University of Cincinnati

Date: 2/23/2018

I, John E. Stegall Jr., hereby submit this original work as part of the requirements for the degree of Doctor of Education in Urban Educational Leadership.

It is entitled:

Qualitative Analysis of Teacher Evaluation from the Perspectives of Teachers within a Public School District

Student's name: **John E. Stegall Jr.**

This work and its defense approved by:

Committee chair: Miriam Raider-Roth, Ed.D.

Committee member: Mary Brydon-Miller, Ph.D.

Committee member: William Gary Pack, Ed.D.

Committee member: Constance Kendall Theado, Ph.D.

Committee member: Vicki Plano Clark, Ph.D.

30694

Qualitative Analysis of Teacher Evaluation from the Perspectives of Teachers within a Public School District

A dissertation submitted to the

Graduate School

of the University of Cincinnati

in partial fulfillment of the

requirements for the degree of

Doctor of Education

in the Urban Educational Leadership Program

of the College of Education, Criminal Justice, and Human Services

November 2018

by

John E. Stegall Jr.

B.A., University of Akron, 2000
M.Ed., University of Cincinnati, 2011

Committee Members:
Dr. Miriam Raider-Roth, Chair
Dr. Mary Brydon-Miller
Dr. Vicki Plano Clark
Dr. Gary Pack
Dr. Connie Kendall Theado

Abstract

This qualitative interview, group level assessment methodology-based, and self-reflection study examined the phenomenon of teacher evaluation from the perspectives of teachers within a local school district serving several suburban communities situated in the metropolitan area of Cincinnati, Ohio, as well as from the professional perspective of the researcher. The rationale for this research study came about as a response to the emergent trend in the United States of implementing of new professional performance evaluation systems within K-12 public education focused on teachers. Absent from the development of these new professional performance evaluation systems was the presence of teacher voice within their development. Research study findings consistent across a review of the literature, thematic analysis of two rounds of interviews of eight teachers, followed by a group level assessment activity with two teachers, and a self-reflection by the author of this research study suggested central themes associated with teacher voice consisting of Teacher Perceptions of Evaluation, Evaluation Validity and Reliability, and The Role of Mentoring in the Context of Evaluation. The implications of these themes suggest several potential areas for additional inquiry and action on local as well as state policy fronts. These include providing additional balance and support for the principal as evaluator and instructional leader, improving alignment between evaluators' professional backgrounds and the teachers they evaluate, expanding mentoring and related support for teachers, promoting greater awareness of inequities within professional relationships and practice, and focusing political action and advocacy to give teachers voice within current and future professional performance evaluation policy decisions.

Keywords: teacher evaluation, interviews, group level assessment, self-reflection, qualitative

Acknowledgements

What began in 2009 as a sudden and unexpected opportunity for educational advancement and leadership that I had long anticipated, but nearly gave up on pursuing, not only happened, but took on the form of an extraordinary journey that has culminated in my completing a doctoral program in Urban Educational Leadership. Along the way, I have met and engaged with many people whose contributions to my personal and professional growth both large and small have earned my respect and gratitude.

My list of acknowledgements begins with Jeff Forester, Jack Culp, and James "Skip" Thomas, veterans of educational leadership, all of whom encouraged and supported my initial candidacy as a Leadership Trainee and encouraged me to strike at the opportunity to earn my master's degree and also earn leadership experience while the iron was hot for it. In addition, Paul Wrobbel, Les Flynn, Mike Berry, James "Jim" Condron, Brent Dearbaugh, Tom Fletcher, Mark Stuart, Kevin Snarr, Dawn Psurny, Sarah Patterson, Sarah Hanka, Maria Boyarko, Garry Gohlke, John Tornes, Patrick Barrett, Scott Kern, Steve Sellers, Nick Wilson, Kristen Kask, and MaryAnn Latimer each in his or her own way served as a mentor, a guide, and an example of leadership in education for me, teaching me both through feedback and example.

Dr. Carri Schneider was the facilitator of my master's degree program in Educational Leadership at UC. She was always strong and firm in her instruction and feedback, but equally as supportive and encouraging of our efforts to learn and grow further through reflection and applied practice. It was Carri who actually encouraged me to make the leap into the Urban Educational Leadership doctoral program, citing my strong writing and critical thinking abilities. I owe her a unique level of gratitude.

I recognize and honor the contributions of the faculty members that welcomed me into Urban Educational Leadership at UC while challenging me and encouraging me to reach new heights, among them Dr. Ann Millacci, the director of the online Educational Leadership program who supported and endorsed my candidacy for admission, as well as Dr. Daniel Milz, of the master's level principal internship program. I further acknowledge Dr. Janet Decker, Dr. James Koschoreck, Dr. Shana Goldwyn, Dr. Nancy Evers, Dr. Roger Collins, Dr. Gregg Tracy, Dr. Rob Harper, Dr. Lionel Brown, and Cheryl Deardorff, each for their contributions to my development and advancement. I further acknowledge Dr. Chris Swoboda, Dr. Leigh Wang, and Dr. Connie Kendall Theado.

I reserve a particular debt of gratitude to Dr. Vicki Plano Clark, Dr. Miriam Raider-Roth, Dr. Gary Pack, and Dr. Connie Kendall Theado, for each agreeing to set aside time within their very active schedules to serve as members of my doctoral committee. Each generously provided me with their constructive suggestions, ideas and feedback at each and every critical step along the way from my initial program plan hearing to my final dissertation defense. I am truly grateful to them for their support and guidance.

Further, I am forever grateful to Dr. Mary Brydon-Miller for her willingness to take me on as one of her advisees and for providing me with her unwavering support and guidance. During a time of great uncertainty about the future of my program with regards to its administration, she ensured that I had a fighting chance to succeed in spite of the technical obstacles that emerged, holding fast to, while also at the same embodying, the higher ideals and model of leadership that originally brought me to UC.

My time spent in and around UC in Clifton yielded many friendships over the course of the three years I lived and worked in Cincinnati. I am forever grateful for the bonds and

connections I established with my fellow grad students, UC alumni, and kindred souls within our extended circle: Cindy Veraldo, Anita Williams, Helen Wang, Portia Watkins, Amy Crouse, Melissa Stuart, Peggy Shannon-Baker, Bertin Ondja'a, Samuel Burbanks, Jennifer Killham, Nikki Bibler Zaidi, Mark Kohan, Eric Dustman, Khaliah Sanders, Susan Tyler, Gail Headly, Janet Walton, Jenny Mast, Abishek Balsamy Kamaraj, Ayça Mazman, Sarah Hellmann, Alan Wight, Karla Gacasan, Doug Stevens and Skip Houck, Anne von Koschembahr, Mary Stroud, David Spatholt, Edwin and Jenny Loayza, and David Abad. I owe a particular debt of gratitude to Doug and Skip for coaching me through my preparation for the Group Level Assessment activity project, and Doug in particular for also helping to organize and facilitate it while I documented the responses of participants, as well as Karla and Stephanie for generously offering me room and board after I moved back to Columbus when I still needed to spend additional time in Cincinnati in order to collect the remainder of my data. I further owe eternal debts of gratitude to Cindy for being a sympathetic, supportive friend and mentor to me nearly every step of the way throughout my program, and to Anne for being both a dear friend and mutual confidante while she endured her own trials and rarer but rewarding triumphs within the medical college. Finally, my cohort colleagues, Sarah Ghee, Pam Theurer, Juanita Darden-Jones and Melissa Brown, have each and together become like four additional sisters I never realized I had until I met them during the day of our program orientation. We grew close during the course of our experiences and challenges both shared and individually endured. They are literally like family to me now, and I hope that our bonds and mutual support will only continue to grow and deepen with the passage of time.

Teachers come in many forms within the classroom of life. I am honored to have known several of them as good friends over the years, among them are Steve Yeargin, Tom Basham,

Mike Hixon, Sean and Jody McClain, Mike Codechini, Carrie and Adam Clark, David Thomas, George and Twyla Herring, Jinhee Cho, Glenn and Jin Young Curtis, and Chris and Jessica Hinsvark. I further acknowledge and honor those individuals who have mentored me from their unique positions in relation to my own: Jennifer Boyer, Bhopinder Parwana, Patricia "Pat" Paquette, Tom Beck, Rick DiClemente and Liza Jane Brown, Ron Stuecher, Ted Senf, Joey and Emma Calhoun, Gertrude Holt, Joyce and Henry Robinson, Leeola Yancey, Mr. Hill, Ms. Morton, Mr. Thomas Royce, and Karen Jones.

Along similar lines, I have drawn a unique form of inspiration from two kindred spirits I have known as friends as well as colleagues in service to humanity, Danielle M. Hagaman Hier and Laura Condos Merriman. Each in her own way has served to ignite and further inspire my passion to further pursue my path and goals in life with the same resiliency, tenacity, and vigor that yet spurs her ever onward along her own respective life course.

My list of gratitude includes many more good friends, honored professional colleagues, and influential acquaintances in either personal or professional spheres of my life. Among them are Andrea Schweitzer, Derek and Lisa Arnold, Julie Anthony, Tom Kemp, Ed Ackerman, Ray Petelin, Rachel Petelin, C. Ann White, Eric Viegel, Lori Roberts, Matt "Chill" Frasier, Shawn Starcher, Steve Garton, Tom and Jenny Arnold, Chris Diefendorf, Kelly Byrnes, Dom Comparetto, Dana Comparetto, Richard and Donna Noble, Yyvette Beach, Chris Boyer, Dreama and James Rice, Cindy Culver, Shunda Wright, Gwynne and William Roberts, Marianne Majewski, Lisa and Ryan Garber, Cole Ford, Andre Green, Zachary Graham and Mea Adams, James Stanford, Brian McGee, Heather and Mike Fidler, Katie Abbott, and Kari and Bill Warfield.

Also noted and honored are Kent Burgert, Lori Acheson, Erin Boron, Patty Whetsel, Patrick Shaughnessy, Charlie Shoenberger, Ed Breehl, Elizabeth Bailey, Eliott Giles, Mat Old, Heather Wilhelm, Jim Gray, Jay Bowman, Jennifer Cornell, Rebecca Rose, Sara Adduchio, Sara Kranson, Marieca Mattox, Rosemary Mullins, Dave Sirkoski, Keith Cole, Vernell Bristow, Deron Obey, Kristy Klare-Robinson, Jeff Grimmett, Max Lallathin, Tim Schmid, Jerrilyn Baumbusch, Linda Schurch, Chris Blose, Crystal Wasmund, Daren Thomas, Andy McGillicuddy, Sunny Schultheis, Erin Schultheis, Heather Townley, Diane King, Emma Baucher, Christine Armbrust, Bev Beutell, Angela Vickers, Ryan Kilpatrick, Cara Shebak, Melissa Fitzsimmons, Kelly Hanna, Kelly McCormick, Tiffany Hunter, Robb Townsend, Alison Cornelli, Jamie Campbell, Jennifer Franzel, Mindy Geschke, Caroline Adams, Brandy Cox, Terri Gossett, Sherry Newmann, Sarah Plasters, Erin Newmann, Amanda Magora, Denise Jones, Jackie Gray, Beth Rollyson, Suzi Zeinski, Brad Scott, Jeff Long, Adam Riggs, Andy Whitmer, Sean McCready, Ruth Miller, Amanda Johnson, Mike Cook, Ashley Hale, Melva Coyle, Danny Sams, Jill Dannemiller, Brittny Pierson, Cathy Blankenship, Rick Teeters, Nichole Reese, Hailey Dunn, Michael Tucker, Susan Ratliff, My-My Huynh, Bethany Immel, Carol Dimoff, Alex Subak, Brittany Dunn, Rob Schofield, Janice Boyd, Mark Palen, Judy Tripp, Justin Pyle, Jesse Thomas, Jeff Smith, Larry Snyder, Ken Kibbey, Brain McWhorter, Kim Woods, James “Jim” Hitt, Kim Petrik, Chris Hitchcock, Ron Ball, Brett Counts, Erica Coffman, Tom Dyer, Grant Kibbey, Stefanie Hunn, Devon Ryan, Ross Miller, David Shock, Kim Miller, Chris Stanley, Sara Szych, Sara Hayes, Deb Tracy, Susann Stanley, Christy Martinez, Ashley Cohen, Lisa Taggart, Adam Clark, Katie Self, Melissa Pitzer, Brent Newbold, Katherine Hetrick, Jason Gary, Conor Dempsey, January Almaguer, Melissa Sullivan, Amanda Charney, Elizabeth “Libby” Campbell, Anna Geiss, Tabitha Stitt, Phil Serina, Greg Stinger, Shawna Pope, Clayton “Chip” Bullett,

