.), *Humanizing research: Decolonizing qualitative inquiry with youth and communities* (pp. 223-248). Thousand Oaks, CA: SAGE Publications.

U.S. Department of Education, Office of Planning, Evaluation and Policy Development, Policy and Program Studies Service. (2016). *The state of racial diversity in the educator workforce.* Retrieved from https://www2.ed.gov/rschstat/eval/highered/racial-diversity/state-racial-diversity-workforce.pdf

Winn, M. T., & Ubiles, J. R. (2011) Worthy witnessing. In A.F. Ball & C.A. Tyson (Eds.), *Studying diversity in teacher education* (pp.293-306). Lanham, MD: Rowman & Littlefield.