Christina Ritchey, Jennifer Patton, Natalie Blake, Peggy Martinez, Andrea Berg, Alaina Way, Deanne Anderson, Jennifer Jerrome, Rebecca “Becky” Farson, Rick Hunyady, Jeff Maglich, Kim Turner, Laurie Wikle, Sheri Dornhecker, Ann and Gale Barnes, Pete Walter, Jim Underwood, Tom Kaylor, Eddie Zoog, Sheree Harvey, Lawana Partlow, Krista Dobbins, Kyle McDowell, Jan Attanasio, Patrick Tingler, Mark Puttick, Sue Bradley, Kathy Martensen, Ryan Crawford, Marcia Speck, Donna Davies, Drew Palmer, Alecia Taylor, Cindy Burrington, Alisa Fisher, Deborah Romans, Jacque Zody, Jacque Konich, Jackie Matheney, Stephanie Goodrich, Jen Erickson, Billie Jo Catanzarite, Robin Hyde, Chris Hoover, James Jones, Mark Stewart, Vicki Bryan, Kenton “Kenny” Lee, Deliana Soto, Denise Sizemore, Ingrid Luft, Joleata Johnson, Michelle Brown, Sara Thomson, Aaron Moore, Niles Riddel, John Harrison, Vincent Tubuo, Mary DeLozier, Mary Barber, Aaron Stamper, Grace Chen, Jathiya Abdullah-Simmons, Ryan McClellan, Thomas Trang, Katherine Carr, Karen Saunders, Tera-Lee Stockdale, Lisa Brandt, Rebecca Vasilakis, Rhonda Tabbah, Christine Hecht, Jeff Cox, Katherine "Katie" Kimnach, Adam Weiss, Kelly Weinfurtner, Melissa Wengerd-Gertsch, Jenny Ramirez-Bello, Coral Bishop, Andrea Tippery, William Lehman, Cheryl Burden, Jackie Nash, and Nancy Sinning.

I offer special thanks to the following:

Coffee shops Muggswigz Coffee and Tea Co. of Canton, Ohio, Angels Falls Coffee and Nervous Dog Coffee Bar of Akron, Ohio, Artisan Coffee of Ellet, Ohio, Grandview Grind and Stauff’s of Grandview Heights, Ohio, Crimson Cup of Columbus, Ohio, Crimson Cup of Tallmadge, Ohio, Phoenix Coffee of Cleveland Heights, Ohio, Rohs’s, Sitwell’s, and Community Coffee, of Cincinnati, Ohio, and of course, Third Perk of Dayton, Ohio, all of whom provided me with the fuel to work through many long days and longer nights.

Stev Guyer, Stacie Boord, Julie Klein, Jimmy Mak, David Whitehouse, Gabriel Guyer, Tom Cardinal, Amy Lay, Katy Psenicka, Nikki Fagin, Brandon Anderson, Noelle Grandison, Leah Haviland, Stephanie Shull, Matthew Hahn, Nickey Winkelman, J.T. "Chops" Walker, Mary Randle, Jennifer Hahn, Andy Ankrom, Harley Wolfe, and the entire cast and crew of Shadowbox Live who helped me to remember how to laugh and de-stress when those switches in my mind lay buried beneath an endless string of data, and no clear way to reasonably sort it all out in the moment.

Dr. Carol Engler, an inspired and inspiring teacher, principal, professor, and lifelong learner, for recognizing, honoring, and placing such tremendous faith and hope within me at our very first meeting, the great potential that I could not yet see in myself at the time. I will forever be grateful to her for pointing me towards the path that I would ultimately walk to become the person that I am today and am still becoming. She will always represent to me the essence of what it means to live and to share your inspiration with others.

Faith Stegall, my former wife and enduring dear friend for her top-notch editing and proofreading skills, ceaseless moral support, offbeat humor and insights, and renewed friendship. I didn't understand until I met her just how much two people and their relationship with each other could go through as well as grow through. I am eternally grateful to her for supporting me through the triumphs and failures, crises and times of smoother sailing.

The faculty and staff of the school district primarily featured within this study for welcoming me into their district and contributing to a process that I hope will only generously and beneficially return in spades for them what it has given to me.

I am forever grateful to my parents, John Sr. and Cheryl Stegall for their unconditional love, encouragement, and support that have served to motivate me throughout my entire graduate

school journey and to never give in and never give up. I am equally as thankful to my sister and brother-in-law, Janet and Rodney Scott, for the inspiration and moral support they have provided to me through their own tenacious and ongoing educational pursuits. I further recognize my cousin, Tai Stevens, as a kindred spirit, who like Janet, Rodney, and I, relentlessly exhibits the drive and determination to expand and improve her professional skills and knowledge, and is all the better for doing so.

I acknowledge the support, love and inspiration I have received on my journey from my extended family, among them, my dearly loved late Aunt Rebecca "Becky" Reese whom I fondly remember for tying up my socks when I was little, my late Uncle Robert and my Aunt Phenon Reese, Uncle Richard and Charmaine Reese, Grandpa Sylven and Grandma Annie Dale Stegall, Uncle LB and Lucy Stegall, Uncle Excel and Aunt Liz Stegall, Uncle Wayne Stegall, cousins Richard, John, Meredith, Mike and Robert “Bobby” Reese III, Kevin, Steve, Tracy, Tiffany, Wanda, and Shalonda Stegall, and the many further extended branches of the Reese, Poole, and Stegall families.

Finally, I dedicate this work to the following:

To the spirit of resiliency and commitment to triumph and thriving even in the face of overwhelming odds, outside disdain, and general adversity that would lay flat most others, evident in my hometown of Akron, Ohio as well as throughout the cities of the Automotive Rust Belt region, Cleveland, Pittsburgh, Youngstown, Canton, Erie, Buffalo, Rochester, Toledo, Detroit and Flint.

To the ongoing improvement of the education profession from within, as achieved through the daily and ongoing practice of those who can only be the most knowledgeable and best informed about it, teachers and educators themselves.

To my niece, Sydney Isabelle Scott, whom I hope will never stop learning, never stop asking questions, and always persevere within her ambitions with the knowledge that education is the key to ultimate success and achievement. May you always carry the torch of its inspiration and share your light with others, and the entire world.

And finally, to my dearly beloved late Grandma, Clara Poole Reese, whose wit and warm words of wisdom were so much an influential part of my childhood and early adult years. Her spirit exists forevermore as a guiding beacon within my mind, heart, and soul.

To all those who supported me along the way in some form or another, it is my genuine hope that I can somehow or another eventually return the favor, in kind, in full, and much more besides. From the absolute bottom of my heart, thank you.

Table of Contents

Chapter 1

Introduction to the Study

Within this chapter, I provide an overview of the problem I intend to explore within this research study, and also a brief discussion of the gap in research that exists. Further, I provide the context in which the problem exists. In addition, I provide a purpose statement, as well as an explanation of intended audiences for this research study, and a definition of terms section. Furthermore, I discuss my conceptual framework and positionality as a researcher. Finally, I discuss the significance of this research study.

Statement of Problem and Research Gap

The problem that I have intended to address through my completion of this research study is accountability within teacher evaluation. Traditionally, school principals have been responsible for completing teacher evaluations (Brand et al., 2007). Until recently, the trend has been for evaluation models and procedures to reinforce this role, making principals the dominant arbiters of teachers' performance, which has included determining the consequences for failing to meet expectations, such as non-approval of tenure and dismissal (Brand et al., 2007). Structured frameworks such as Danielson's (2000) and Marzano's (2012) have provide a constructivist approach to teacher performance evaluation, but still primarily emphasize the role of the principal within the teacher evaluation process. During the recent period, however, value-added measures of teachers' instructional effect on student achievement have increasingly been incorporated into teachers' professional evaluations, supplementing or even complementing principal observations as a summative measure of teachers' performance (Guarino et al., 2014). There appears to have been limited research conducted examining the potential for teacher input within the performance evaluation process, however.

In simple terms, I wanted to study accountability within teacher evaluation from the perspectives of teachers. How better to understand the concept in the greatest amount of depth than from the very individuals who experience it foremost as a critical component of their employment? Further, the lack of previous research into teachers' perspectives on accountability within professional methods of evaluation suggests a blind spot within such methods and the policies driving them. Research into this gap potentially raises new questions with regards to how teacher evaluations are facilitated and may lead to changes in approach and policy that benefit the education field overall, while further supporting teachers' professional needs.

The Ohio context. The concern that led me to examine this issue was based upon my observation that K-12 public school teachers in Ohio lacked substantial input within the decision to develop and implement of a statewide teacher evaluation system that would measure and judge them on the quality of their instructional performance. This system, known as the Ohio Teacher Evaluation System (OTES), is a standardized, state-wide system of teacher evaluation that requires all Ohio public school districts to evaluate the overall performance of their teachers using a combination of performance ratings based upon the observations of evaluators and quantitative student growth measures based upon standardized test score results (Ohio Department of Education, 2016). The Ohio General Assembly and Governor John R. Kasich codified legislation in 2012 detailing the structure and procedures of OTES, which became effective statewide at the start of the 2013-14 school year. Several other states in the US have adopted similar systems of evaluation that also incorporate the use of student scores on standardized assessments to measure and summarily determine teacher performance (Darling-Hammond et al., 2012).

Of particular note to me with regards to the development of OTES was the lack of research citing valid and reliable use of quantitative data within performance evaluations, most notably student test scores. I found this research gap to be rather surprising, and it further suggested to me that Ohio's use of such data within OTES was occurring without any sound basis in available research. It also meant that public school teachers would be subject to an untested, and perhaps unreliable and invalid system of performance evaluation that would have bearing on their professional reputations as well as their continued employment and licensure within the state.

In my view, accountability teacher evaluation methods warrant considerable scrutiny. Performance evaluation systems should be as reliable as possible while also producing valid results during each instance of application. In fact, existing research suggests that quality teacher evaluations are not narrowly focused on one or two measures of teacher effectiveness, but rather are comprehensive in their approach, relying on multiple measures that consider a teacher's overall impact on student performance and development (Goe et al., 2008). A number of researchers have challenged the use of student test scores on standardized assessments as a valid and reliable indicator of teacher performance, citing inconsistencies within results and suggesting the presence of confounding variables for which many teacher performance evaluation systems do not properly account (Brown et al., 2010; Darling-Hammond et al., 2012; Hill et al. 2012; Huang et al., 2009; Rivkin et al., 2005, Sanders & Horn, 1994; Stronge et al., 2011). As such, state officials failed to ensure that a key performance assessment system is even capable of producing valid and reliable results. Of further concern are the potential legal ramifications for schools, districts, and the state and federal governments that either actively use, support, or have

endorsed this particular narrowly-focused approach to summative measurement of teacher performance (Baker et al., 2013).

In short, this research study is important because it examines a methodological approach to facilitating professional accountability for teachers that may or may not be producing valid and reliable results. The role of teachers within any such system derived from this approach warrants due consideration, given their status as subjects that can be affected by its results. Their direct input is what I seek in order to understand what would best serve teachers within the context of professional evaluation.

As I prepared to conduct this research study on accountability within evaluation from teachers' perspectives, I considered the possibility that many teachers—possibly even a majority—would like and even prefer the new teacher evaluation system, and are okay with it as-is. As such, I realized that I had to be open to receiving perspectives through my proposed research study that might challenge my concerns about the lack of teacher voice within the development and implementation of teacher evaluation systems that include quantitative assessment data. I also came to appreciate the need to be receptive to diverse opinions, including those that are perhaps conflicting and contradictory.

I decided to proceed with my inquiry into teacher evaluation from the perspective of teachers, setting aside for the time being my own opinion that teachers should be concerned about quantitative assessment measures within teacher evaluation—and OTES in particular within Ohio—as an untested and potentially invalid approach to performance measurement, as well as what these may mean for their careers. In order to make space for their voices, I understood that I, as the researcher, needed to step aside and allow them to speak with their voices and allow their opinions to be heard and known.

I must note that since I initially began to conduct my research study during the spring of 2014, the Ohio General Assembly and Governor Kasich elected to modify OTES for the 2014-2015 school year to include an additional, alternative evaluation framework along with alternative components of evaluation within this structure as potential options for districts, schools, evaluators and teachers to use. These options include teacher self-evaluation, peer review, student portfolios, student surveys, as well as local district discretion in developing and using a unique alternative component. Whereas the original version of the OTES evaluation framework featured teacher observations and student growth measures with each contributing 50-percent towards teachers' overall summative ratings, the newer, alternative framework reduces these components to 42.5-percent each, with the alternative measures accounting for the remaining summative rating balance of 15-percent.

This update and institution of an alternative framework did not immediately influence the planning or execution of my research study as my research study at this point in time was already within its advanced stages of facilitation. Further, these changes did not represent substantial, fundamental changes to teacher evaluation as mandated within the state of Ohio and the school district featured within this study, on a local level. I further note that in September 2015, the Ohio General Assembly and Governor Kasich elected to make additional modifications to the OTES alternative evaluation framework for the 2015-2016 school year. These modifications consisted of returning the weighting of teacher observations to 50-percent, while lowering student growth measures to 35-percent and keeping the additional optional measures possible at the remaining 15-percent of teachers' summative ratings.

Purpose Statement

The purpose of this research study is to gain understanding of the perspective of K-12

public school classroom teachers regarding teacher evaluation. By conducting this research study, I seek to answer the following central questions:

- What are teachers' perceptions of teacher evaluation?
- What are teachers' concerns regarding the validity and reliability of teacher evaluation?
- What are teachers' views regarding the role of mentoring in the context of evaluation?

These qualitative questions served as the guiding basis of my research inquiry within the development of an interview protocol that I used during interviews with teachers. Further, the findings that resulted from this interview protocol and process informed the development and implementation of a Group Level Assessment (GLA) activity to inform teachers as to items for advocacy and action within their district. Ultimately, I discuss my overall findings and their implications for the school district's teachers, as well as for teachers with respect to their voice and teacher evaluation.

Intended Audiences

My primary intended audience for this study is elected state government officials, particularly the General Assembly and Governor of Ohio that together passed OTES into law, and would also be respectively responsible for making any changes to it in terms of its legal definition and execution. I also anticipate that the Ohio Department of Education will have an interest in this study since they carry the task of overseeing the implementation and administration of OTES as the designated education administrative unit of the state executive branch, and thus may be concerned about the perceptions of the validity of OTES from an overarching, statewide administrative perspective. Further, I expect that school officials at the

district level, including superintendents and executive-level administrators such as directors of human resources and of curriculum, will also have an interest in the findings of this study based upon their roles and interests in managing personnel, perhaps with collective bargaining agreements and the legalities of making decisions about teacher retention in mind.

Ultimately, however, I expect that Ohio public school teachers themselves, particularly those whose schools and district feature within this research study, will have an interest in my findings, as it will give them insight into perspectives on teacher evaluation that they may or may not share with their professional colleagues. Further and most importantly, the findings within this research study may provide teachers, if they elect to pursue it, a staging point for advocacy and action in terms of collectively seeking to either reform or eliminate teacher evaluation as it currently functions within state and local education policy, according to the perspectives and ideas to which their colleagues have given voice.

Definition of Terms

Here I provide a brief overview of terminology used or referred to within this document. One or more of these terms may not be familiar to readers who have little or no practical experience as a professional serving within K-12 education in Ohio. Accordingly, this section is intended to serve as a point of reference.

Banking. The banking approach is a conceptualization of education in which teachers literally deposit facts and bits of knowledge within students' minds for them to regurgitate when prompted, much like retrieving a specified amount from one's bank account via a teller. It is a grossly simplified view of learning that denies the humanity and innate potential of students themselves, reducing them to mere vessels that a teacher must labor to fill up with knowledge, rather than to inspire to engage in learning (Freire, 2000).

Ethic of Care. A conceptual frame in which the needs of others are considered and prioritized within any actions or decisions made of potential consequence to the parties affected. In education, the ethic of care guides educators to consider how their actions and behaviors affect students, and to accordingly make the best possible choices that benefit students' well-being (Shapiro & Stefkovich, 2011).

Ethic of Profession. A conceptual frame in which professional values are used to model actions, behaviors, and decision-making with an emphasis on maintaining accountability with regards to one's accepted role and responsibilities, rather than not doing so. In education, the ethic of profession guides educators to continuously learn and grow to improve themselves and their instructional effectiveness, thereby maintaining professional accountability (Shapiro & Stefkovich, 2011).

Neoliberal Corporatist Paradigm. The modern-day top-down socioeconomic system driven by free market capitalism and large, multinational corporate interests that wields a powerful, pervasive influence over nearly all aspects of life within the United States and much of the world. Within this system, democratic government is generally consigned to a role that is hands-off, if not subservient, to the large multinational corporations and their interests, rendering weak and ineffectual the interests of the public, particularly at the local level (Boas et al., 2009; Jones et al., 2005; Wiarda, 1978).

Positivist Rationalism. The theory of positivism holds that what we see, learn, and know is based upon observation, or literally, what we can see and observe (Miles & Huberman, 1994). Accordingly, if we cannot see something, or else establish a rational basis for its tangible existence, then the subject in question must be viewed as suspect, rather than accepted at face value (Pring, 2000). Positivist rationalism is thus the strict focus on logical explanations for

phenomena to the exclusion of other, less definitive possibilities or the exploration thereof. Its role within education during the recent period has emphasized assessment data (e.g., student test scores) as the *sine qua non* of measuring school performance, and has further led to correlations being made regarding student test data and teacher effectiveness (Craig & Ross, 2008; Goe et al., 2008).

Resident Educator Program. Ohio's program of mentoring for teachers new to the profession who have earned their bachelor's degree, hold a one-year provisional license, and have been hired to teach at least a minimum of two classes or 25 percent of what would constitute a full-time position within the content area of their licensure (e.g., High School Biology, Middle School Mathematics, etc.). Resident Educators are assigned a mentor who assists them with the process of completing their requirements. Resident Educators must successfully complete all program requirements, including passage of the Resident Educator summative assessment, in order to qualify for a five-year professional license and exit from residency (Ohio Department of Education, 2017b).

Student Learning Objectives (SLOs). A process associated with the Student Growth Measures portion of Ohio Teacher Evaluation System (OTES). Through SLOs, teachers are evaluated based upon how students perform on assessments that teachers either individually design and select themselves, or else the content-area departments to which teachers belong (e.g., High School English Department) designate for them to use and be evaluated on as a group. Teachers' performance on SLOs count for the entire Student Growth Measures portion of OTES which currently counts for either 35-percent or 50-percent of teachers' summative rating, based upon whether or not their school district has opted to include Alternative Components (e.g.,

surveys, self-evaluations, student portfolios, or a district-determined component) as 15-percent of teachers' summative rating.

Teacher Evaluation. The process of supervising teachers with regards to assessing their instructional impact on students and providing support (e.g., instructional leadership) for their professional development and growth (Danielson, 2009; Marzano et al., 2011; Sergiovanni & Starratt, 2007). Traditionally within the U.S., teacher evaluation primarily consisted of formal (e.g., summative) and informal (e.g., formative) observations by principals or designated supervisors for the purpose of making decisions regarding retention or dismissal (Sergiovanni & Starratt, 2007). Since the mid-1980s, however, the field has increasingly emphasized teacher evaluations as opportunities for principals to provide instructional leadership in the form of professional support and development (Sergiovanni & Starratt, 2007). During the more recent period, teacher evaluation systems in school districts and states across the U.S. have been expanded and revised to incorporate student performance assessment data as a measure of teachers' instructional effectiveness (Goe et al., 2008).

Teacher Mentoring. A one-to-one relationship in which a teacher receives support and guidance from a more experienced professional colleague (Marzano et al., 2011). Mentors can support new teachers, as well as more experienced teachers who may be struggling in a particular area. Increasingly, in Ohio, mentoring has been regarded as a means of supporting new teachers, a fact reinforced with the advent of the Resident Educator program (Ohio Department of Education, 2017b).

Value-added Measurement. A component of teacher evaluation in which student test scores from the current school year are compared with student test scores from the previous school year. In Ohio, value-added measures may count between 10 and 50 percent of teachers'

summative evaluation rating based at the discretion of each school district (Ohio Department of Education, 2017d).

Conceptual Framework

Within this section, I discuss the conceptual framework I used to approach my study of teacher evaluation from the perspectives of teachers. This framework consists of the four theoretical paradigms most prevalent in education today. These theoretical lenses, as I understand them, are positivist rationalism, critical theory, and feminist theory. Finally, I discuss my personal theory of urban educational leadership that I have derived from these perspectives, and then state how this relates to teacher evaluation.

Positivist rationalism emphasizes and prioritizes quantifiable results that are observable, tangible, and logical. If something cannot be conclusively measured, it is considered suspect when viewed through this theoretical lens and given a back seat to other items better suited to its application (Miles & Huberman, 1994; Pring, 2000). Within education during the recent period, positivist rationalism has emerged to wield substantial influence within school performance measurement and teacher evaluations through an emphasis on student assessment data and teachers' contributions to it (e.g., value-added measures) (Goe et al., 2008). Neoliberal corporatism is essentially an economic system and ideology that prefers unfettered free market capitalism to other alternatives (Boas et al., 2009). It has featured prominently within the postmodern United States and developed world as an ideological lens through which policies are developed for addressing issues within multiple aspects of society, including education (Jones et al., 2005; Wiarda, 1978; Fang, 2014). With regards to education, free market capitalism, corporations and corporate investors have taken on increasingly prominent and influential roles

as agents of change and reform, displacing and diminishing in the process, the voices, knowledge and expertise of actual educators grounded within the field (Attick & Boyles, 2016; Fang, 2014).

Critical theory, a derivative of neo-Marxism, challenges the fundamental aims of the dominant corporate capitalist paradigm and the powerful influence it wields over education through its agents and advocates' manipulation of culture, media, politics, and government policy (Anyon, 2011). Critical theory examines how this paradigm perpetuates the very system and structures of inequity that limit and reduce, rather than provides, the opportunities that proponents of free market capitalism claim with regularity are the forthcoming benefits of current economic structure and policy (Anyon, 2011).

Feminist theory, also derivative of neo-Marxism, challenges the status quo of the dominant masculine, patriarchal paradigm by seeking to view and interpret relationships through what is in essence an ethic of care and of caring (Noddings, 1984). Feminist theory can assist with identifying the people and groups undervalued and unheard within the dominant masculine, patriarchal paradigm. Further, it can also aid in granting them a voice, thereby challenging the normative structures and narrative of the status quo (Noddings, 1984).

In the following paragraphs, I will briefly discuss the critical and feminist perspectives and their relationship with positivist rationalism and the neoliberal corporatist paradigm.

Giroux and the Frankfurt School: Foundational critical theory and dialectical thought. Henry Giroux (2009) discussed the evolution of critical theory within the Frankfurt School, which consisted of Adorno, Horkheimer and Marcuse, among others, and their efforts to reconstruct and rethink notions of human emancipation. The Frankfurt School asserted that "all thought and theory are tied to a specific interest in the development of a society without injustice" (Giroux, 2009, p. 35). Giroux deduced from this assertion that critical thinking was a

necessary component of human freedom and the creation of a better world (2009). Further, critical thinking and critical theory served as a critique and counter-balance to positivist rationalism, which itself was uncritical of science and knowledge, suggesting that a *banking* method of accumulating facts and bits of information is sufficient to constitute learning (Giroux, 2009). The Frankfurt School believed that positivism had supplanted the critical rationality of the Enlightenment, however, and that this in turn resulted in mass social repression through the inherently anti-critical nature of positivism (Giroux, 2009).

Recently, positivist rationalism influenced United States law and policy as evidenced within the No Child Left Behind Act of 2001 (NCLB) and its high-stakes student testing mandates under the administration of President George W. Bush, as well as within the Race to the Top (RttT) teacher evaluation reform initiative in 2009 under the administration of President Barack Obama. NCLB and RttT each represent the triumph of what Giroux (2009) termed "methodological correctness," or a certainty about matching facts to corresponding details, in this particular case, the performance of students, teachers, and schools, without a further, in-depth analysis. Preparing students to pass the test, based upon teacher evaluation emphases, appears to be more important than preparing them to think critically, while the performance ratings that teacher receive on their evaluations do not account for the socioeconomic challenges that individual students face outside of the classroom (Goe et al., 2008). Positivist rationalism thus is consistent with fact-based learning as it emphasizes the importance of certainty as opposed to subjectivity and critical thought (Giroux, 2009). Giroux (2009) asserts, however, that dialectical thought, based in critical thinking and critical theory, can serve to illustrate where established facts and logical systems are incomplete through critique, theoretical reconstruction

and historical analysis. This would seem to form the basis of a possible non-banking approach to education.

Critical: Dewey and holistic education. An additional challenge, based within critical theory to the positivist rationalist approach to learning and measuring the performance of students, teachers, and schools comes from John Dewey (1938). He suggests that students' lived experiences are essential to their education. Accordingly, students' educational experiences must be grounded within the places where they live in order to make education holistic and, therefore, purposeful for them. Further, he suggests that such experiences must integrate principles of interaction and continuity that are critical for proper growth and personality integration within human beings. He asserts that people are developmentally influenced by the sum of their previous experiences, and this, in turn, affects their future. Ideally, this holistic approach to education results in well-integrated human beings that are wholly capable of creating a society that reflects their development.

It is difficult to see how a positivist approach to teacher evaluation in particular would align with Dewey's (1938) philosophy of holistic education. He might therefore question the reductionist approach to learning that serves as the basis for positivist rationalist educational reforms like teacher evaluation embodied within RttT. He might further question how a teacher preoccupied with concern over meeting standards-based instructional requirements or a detailed, graded evaluation rubric, might reasonably be able to inspire students to seek knowledge and understanding beyond these government-mandated basics (Dewey, 1938).

Critical: Aronowitz and sites of control vs. sites of education. Stanley Aronowitz (2009) asserted that schools must play a dual role of preparing their students for the world of work and citizenship while conveying to them the knowledge and awareness of the

Enlightenment. The result of this expectation, however, has been somewhat less than ideal in practice. In reality, schools are underfunded, under-resourced, enforcers of class distinctions and privilege, and purveyors of anti-intellectual bias because of the current prevalence of positivist rationalist methodologies within state and national policy, which include teacher evaluation (Aronowitz, 2009). He suggested that the most immediately effective methods to reverse these trends would be to end high-stakes testing and the banking concept of education (Freire, 2000) along with it, reinforced by RttT and current approaches to teacher evaluation that emphasize quantitative measures of teacher performance. He also suggested that schools must end their relationships with corporations in which they merely serve corporate needs to the detriment of their students' needs, and return to an instructional focus that emphasizes the needs and well-being of their students (Aronowitz, 2009). He would likely find common ground with Giroux and the Frankfurt School in advocacy for critical thinking and critical theory over the positivist rationalism that supports the banking concept of education (Aronowitz, 2009; Giroux, 2009).

Feminist: Noddings and the ethic of care. Noddings (1984, 2007) challenges the positivist rationalist approach to education from a critique based within the feminist ethic of care. She asserts, with regard to concern for the well-being of others, that caring is "reactive and responsive" (1984, p. 19). Further, an individual who cares is present within their actions and deeds to facilitate caring and fulfillment of the relationship. This is true in all relationships and interactions, including within education, between teachers and their students. Her ethic of care (1984) is diametrically opposed to the positivist rationalist approach that perpetuates a world of winners and losers maintained by systemic inequalities and narrowly focused performance measurement systems, as many approaches to teacher evaluation are today. In contrast, the ethic of care fosters an environment in which equality is wholly possible through the valuing of

individuals and the relationship we maintain with them, not as superiors or inferiors, but as equal human beings.

Furthermore, Noddings (2007) directly critiques the modern structure of education and recent reform measures embodied within NCLB, arguing that both are morally flawed because of their unequitable regard for students and their highest individual potential possible. She asserts that mainstay features of modern comprehensive high schools, such as academic tracks and required courses (e.g., Algebra), are designed to achieve standardized outcomes measurable by tests, rather than to holistically nurture students' intellectual and artistic capacities. Standardization and equal treatment of students—expecting all students to achieve the same result—are, of course, hallmarks of the underlying positivist rationalism that has strongly influenced modern education. She summarily rejects these aims, arguing instead for reforms that honor human uniqueness and creativity, and that further constructively support teacher professional development and student achievement. The underlying feminist principles of this reformed approach to modern education that stand in contrast to the reductionism evident within the positivist rationalist view of student and teacher education, are trust and care.

Conceptual framework summary. Within this section I briefly discussed the critical and feminist theoretical perspectives and their relationship with the dominant theoretical perspective of positivist rationalism in education. Both critical and feminist theories challenge positivist rationalism, and in turn, also serve to illustrate the inequitable, stifling and dehumanizing nature of positivist rationalist approaches in education, including within teacher evaluation. These approaches fail to place consideration of individuals' needs and the common good first and foremost, and are the subject of critiques served through the lenses of critical and feminist theoretical perspectives, traditions, reflective practice and similar actions.

The critical and feminist theoretical perspectives support my study on teacher evaluation by providing me with alternative lenses through which I am able to examine it. Critical theory enables me to examine teacher evaluation with respect to how it, as practiced during the recent period, perpetuates structural inequities within education that are detrimental to students, educators and society. Similarly, feminist theory enables me to examine teacher evaluation with respect to how its implementation silences or otherwise minimizes the voices of those who are subject to it. Specifically, the perspectives of Dewey (1938) and Aronowitz (2009) support me in conceptualizing alternative approaches to education based within critical theory, while Nodding's ethic of care (1984; 2007) and also trust (2007) support me in conceptualizing alternative approaches to education based within feminist theory.

Positionality

Within this section, I discuss my positionality as a researcher with regards to this study. This section consists of two subsections. Within the first one, I explain my personal theory of urban educational leadership and how it relates to my objectives. Within the second section, I discuss the rationale behind my research interest and focus. As a researcher, I feel that it is important to openly discuss each of these items in order to establish and clarify my personal relationship and involvement with this research study.

Personal theory of Urban Educational Leadership. We live in a rapidly changing world today in the 21st century, one that is shifting demographically and culturally, propelled forward in both regards by technological innovations within information technology and the reorientation of masses of people towards urban centers as a means of networking socially, culturally, and professionally in order to be a functioning part of society. Looming on the horizon are additional large-scale changes in how people relate that will result from growing

scarcity of vital resources (e.g., water, petroleum, phosphorous) as well as climate change, all of which will force people to relocate and cluster in close quarters within urban environments where resources are available and more readily accessible. It is likely that the hierarchical status quo will seek to maintain control over society and prevail despite emerging new trends that may lead to greater social interaction, and in turn collective action seeking social justice and equality where these concepts remain compromised by imbalance and inequity.

As an urban educational leader, my dedicated purpose is to always serve, negotiate for, argue for, and fight on the side of social justice and equality for all students in education and the communities and peoples of their origin in order to make them universal and triumphant over the existing status quo hierarchy. I do this not simply regardless of the race, ethnicity, gender, religion, or sexuality of my students and the communities and peoples to which they originally belong, but in full awareness of them and with respect to valuing the human diversity that each group and individual contributes towards and enhances by being present and honored as they uniquely are. Schools as educational organizations regardless of their location within metropolitan areas urban, suburban, or rural must become safe places for learning and nurturing of student development and growth into social consciousness to occur unhindered by the corporate hierarchical agenda driven by quantifiable outcomes (e.g., standards-based testing).

My role as an urban educational leader, regardless of my actual position of employment within a given school or place, is to seek to make my school into a safe, beneficial, well-resourced hub of the community. In this regard, school must be a place where authentic learning can and will take place and where all students can benefit fully from participating within it. This must be the case no matter what technological and environmental shifts take place, but just as

much so in complete awareness of these shifts as they are happening, for authentic learning only occurs through comprehensive awareness.

My personal theory of urban educational leadership leads me to investigate and shed light on any impediments to this vision of schools and their potential to serve their students and community in the best way possible. With regards to this research study, my theory of urban educational leadership is concerned with how and why teacher evaluation may be failing to support schools and teachers with the realization of their greatest potential. Also, it is my intent to apply it towards recommending potential solutions.

Research interest and focus. At the time, I began pursuing a subject for my dissertation research study, I was working professionally as an educational administrator within an online charter school based in Ohio. My formal role was "Instructional Supervisor," which entailed evaluating teachers according to a rubric designated by the school's administration as the metric for assessing teacher performance across a number of categories deemed relevant to instructional practice. I took my role very seriously as an evaluator of teachers. My philosophical approach to my role was one based on a personal and professional desire for all teachers to grow professionally and become better at educating—or better still, inspiring self-inspired learning—within their students. Teacher evaluation for me was never about playing "gotcha" and trying to catch teachers in the act of missing the mark and poorly performing on some aspect or another of their job, much to the contrary of the neoliberal corporatist paradigm that seems to hold that teachers intentionally seek to fail their students. I truly believed—and still believe—that all teachers can be successful, if only they have the all the support, resources, and guidance they need in order to fully develop and realize their potential.

My work and philosophical ethic led me to research teacher evaluation in all forms, as well as in different school settings and learning environments, both online and traditional. I reviewed the prominent evaluation systems and philosophies of Charlotte Danielson (2000) and Robert Marzano (2012), and I sought to become thoroughly knowledgeable about current and ongoing developments in teacher evaluation, believing that such knowledge and awareness would give me the resources and power to become a more effective, better teacher evaluator from a constructive, supportive perspective, consistent with my personal theory of urban educational leadership. I believed that I could use evaluation constructively to support teachers through my knowledge, rather than divisively and destructively, as the neoliberal corporatist paradigm seemed bent on imposing in public education through abstract and unproven reform measures supported by prominent captains of industry and politicians. President Obama's 2009 Race to the Top initiative seemed to represent the historical watershed moment in which the positivist rationalist paradigm began to directly impact teachers with its accompanying neoliberal corporatist ideals of quantifiable efficiency, represented within teacher evaluation reforms as well as pressures upon teacher evaluators, such as myself at the time, to conform to the new approach. As such, I feel that I developed my philosophical approach to my role in direct response to the positivist rationalist paradigm's onslaught against education undermined it as a means for addressing inequities and inequalities, and also for improving the human condition.

My admission into the doctoral program in Urban Educational Leadership through the University of Cincinnati then seemed to occur as a timely opening of a door that would enable and empower me to explore my emergent philosophy and awareness therein as it might pertain to my ability as an educational leader to improve the quality of instructional practice from teachers' perspectives, as well as learning from students' perspectives. Indeed, through a fortuitous

meeting with Dr. Gary Pack, the facilitator of the Superintendent Licensure Program I pursued as part of my self-guided path through the UEL program, I gained access to a man highly skilled and knowledgeable within his practice of school district leadership, and in turn, his district itself. I opted to pursue an internship within a Cincinnati-area school district during the Fall of 2012, and this resulted in my gaining of access to a collection of traditional schools, and more importantly, their teachers, staff, and students. It soon became evident to me that my professional work and research interests were aligning as I took note of this school district's participation within the federal Race to the Top program. I requested to join the committee and received full, unfettered access to RttT meetings by the Associate Superintendent and the district's Director of Human Resources.

Through my work and contributions made on the Race to the Top committee, I gained additional insight into teacher evaluation within a traditional school district and the considerations that the district, its administrators, principals, teachers, and staff, as well as parents and even students, all needed to carefully review and process through. Such considerations included professional development time for teachers to become familiar with and acclimated to a new evaluation system, as well as the negotiations that needed to take place between the district and the teachers' union regarding changes to seniority as a priority within decisions about teacher retention, within the collective bargaining agreement. At the same time, I became even more keenly aware of teachers' needs for developmental support, and also, in particular, to be heard.

The concept of teacher voice, and the role it maintained within the school district during the district's development and implementation of a new teacher evaluation system as a result of its participation within RttT resonated with me. I observed that teachers had a representative

voice on the RttT committee, which seemed normal and was what I felt most people as professionals in education would expect to see happen at the district level. What emerged as a concern to me, however, was the apparent lack of concern for the inclusion of teachers and their voice within state education policy in Ohio, particularly with respect to the state's development and implementation of the statewide Ohio Teacher Evaluation System. This concern grew into a desire to formally research teacher evaluation—from the perspectives of teachers, and led to the broad search for and review of pertinent literature that I assembled within the next section.

Significance

To summarize, I began this chapter by stating my intent to address, through this research study, a lack of teacher voice within professional evaluations. As an urban educational leader and based upon my research interest and focus, it is important for me to critically investigate this issue for the purpose of helping schools and teachers to serve their students and communities in the best way possible. Through my application of critical and feminist theoretical lenses, I intend to investigate how positivist rationalism and neoliberal corporatism may have contributed to inequities within teacher evaluation, and in turn, within public education. To achieve this aim, I intend to create opportunities to listen to and amplify the voices of teachers who are subject to evaluation approaches that reflect the values and assumptions of the current dominant paradigm, and that may function to their detriment with regards to teachers' ability to best serve their students and the school community.

Chapter 2

Conceptual Framework

The subject I have selected as my area of interest for my research focus and agenda is teacher evaluation from the perspectives of teachers. Specifically, I am interested in teacher evaluation from the perspective of teachers within primary and secondary public-school education—grades K-12—within the United States. I wish to understand the concept within the context of its role within public education. I note that it is common knowledge that teacher evaluation in public schools is one of the prominent issues within the field today.

By developing an understanding of this subject, I intend to contribute to the ongoing development of knowledge in this area, which may in turn influence education policy. Accordingly, I have conducted a review of literature, from which I developed three themes relevant my analysis of teacher evaluation. These three themes consist of Teacher Perceptions of Evaluation, Evaluation Validity and Reliability, and The Role of Mentoring in the Context of Evaluation, and these themes also form the basis of my underlying research questions and study purpose. What follows is a discussion of the literature that I have assembled. Each piece of literature corresponds with one of the three emergent themes that have served as points of inquiry and analysis within this research study.

Teacher Perceptions of Evaluation

The first thematic category, Teacher Perceptions of Evaluation, represents teachers' perceptions of evaluation with regards to the quality of their professional evaluations in two key aspects. First, this theme addresses how teachers perceive the evaluation process specifically with respect to the quality of the relationship that exists between them and their evaluator, typically a principal, or in some cases, an experienced teacher or retired administrator and former

teacher assigned to the role of supervisor. Secondly, it addresses the extent to which teachers view the evaluation process itself, as being fair, relevant, and thus beneficial to their professional development and growth. My review of research literature resulted in six articles for this theme.

With regards to teachers' perceptions of the effectiveness of teacher evaluations, Zimmerman and Deckert-Pelton (2003) conducted a qualitative study consisting of 86 Northwest Florida educators. The educators were surveyed regarding their perceptions of the effectiveness of principals as evaluators, suggest, based upon findings, that there are four essential components of teacher evaluations that influence their perceptions. These include educator and principal interactions, consistency of evaluations, the commitment of principals towards facilitating professional teacher evaluations, and the background that principals bring to the table with respect to content knowledge, pedagogical knowledge, and professional evaluation knowledge. When principals oversee conducting evaluations, teachers look to them for constructive feedback and support relevant to their individual needs. The authors' overarching findings suggest the evaluation process, when it incorporates the four essential components, is relevant and purposeful for teachers, and also encourages greater buy in from them. Further, they aid principals in focusing on the quality of their interactions with teachers with respect to human relations, and also promote the relevancy of the technical knowledge that evaluators themselves bring to the professional evaluation process.

Within a qualitative case study Kimball (2002), examined how three different school districts used the same evaluation system, how teachers responded to evaluation feedback, as well as their perspectives regarding enabling conditions and the fairness of the evaluation system to which they are subject. Based upon findings, the author suggests the extent to which teachers view evaluations as being fair and constructive, based upon principals' facilitation of them. Also,

they suggest the extent of teachers' trust within their evaluating principals influences their perceptions of evaluation fairness, suggesting a relationship. Further, an enabling school culture and climate can enhance the overall quality of the teacher-principal relationship, something which principals are in the best position to create and maintain within the buildings for which they are responsible as administrators.

In a mixed-methods (QUAN + qual) study, Colby et al. (2002), compared teachers' perceptions of traditional state-mandated evaluation systems versus locally-developed alternative evaluation systems within 21 Northeastern North Carolina school districts. They suggest teachers perceive local evaluation systems as being more constructive and fairer with respect to their professional growth and development, as well as making a stronger contribution towards school improvement efforts than do state-mandated evaluation systems. They further suggest that locally-developed evaluation systems support stronger overall connections between student learning, teachers' professional development, and school improvement efforts, resulting from the fact that these systems take into consideration local school districts' unique circumstances and needs. This highlights the importance of considering teachers' needs within professional evaluations as they are most immediately affected by their results They are also in the best possible position to understand them and to also identify aspects for refinement, much more so than distant policymakers removed from active practice.

Kyriakides et al., (2007), using a quantitative survey, examined the political dynamics of teacher evaluation system implementation in Cyprus schools. Their findings suggested that teachers perceive self-evaluation approaches as being more constructive and empowering than external evaluations facilitated by supervisors in which teachers had less power and autonomy. Such self-evaluations also provided greater opportunities for professional growth. Essentially,

self-evaluations address teachers' need for professional evaluations that are relevant to them because they are embedded within instructional practice, requiring teachers to apply their professional knowledge through self-reflection, as they evaluate their own instructional performance. The author also suggests that the power and authority wielded by teacher evaluators is a potential factor that influences their ability and capacity to provide meaningful and constructive evaluations for teachers.

Ovando (2001), within a qualitative study of elementary teachers within a south-central Texas school district, examined teachers' perceptions of a locally-developed system of teacher evaluation with regards to its ability to support their professional development as a learner-centered system. The researcher's findings suggest that teachers believe that a learner-centered evaluation system is holistic and constructive to their professional growth and development. Further findings suggest that teachers find support for their professional development at both school campus and district levels within this type of system. Some findings expressed by teachers raised concern about subjectivity within the evaluation system with regards to performance ratings, and in general. The researcher suggests a need for additional studies that incorporate teachers' perceptions of evaluation systems.

In a qualitative case study of a teacher supervisor and her subordinates, Burns & Badiali (2015) examined the perceptions of teacher candidates with respect to the dual relationship of evaluation and supervision commonly embodied within instructional supervisor roles. The authors' findings suggest supervision can become synonymous with evaluation by essentially supplanting the latter when instructional supervisors are less-versed in knowledge and skill to provide constructive teacher evaluations. Further findings suggest teachers view instructional supervisors as gatekeepers whose requirements they had to meet in order to receive satisfactory

evaluations. Teachers did not see these same evaluators as constructively facilitating their ongoing growth and professional development. This finding supports establishing a clear distinction between supervisors and evaluators, suggesting that the roles should be separated, which would result in teachers' engagement with professional evaluations becoming a wholly constructive, supportive experience.

To summarize, Teacher Perceptions of Evaluation addresses teachers' perceptions with respect to the quality of their professional evaluations. It highlights the relationship between teachers and their principals or supervisors serving as evaluators, both in terms of what they bring to the evaluation process with respect to their professional knowledge and experience, as well as in terms of the quality of human interactions they are able to facilitate (Kimball, 2002; Zimmerman & Deckert-Pelton, 2003). It also highlights teachers' perceptions of evaluation processes as being fair, relevant, and beneficial to their practice in terms of their professional development and growth, and also regarding the evaluation practices that empower teachers (Burns & Badiali, 2015; Colby et al., 2002; Ovando, 2001; Kyriakides et al., 2007). In general, professional evaluations that prioritize teacher professional development and growth are likely to be the most well-received among teachers than those that do not provide teachers with fair, relevant, and beneficial professional evaluations.

Evaluation Validity and Reliability

The second thematic category, Evaluation Validity, addresses the validity and reliability of the teacher evaluation processes based upon the consistency with which administrators conduct teachers' evaluations based upon standardized criteria. My review of peer-reviewed research literature revealed nine articles highlighting this particular theme.

Providing historical perspective, as well as a theoretical discussion with regards to teacher evaluation and the concern for establishing and maintaining validity within it, Musella (1970) examines inter-rater reliability among teacher evaluators, an issue that continues to feature prominently within the field. The author finds that teacher evaluator subjectivity is influenced by unique individual characteristics that cannot be controlled for in their entirety. As such, he suggests an alternative approach in which evaluators and teachers collaborate within a process focused on teachers' self-reflection and self-evaluation of their performance supported by data and input from the evaluator's observations. This empowers teachers, by giving them the ability to self-evaluate themselves with their principals serving as consultants throughout the process. Accordingly, it can be an effective alternative means to balance and minimize principal subjectivity aside from the goal of achieving inter-rater reliability.

I included this article within my literature review to highlight a potential solution for the issue of teacher evaluation validity that is fundamental to how evaluations are conducted, as the author suggests. This study serves to provide a historical perspective on teacher evaluation that seems to be as relevant today as it was when the author first produced it. What seems missing and needed within teacher evaluations today is the constructive involvement of teachers themselves. As the researcher's findings suggest, teacher evaluation could possibly become more constructive and beneficial for teachers if they are able to participate more directly within the process by conducting their own self-evaluations for at least part of the overall measure of their performance.

Ellis (1986) analyzed five studies on teacher evaluation to synthesize a broad understanding of the concept and challenges to its successful implementation. The researcher discusses characteristics of constructive teacher evaluations and also considers the practical

aspects of evaluation system implementation, such as the validity of measures and their ability to withstand court challenges. He concludes with the general suggestion that schools and their leaders opt for formative rather than summative evaluations whenever possible, noting that summative evaluations are still a practical necessity for personnel decisions. As a historical piece, the researcher's work provides insight into the key concerns with regards to teacher evaluation during the period of publication, with the question of summative versus formative evaluation standing out as the central subject of debate. This issue remains a concern today with regards to the validity of teacher evaluations.

Within a quantitative study of the Nebo School District in Utah, Peterson (1987, examines teacher evaluation that incorporates multiple lines of evidence to determine teacher quality. The researcher's purpose in doing so is to find a better, more valid alternative to the prevailing use of principal observations to assess teacher performance. The researcher suggests that this method may not be the most valid and beneficial approach towards facilitating measurement, growth, and professional development. The evaluation system that the author examines incorporates parent surveys, student reports, peer review, teacher tests, student achievement, professionalism documentation, administrator report, and additional measures of teacher performance. The researcher's findings suggest a benefit to incorporating additional and perhaps contrasting lines of evidence that measure and provide insight into teacher performance, even when these diverge from each other with regards to their results. Although the research work of Peterson (1987) is older, it remains useful for the historical perspective it contributes, as well as from the standpoint of providing insight into the holistic application of other measures of teacher performance within a system of teacher evaluation and their validity.

Harris et al. (2014), within a quantitative study of 30 schools, examined a correlation between teacher performance on value-added measures and principals' ratings of teachers. The researchers' findings suggest principals' observations of teachers' instructional practices are weakly correlated with how teachers perform on data measures (e.g., test scores). This suggests different components of teacher evaluation may yield unique insights into aspects teachers' performance, even when they are measuring the same phenomenon (e.g., teacher instructional effectiveness). This also raises questions about the reliability of each measure, including the contributions of principals as evaluators, as well as the validity of each measure, and the validity of evaluation systems in general with respect to their ability to accurately measure teacher instructional effectiveness.

Papay (2012), within a theoretical discussion, examines evaluation tool quality. He suggests that assessment of tools is necessary in order to determine their accuracy in measuring teacher performance and how they may contribute to teachers' professional development and growth. The author associates evaluation accuracy with inter-rater reliability, suggesting multiple factors may be responsible for variations within the ratings that evaluators report. These factors include principals' individual approaches taken towards conducting evaluations, as well as the personal standards they maintain, and also unique differences within the lessons they observe being taught. Accordingly, there are many small factors can lead to variations within teacher evaluation outcomes, not all of which can be controlled to perfection within evaluation systems, or between evaluators themselves, which is the goal of achieving inter-rater reliability.

Darling-Hammond et al. (2012), conducted a review of research on value-added models and similar data measures used within summative teacher evaluation and suggested that such systems are flawed as they cannot control for wide variations in student population

characteristics from one year to the next. Factors influencing student performance that are beyond the teacher's control include home and community support systems, class sizes, attendance, peer culture, and loss of knowledge during summer break. The researchers suggest that value-added measures may still be used more formatively to validate more reliable, standards-based measures, such as National Board Certification or new teacher licensure or residency programs (e.g., Ohio Resident Educator).

Baker et al. (2010), within a briefing paper, examined problems resulting from use of student test scores through value-added modeling to evaluate teachers. These included significant year-to-year inconsistencies between tests and student outcomes that run counter to ideas of relative consistency in teacher performance. The authors suggest that there are other factors accountable for these inconsistencies that have little if anything at all to do with teacher performance, such as school attendance, economic mobility, health, and the qualities of home and neighborhood environments. They concluded that granting student test scores exclusive or considerable weight in teacher evaluations is both unfair and can have detrimental and damaging effects on teacher morale and the educational profession in general by deterring talented, capable and potentially effective teachers from entering into or remaining within the profession. They also asserted that law and policy makers should avoid focusing on testing as an exclusive or substantially-weighted measure of teacher performance, and that perhaps alternate, they should consider it as only one measure balanced with multiple other measures of teacher and student performance.

Firestone (2014), within a theoretical discussion, examined deficiencies within the dominant view of what constitutes teacher evaluation in an attempt to suggest an overarching theory that more accurately defines the functionality of teacher evaluation. Specifically, the

researcher compared extrinsic motivation using money as a stimulus, against intrinsic motivation. He suggested that, between the two approaches, extrinsic motivation does not appear to be particularly effective, while intrinsic motivation holds greater potential, but warrants additional observation and research in order to determine its effectiveness in application. He concludes that effective teacher evaluation should combine balanced and meaningful measurement with a clear interpretation and application of theory to support teacher professional growth and success.

Firestone and Pennell (1993), within a theoretical discussion, explore a framework for measuring teacher job commitment for retention, promotion and incentive compensation consisting of measures for feedback, job characteristics, participation, resources, autonomy, collaboration, and learning opportunities. They suggest that teachers are suspicious of bias within observations based upon evaluators' personal preferences and favoritism towards some teachers over others. Further, they suggest evaluator experience, instructional knowledge, and content knowledge may negatively influence evaluation outcomes when evaluators are unfamiliar with their evaluees' area of expertise. In addition, the authors suggest the number of evaluations conducted can also potentially influence evaluation outcomes. Thus, there are multiple factors that can lead to variations within teacher evaluation outcomes, although evaluators can themselves seek to minimize bias and favoritism within the evaluation process, and also within the general context of the principal-teacher relationship.

Smagorinsky (2014), within a theoretical discussion, examines the considerations that would make teacher evaluation systems legitimate and would further contribute to the strengthening of the field. These include commitment to reliability and validity in evaluations, stakeholder buy-in, teacher development, participant utility, and use of multiple measures in a

formative, rather than summative, context. The author posits these items as an essential framework capable of challenging the current paradigm in which student test scores have become the quintessential measure of teacher effectiveness. The author suggests that the current prevailing focus on using a single quantitative measure of teacher performance within evaluations is driven by an impulse within the current paradigm to isolate and remove poor performing teachers from the profession. The author further suggests that teacher development must instead be the focus of performance evaluation and that embracing a system incorporating multiple measures will produce a more positive overall performance result.

Croft, Roberts, and Stenhouse (2016), within a case study of education reforms in Georgia, examine the results of neoliberal policy efforts originating with the No Child Left Behind Act of 2001. They suggest that these efforts, which have required mass testing of students and summative teacher evaluation systems, have not only fallen short of their aims, but that they have actually expanded the inequities such efforts were intended to resolve. Further, these policies have served to undermine confidence in public schools and teachers. They have also negatively impacted teacher professional development programs and teacher evaluation by emphasizing measurable performance within teaching over all other considerations necessary for teacher preparation, raising questions about the validity and reliability of such a narrow policy focus. The authors conclude that the neoliberal platform in education must be opposed with alternative policies that better support teachers and public schools.

To summarize, Evaluation Validity and Reliability addresses the validity of evaluation measures as well as the consistency with which individual evaluators rate teacher performance based upon standardized criteria. It further accounts for efforts to control for variations within evaluation results. Inconsistencies between evaluation measures, and also between evaluators

themselves, can limit inter-rater reliability and raise concerns about the general validity of a particular evaluation system (Harris et al., 2014). Research has also challenged the validity of value-added models and data measures in teacher evaluations (Darling-Hammond et al. 2012). Accordingly, these concerns suggest a need to grant greater weight and priority to formative evaluations instead of summative evaluations (Croft et al., 2016; Ellis, 1986; Smagorinsky, 2014) as well as evaluation systems that incorporate multiple measures (Croft et al., 2016; Peterson, 1987; Smagorinsky, 2014). It may not be possible to control all threats to evaluation validity and inter-rater reliability, however (Papay, 2012; Firestone & Pennell, 1993). Principals serving as evaluators, however, may be able to strengthen inter-rater reliability by seeking to minimize their own biases and favoritism within their relationships with teachers (Firestone & Pennell, 1993). Finally, empowering teachers within the evaluation process by granting them self-evaluations with principal support can also address teacher concerns with inter-rater reliability (Musella, 1970). While achieving evaluation validity and reliability within teacher evaluations may challenging to fully achieve, alternative models and approaches to facilitating teacher evaluation exist that can potentially support realization of such aims within education.

The Role of Mentoring in the Context of Evaluation

The final thematic category, The Role of Mentoring in the Context of Evaluation, addresses the coaching as professional support teachers receive that is directly or indirectly related to the teacher evaluation process, with respect to fulfilling their professional growth and development potential based upon performance expectations. Six research articles contributed to my understanding of this theme.

Danielson (1999) addresses mentoring for beginning teachers. She discusses the importance of using a defined strategy for mentoring, and suggests the problems that can result

when beginning teachers are not mentored. Further, she discusses the experiences of teachers new to the profession, how they can grow through mentoring, as well as what a well-designed mentoring program looks like. Finally, she suggests that mentoring can serve as professional development for new teachers, as well as for their mentors. It is noteworthy that Danielson has incorporated consideration for mentoring into her evaluation framework, suggesting that it "may be used as the foundation of a school or district's mentoring, coaching, professional development, and teacher evaluation processes, thus linking all those activities together and helping teachers become more thoughtful practitioners" (The Danielson Group, 2017)

Marzano and Simms, within a professional development book (2012), and webpage (Marzano, 2017), examine coaching as a non-evaluative means of helping teachers to improve their instructional performance, and also provide strategies for doing so. The authors suggest that coaching as formative mentoring helps teachers to grow and improve because they feel more comfortable examining their practice with a peer teacher who is not summatively evaluating them, rather than a supervising administrator who is doing so. The authors further suggest that the goal of coaching must be to help teachers improve their practice in one particular area at a time in order to achieve and sustain a constructive result. It is noteworthy that the theory of mentoring that the authors address within this book is also incorporated into Marzano's (2011) evaluation framework, most notably within *Domain 4: Collegiality and Professionalism*, suggesting that professional interactions with colleagues and support sought or received from them are important to professional growth and development.

Smith and Ingersoll (2004), using the 1999-2000 Schools and Staffing Survey (SASS) conducted by the National Center for Education Statistics, examine data from a sample of 3,235 new teachers generalizable to the population. The authors seek to determine whether new

teachers receiving mentoring and related supports during their first year were more likely to remain in their positions or leave them at the end of that year. SASS contains questions that address teachers' perspectives on mentoring, induction, and similar supports that they may or may not receive during their first year within the classroom. The researchers use the survey data to address the extent of mentoring programs within the U.S., as well as teacher participation in these and similar mentoring activities, new teacher rate of turnover, and the effect of mentoring and similar supports on teacher retention. The researchers' findings suggest that new teacher participation in mentoring programs increased between 1990 and 2000 from 40 percent to 80 percent of the population, while approximately two-thirds of these teachers were assigned a mentor to work closely with. During the same period, however, the rate of attrition among new teachers increased from 11 percent to 26 percent. The researchers also suggest that teachers with mentors assigned to them within their field (e.g., content or instructional knowledge area) reduced the rate of turnover by approximately 30 percent, while having access to colleagues with regards to instructional planning reduced new teacher attrition by approximately 43 percent. Overall, teachers appear to benefit from mentoring and the presence of collegial support with regards to their professional success and retention.

Johnson and Birkeland (2003) used a longitudinal interview study of 50 Massachusetts teachers new to the profession to understand why beginning teachers either remained within their public school of original employment, departed this school for a position within another school, or else left the teaching profession altogether within their first 3 years. The researchers found consistencies between the professional experiences of present day teachers and teachers of the past, such as limited pay, resources, and opportunities for advancement, as well as isolation and lack of authority to make decisions, suggesting these issues within the profession have not

significantly changed over time. Additional findings suggest factors such as prior work experience, gender, and race also influence teacher retention. The findings that are most central to this theme suggest that teachers are concerned about being evaluated by mentors, and as a result, may be unable to confide in and seek the help from them that they need in order to be successful. Collectively, these findings suggest multiple possible avenues for improving teacher retention within public schools, while highlighting the need for the mentor and evaluator roles to be separated in order for teachers to constructively benefit from both roles.

Holloway (2001), within a theoretical discussion, addressed the benefits for benefits for novice teachers as well as experienced teachers. For novice teachers, mentoring can support them with addressing their professional challenges, while for both novice teachers and experienced teachers, it can serve as a professional development opportunity, although mentors will need their own training to serve as effective mentors. In addition, a mentoring program that is well-designed can help to reduce teacher turnover. The author concludes by suggesting that mentoring programs can professionally benefit inexperienced teachers, and in turn, help students as well as their teachers improve, while also providing professional development for experienced teachers.

Ingersoll and Strong (2011), within a research review article, conduct a critical examination of 15 empirical studies regarding new teacher induction programs. The authors suggest that most of the studies indicate that new teacher induction, including induction that involves assigning new teachers to work with mentors, achieves positive outcomes for new teachers with regards to three areas: teacher commitment and retention, student achievement, and teacher classroom instructional practices. One notable exception was a randomized controlled trial study that took place in a large, low-income urban district, in which findings suggested the

new teacher induction program had a positive effect on student achievement, but did not have a noticeable impact on teacher classroom practices or teachers' commitment and retention. The researchers suggest that most of the existing research does not examine the fundamental causal factors within the success or failure of induction programs, while further suggesting that the duration of induction programs may be a factor with regards to the ones that succeed. This article is relevant to this research study from the standpoint of providing insight into the benefits of teacher induction programs for new teachers.

To summarize, The Role of Mentoring in the Context of Evaluation encompasses coaching and support for teachers that is either directly or indirectly associated with teacher evaluation. Mentoring can serve to help teachers improve their professional practice outside of teacher evaluation so that teachers are better able to perform and meet evaluation standards (Danielson, 1999; Marzano & Simms, 2012). For new teachers, mentoring can help retain them within the profession while preparing them for the expectations of performance they must meet and be evaluated upon (Holloway, 2001; Ingersoll & Strong, 2011; Johnson & Birkeland; Smith & Ingersoll, 2014). Mentoring thus serves as an avenue for teacher professional development, growth, and improvement that is both immediately and indirectly related to the evaluation process, but in either instance serves as an essential part of it.

Summary

The focus of this conceptual framework and review of research literature was the subject of teacher evaluation. I selected this subject because of my interest in teacher evaluation from the perspective of teachers within primary and secondary public school education in grades K through 12 within the United States. My intention in conducting this literature review was to

better understand the concept of teacher evaluation from the standpoint of the contexts of its role and priority within US public education.

Each of the three categorical themes, consisting of Teacher Perceptions of Evaluation, Evaluation Validity and Reliability, and The Role of Mentoring in the Context of Evaluation, suggests an area that may be of interest with regard to my study of the subject of teacher evaluation. In preparing for this literature review, I found rather troubling the trend of corporate concepts such as quantifiable results and efficiency taking precedence within education over more traditional, fundamental democratic principles and values (Attick & Boyles, 2016; Dewey, 1938). This is particularly disturbing to me in light of research suggesting that test scores such as quantitative value added measures of student performance—and substantial components in many new and reformed teacher evaluation systems—are neither valid nor reliable, and may even be exposing schools and state governments to lawsuits, as teachers are wrongfully dismissed for low student performance on tests over which they have no actual control (American Statistical Association, 2014; Baker et al., 2010). Further, in many instances teachers were apparently either granted limited input or completely excluded from the decision-making processes that resulted in the implementation of these questionable reforms (Baker et al., 2010). While I cannot personally fathom how any teacher would view the current predominant trend in performance evaluation as being helpful either to them or for the teaching profession, I have remained open to the possibility of receiving alternate perspectives on teacher evaluation in its current form that may provide a more comprehensive picture of teachers' experiences with it, including those that are constructive and positively received.

Nevertheless, I have found as a result of my own personal summative reflection and review of the literature on teacher evaluation that I am all the more inspired to pursue my

commitment to social justice in education. From the perspective of a researcher, as well as being a teacher myself, I see a clear and pressing need for teachers within K-12 education to have a voice within research on teacher evaluation specifically, but also in all aspects of our profession. Policy decisions have taken place with respect to education during the current period that did not meaningfully include the voices and professional knowledge of the people best suited to inform and direct them, namely teachers. As a result, teachers are now subject to questionable policies that are not only poorly serving them and their instructional practices, but are almost certainly undermining and hurting them professionally and in ways that harm the educational opportunities of the students they serve. Through my research, I intend to do my part to challenge the existing paradigm and hopefully give teachers a stronger and more influential voice within educational policymaking.

Chapter 3

Research Methods

Within this chapter, I discuss the research methodology I used for conducting my research study into teacher evaluation from the perspectives of teachers. This includes my methodological approach, procedures, participant recruitment and sampling, as well as site selection, ethics, and teacher positionality. Further, this section details the research methods I used to conduct this study, consisting of participant interviews, Group Level Assessment (GLA) methodology, and first-person reflection. Finally, I also discuss the processes of data analysis I used, as well as my approach towards establishing validity and trustworthiness.

Methodological Approach

Facilitating this exploratory research study into understanding teacher evaluation from the perspectives of teachers necessitated an open-ended methodological format. Accordingly, I selected an emergent qualitative multimethod approach in which the research study would be conducted in several stages (Hunter & Brewer, 2015). Each phase, and my findings therein, would inform the subsequent phase and choice of methods therein for the purpose of validating findings and establishing sufficient depth of understanding regarding the subject of my study. The three research themes I developed from my review of research literature, consisting of Teacher Perceptions of Evaluation, Evaluation Validity and Reliability, and The Role of Mentoring in the Context of Evaluation, each serve to organize my findings across each of the phases of this research study. Further, by addressing these three themes during each phase of my research study, I am also addressing the underlying research questions for this study which I developed from my review of the literature and subsequent establishment of the three themes.

I must note here that findings from this research study are not intended to be generalizable to the broader population of public school teachers in Ohio or elsewhere, based upon the methodology used to collect and analyze data ultimately leading to these findings (Rubin & Rubin, 2012). Instead, the methodology used within this study is designed to support validation and trustworthiness of findings specific to this research study, and in turn, the recommendations and conclusions drawn from it (Hatch, 2002).

Procedures. Because of the nature of the qualitative methodological approach, my sample size did not need to be particularly large. In speaking with my committee with regards to the precise size of this sample, we determined that between eight and ten teachers representing the school district would suffice for the original Phase I set of interviews I conducted, as these interviews would together constitute sufficient depth for qualitative data analysis (Rubin & Rubin, 2012). Accordingly, facilitation of eight interviews became my target goal, which I was able to achieve. I sought out interview participants on a selective, voluntary basis, with a goal of assembling diverse perspectives across general teacher demographic characteristics (e.g., age, race/ethnicity, gender, experience).

Upon reviewing the original data I collected for the eight teachers I successfully recruited and discussing my initial findings with my committee, we determined that the strategy of expert review, or member-checking interviews, with the teachers I had previously interviewed would provide depth to the data while also serving as a method of validation. This follow-up stage constituted Phase II of my research study.

After collecting and then reviewing the member-checking interview data, I discussed my findings with my committee. We determined at this point that facilitating a GLA activity among teachers within the district where I had conducted the interviews would serve to provide an

additional layer of depth and validity for this study. The GLA activity thus constituted Phase III of my research study.

Finally, after reviewing my original data, the member-checking data, and the GLA activity data, I once again discussed my findings with my committee. We determined that conducting a personal self-reflection as Phase IV of my research study would provide one final layer of depth and validation for my research analysis, based within my own professional perspective. This is consistent with reflective practice action research, which supports practitioners with understanding our collective construction of reality, as well as practitioners' own relationships with this collective construction (Taylor et al., 2008). Validity is established through practitioners' act of understanding the issue, reflection upon this issue, and then acting to address it.

Site selection. I initially considered including several states within the scope of my research study, particularly Illinois and Pennsylvania, in addition to Ohio, because of similarities within their evaluation methodology. Constraints regarding time and money for travel, however, required me to restrict my research focus to Ohio, and more narrowly, a single metropolitan area within the state.

With my research focus narrowed to Ohio, the participants that I sought for participation within this study were Ohio-licensed teachers actively working as classroom teachers within Ohio's traditional (e.g., "brick and mortar") public schools within the Greater Cincinnati region, and within the school district that became the focus of my research study in particular. These teachers, by state law, were subject to OTES. As such, their professional performance evaluation consisted of two components, periodic supervisor (e.g., principal) observation of professional practice, and annual tracking of student growth achievement, weighted equally (50-percent for

each measure during 2013-2014) and combined into a single summative rating each year. OTES became effective by law at the beginning of the 2013-14 school year, standardizing evaluation procedures at a statewide level that were previously left to the discretion of each of Ohio's 613 public school districts. Accordingly, all teachers that participated within this study were evaluated using the same evaluation system and methodology for its execution for nearly one school year as of the time my research began.

Although I had originally planned to recruit teachers within multiple Cincinnati-area school districts, the district I selected proved to be my best option for the purpose of completing this research study because of previous work I had completed there as part of my doctoral program through professional contacts I established with the Superintendent of the district and other staff and faculty. Further, at the time, this district was a candidate for Race to the Top funding, a 2009 initiative of the Obama Administration intended to encourage educational reform, particularly with regards to teacher evaluation (U.S. Department of Education, 2016). With the issue of teacher evaluation at the forefront of the district's priorities, it was a logical choice to seek to conduct this research study within it.

Participant recruitment and sampling. Once I established the metropolitan region and subsequently the school district that constituted the stage for my execution of this research study proposal and plan, I designed a general recruitment letter to the district superintendent. I submitted this letter as part of my research proposal to the Institutional Review Board (IRB) and received official approval from to proceed with my research study during Spring 2014. Upon obtaining the documentation representing the written consent of the district superintendent to proceed with my research study within their district, I placed it into a secure file I created for the district. I then proceeded to coordinate teacher recruitment through the district staff and

administrators that the superintendent designated as points of contact and facilitators of communication with teachers.

In order to facilitate the recruitment of teachers, I designed a similar general recruitment letter as the one I created for communication of my research interests and purpose with the district superintendent (Appendix A). This IRB-approved letter also outlined the focus of my research study, as well as my research interests, professional background, and purpose for undertaking the research study, and further contained a brief overview of my research study's components, data security, and participant confidentiality, in addition to describing the potential means by which I intended to report my research findings, as well as the target audiences for my research study findings. I used the unaltered contents of this letter within e-mail communications to teachers that district personnel distributed with the superintendent's approval. As such, the e-mail message contained the same detailed overview of the research study purpose and indicated the consent of the district granting facilitation, as well as a notice of confidentiality detailing teachers' rights. Further, it also provided an overview of my background as a researcher, and included my research interests and professional background, my contact information, and a detailed description of the security procedures I used in order to maintain the confidentiality of data.

Ethics. Upon meeting with teachers to conduct qualitative interviews with them, I provided them with information on research data security and confidentiality within an IRB-approved consent form. This document contained a description as to how I would separate teachers' identifying information from their responses and replace this information with unique tracking identifiers, while placing the corresponding names within a separate and secure location, within password protected and locked storage. I also explained to participants within the consent

form that I would not share this information with their supervisors or colleagues, with the one exception being reasonable evidence that harm is occurring to the participant by another individual or that the participant is causing harm to someone else. Further, I explained within the consent form that I will destroy and securely dispose of any identifying data three years from completing my initial report on the findings from my analysis of their data that I have collected (Appendix B, Appendix C, Appendix D.

During recruitment I informed participants of their right to opt out. I wanted all participants to be aware in accordance with IRB protocol that they had the right to opt out of the study at any time, for any reason, or for no reason at all, without fear of repercussion from the researcher, supervisor, or place of employment. In addition, I also provided participants with my phone number and contact address information, as well as the phone number and contact address information for my program and research advisor, in the event that they wanted to opt out of the study, before, during, or after they have participated within it. I asked all participants to read and sign such forms outlining and detailing their rights as participants, as well as documenting their consent to participate within the study and to allow me to record and use their information prior to involving them in any direct aspects of the research study.

With regard to the possibility of participant coercion taking place, I explained to teachers using IRB-approved literature, as well as during their receipt of the consent form from me, that they were under no obligation to participate in the research study at all, and should not have experienced being made to feel in any way that they would be subject to punishment or sanctions of any kind for refusing to participate within my research study. In addition, I advised participants that I would secure their data in a manner that would protect their identity and prevent possible retaliation against them by a third party for their participation within my

research study. This information included their contact info—phone and email address—as well as their licensure/teaching assignment and grade level. I limited my presentation of this info only to what was absolutely necessary in order to further protect the identities of participant teachers and reduce the chance that they could be identified.

I must also note here that very late in my data collection activities, I recruited the help of Dr. Doug Stevens to help facilitate the GLA activity that I incorporated into my research study at the suggestion of my committee, based upon his knowledge and prior experience with facilitating group level assessments. In order to maintain compliance with the IRB, Dr. Stevens completed all required forms and training updates, and we submitted all documentation that the IRB requested to it.

Teachers' positionality and study participation implications within a positivist-rationalist neoliberal corporatist paradigm. In today's climate surrounding public education in the United States, teachers face challenges on multiple fronts. For decades, they have been dealing with issues such as poor pay and often challenging working conditions. More recently, they have also had to deal with reduced respect and autonomy, as well as increasingly harsh criticism regarding their professional contributions, value, and effectiveness based upon questionable measures (American Statistical Association, 2014). Ultimately, amidst all of these issues, however, teachers today are dealing with neoliberal corporatist pressures to privatize public education, resulting in school program commodification and market competition taking priority over traditional values based in democratic ideals (Blakely, 2017; Dewey, 1938).

The result is a situation in which teachers, both collectively and as individuals, figuratively find themselves with their backs against a wall and in the sights of a hostile ideology that threatens their undoing and the end of public education as it has been. On a daily basis,

teachers are challenged to maintain and uphold their core ethical, professional values (Shapiro & Stefkovich, 2011). At the same time, they face new and renewed imperatives—educating against hate and intolerance, promoting educational opportunity and promise for their students, and promoting learning as a democratic right (Strauss, 2017).

It is worth considering that teachers' participation within such a study could be viewed as an act of defiance against the positivist-rationalist neoliberal corporatist paradigm. At the same time, teachers' participation within this research study may also be considered as an affirmation of the core ethical, professional values that are at stake in the present climate. In either regard, I anticipated that my research study would highlight the positionality of teachers with regards to the current positivist-rationalist neoliberal corporatist paradigm and their response to it, both collectively, and as individuals, across each of its four phases.

Phase I: One-on-one interviews. I preferred to use one-to-one semi-structured interviews in order to complete this initial phase of my research study because I wanted to maintain the focus of interviews around specific questions relevant to the study, which the semi-structured interview approach permits (Rubin & Rubin, 2012). At the same time, however, I also wanted to give participants the opportunity to share their thoughts, and the semi-structured interview approach provides some flexibility for divergence from the questions, as well as exploration of an emergent topic of interest related to one or more of the questions. The protocol (Appendix E) that I followed within the creation of this instrument included an icebreaker introductory question, followed by a series of open-ended questions that would ideally support my efforts in terms of producing and collecting high-quality qualitative data (Rubin & Rubin, 2012). In addition, there was room for probing questions in order to further support my gathering of thick and descriptive data. As such, I added sub-questions under two main questions.

Recruitment, participants, data collection intent and procedures. Teachers throughout the school district were recruited to participate within this research study based upon the information communicated to them by the district from the letter to the Superintendent (Appendix A). There were no exclusionary criteria for recruitment, and the only inclusionary criteria was that teachers needed to be actively licensed as Ohio teachers and actively teaching within a classroom. A total of eight teachers were ultimately recruited across grade levels, subject areas, and professional levels of experience for Phase I.

The interview protocol instrument for Phase I (Appendix E) contained a total of nine primary questions as well as two written probing questions (with the probing questions appearing as extensions of questions 2 and 7). I used the instrument to provide a guiding structure for the interview conversations I held between myself and the teachers that participated in this research study. I did not necessarily hold fast to the protocol questions if, within a particular interview conversation, a topic or idea came up that the teacher wanted to expand on, or that I was curious to learn more about; prompting us to deviate from the protocol and discuss it. At the same time, I used the protocol to maintain a general structure for all of the interviews I conducted in order to make it easier to reference data across interviews.

Immediately following each interview, or as soon as possible, I transcribed the notes I recorded within the protocol form onto a Word document in order to clarify and preserve my understanding of the data recorded. I also then assigned a unique case id to the document for tracking purposes, and stored it securely on the computer containing all of the data, sans identifiers, for my research study. The data from this initial phase, based upon my initial analysis and findings, would inform my priorities for the subsequent phases I would undertake.

Phase II: Follow-up interviews. The second phase of my research study, consisting of follow-up interviews with the teachers I previously met with, mirrored the interviews I conducted during the first phase. Essentially, the procedures I followed during this phase were identical to those that I used during the initial phase. Furthermore, no new teachers were interviewed during this phase.

Recruitment, participants, data collection intent and procedures. With regards to the recruitment of participants for the second phase of my research study, as this particular phase represented a validation of the methods and data collected during my first phase, I simply followed-up with all eight of the original participants during Phase I to interview them again for Phase II, using their contact data. As during Phase I, there was no exclusionary criteria for recruitment during Phase II, and the only inclusionary criteria was that teachers still needed to be actively licensed as Ohio teachers and actively teaching within a classroom. Of the eight original teachers, seven participated within Phase II, while one subject declined further participation within this research study. The data I collected from this second phase would inform my priorities for the subsequent phases I would undertake.

The interview protocol instrument for Phase II (Appendix F) was identical to the interview protocol instrument I used for Phase I. It contained a total of nine primary questions as well as two written probing questions (with the probing questions appearing as extensions of questions 2 and 7). As during Phase I, used this instrument to provide a guiding structure for the interview conversations I held between myself and the teachers that participated in this research study. Also, as before, I did not necessarily hold fast to the protocol questions if, within a particular interview conversation, a topic or idea came up that the teacher wanted to expand on, or that I was curious to learn more about, prompting us to deviate from the protocol and discuss

it. At the same time, once again, I used the protocol to maintain a general structure for all of the interviews I conducted in order to make it easier to reference data across interviews.

As during Phase I, immediately after each interview, I transcribed the notes I recorded into the protocol form onto a separate document to support further clarification of the data recorded. I also assigned a unique case id to the document for tracking purposes, and stored it securely on the computer containing all of the data for my research study. The data from this second phase, combined with the data I collected during my first phase, would inform my priorities for the remaining two phases of this research study. At the conclusion of phase two, I combined the data from both phases to enrich the overall quality of interview data within this research study.

Phase III: Group level assessment activity. During my third phase of data collection, in May 2015, I conducted an activity informed by GLA methodology for the purpose of gathering additional feedback from teachers based upon the themes I coded from my phase one interviews and validated during my phase two interviews through the process of member checking. The GLA procedures I adapted for this phase consist of participants providing written responses to prompts within an open setting, followed by a debriefing in which they discuss their responses. This GLA activity was open to all teachers within the school district, and an invitation for voluntary participation was accordingly communicated by the Interim Superintendent to the district. Two teachers responded to the invitation. Included within this count was one teacher completely new to my research study, and one whom I had interviewed during phases one and two of my data collection efforts. Although teacher turnout for this particular phase was not what I had hoped for or anticipated, I was nonetheless able to collect a generous amount of data from

each of the two participating teachers during the two-hour span in which the GLA activity took place at the school district office.

Recruitment, participants, data collection intent and procedures. Teachers throughout the school district were recruited to participate within this research study based upon the information communicated to them by the district from the letter to the Superintendent (Appendix A). There were no exclusionary criteria for recruitment, and the only inclusionary criteria was that teachers needed to be actively licensed as Ohio teachers and actively teaching within a classroom. A total of two teachers were ultimately recruited across grade levels, subject areas, and professional levels of experience for Phase III.

The interview protocol for Phase III (Appendix G) consisted of six prompts that were arranged on separate large wall sheets throughout a meeting room at the district main office in which the GLA activity took place. Teachers were given the opportunity to walk around the room for approximately 30 minutes and respond to any and all of the prompts of their choosing. Afterwards, a discussion took place in which participants were prompted to discuss their responses as well as any additional insights they could offer with regards to the posted prompts. While this discussion was taking place, I recorded notes based upon the teachers' responses to the prompts, which I later transcribed in order to support further clarification of the data collected during this phase.

At the conclusion of the GLA activity, I collected the large wall sheets containing the prompts and reviewed them, along with my notes based upon my discussion with the teachers. This information would be used to further validate the data I collected during the earlier phases of this research study.

Phase IV: Self-reflection. Finally, I completed a first-person reflection of my professional experience in education, consistent with a reflective practice action research approach (Taylor et al., 2008). This reflection was based upon a protocol (Appendix H) and generated relevant data for this research study based upon the emergent themes within my data analysis and literature review. It served to triangulate these same themes through the consistencies in which they manifested within my own professional experiences as a public-school teacher subject to being evaluated.

I anticipated that my findings of this research study would provide insight into teacher evaluation from the perspective of teachers. I also anticipated that these findings would provide insight into how teachers can best approach and respond to professional evaluation. At the same time, however, it is my hope that the process that I enlisted here to establish them may be pursued by other researchers within schools and districts across the country and abroad who seek to understand teacher evaluation from the perspective of teachers within educational organizations local to them.

Intent, data sources, and procedures. The intent behind incorporating a self-reflection phase into this research study was to provide additional validation for the data collected during the previous three phases through a process of practitioner self-reflection (Taylor et al., 2008). The data source for this phase is my own perspective, based within my knowledge and experiences as a professional educator. The data for this phase was generated using a self-reflection protocol (Appendix H) that I used to record my insights and reflects for each of the three themes addressed within this research study.

Data analysis. To facilitate data collection, I used an interview protocol form for note taking as well as digital audio recording software on my personal iPad to record the interview. In

all instances I used the AudioMemos software on my iPad to record interview conversations, and I used copies of the interview protocol form for documenting conversation notes and participant responses to questions. The AudioMemos software created an audio file that I could download and secure on the computer that I used to store the data that I collected for the purpose of completing this research study. I used these audio files to complete the transcription process, which consisted of transferring the recorded audio into typed transcripts that I used during the analysis phase of my research study.

The analysis of qualitative data I conduct within Chapter 4 is based upon the common themes I developed during my review of research literature on teacher evaluation, in Chapter 2. These themes consist of Evaluation Perceptions, Evaluation Validity and Reliability, and Mentoring, and serve as the basis of the topic coding process (Hatch, 2002) I used to support my analysis of qualitative interview data. I reviewed printed and digital transcripts for each interview and coded for each of the three themes within each transcript. Also, I documented and tracked these codes for each transcript within an Excel spreadsheet that I created, maintaining and securing them within the Excel sheet on the same computer where I stored my research study files and data, while also keeping track of primary codes for reference and further analysis (Appendix I). This initial stage of data analysis, of the phenomenon of teacher evaluation from the perspectives of teachers using teacher interview data, serves as the basis for the deeper analysis I conduct within the next chapter. This consists of examining each of the three themes as they appear among the interview data and exploring their relationship with the central phenomenon of my research study, professional evaluation as teachers relate to it.

Validity and trustworthiness. In order to ensure the validity of qualitative data from interviews, I sought to obtain thick, rich descriptive data. Accordingly, I set out to conduct

interviews during my first phase of data collection, in Spring 2014, that were at least 30 minutes in length, and in five of eight instances I was able to meet or surpass this goal, with one particular interview lasting well over two hours. The shortest interview during this phase was 19 minutes. The overall average interview time per teacher was 53 minutes.

Within my second phase of data collection in Spring 2015, I conducted member checks with the teachers I interviewed during my first round of data collection. Member checking is the act of verifying previously gathered data with study participants in order to determine if they agree with their contributions (Hatch, 2002). I sought to conduct interviews that were at least 20 minutes in length. I actually met or surpassed this goal for five of the seven interviews I was successfully able to complete during this phase, which consisted of me sharing my study findings—the emergent codes I found—with the original teacher participants in order to gather their feedback on the validity of these findings while also reviewing the original questions I asked them during our first phase interviews.

The longest interview spanned 64 minutes while the shortest interview lasted only 12 minutes. This resulted in an average of 32 minutes among the seven interviews I completed. The eighth teacher exercised her right to opt out and decline any further involvement in my research study when I sought her participation within this portion of my research study.

To serve as an additional measure of validity, my first-person self-reflection draws upon my personal experience as a practitioner. It further incorporates my understanding of the theoretical framework that I have developed to support my inquiry into the topic of teacher evaluation. This enables me to clarify and establish relationships between various findings, ideas, and concepts within the context of this research study.

Limitations. There are three limitations that I have considered with regards to the scope

and relevancy of the findings I present within this research study. I discuss them within this section as follows.

First, I note that the majority of teacher interview participants within my research study are veterans of the profession. I anticipated that I would likely receive a stronger overall response from established veteran teachers within the district who perhaps had achieved a sufficient level of comfort within their role to feel that they could contribute to my research study. At the same time, I actively sought out the participation of mid-career and beginning teachers as well, acknowledging that they would be able to provide insights for my research study based upon their particular level of professional experience, as well as based upon their time of entry into the field as professionals and being influenced in outlook and approach by the particular trends and initiatives in effect within public education during that time.

Regrettably, despite my recruitment efforts, I was not able to secure the participation of any beginning teachers employed within the school district during the interview facilitations that took place during my data collection phase for this research study. As such, the perspectives of beginning teachers on teacher evaluation were not directly featured within the interview data I collected for this research study. I note, however, that I was able to secure the participation of a first-year teacher during the GLA data collection activity that I facilitated at the end of my data collection phase. As such, I was able to include perspective data from a beginning teacher in a somewhat more limited form than what a one-on-one interview might have yielded to me.

Second, I also regret that I was only able to secure the participation of one male teacher, and then only at the secondary instructional level. I acknowledge, however, the rarity of male teachers in elementary education today. At the same time, I feel fortunate in being able to secure the participation of three non-White minority teachers, which enabled me to gather and provide

insight into teacher evaluation from the perspective of teachers in education belonging to racial or ethnic minority groups.

Third, although participant withdrawal was very low for this research study, I nonetheless lost one participant who voluntarily opted out of further participation within this study during the second round of interviews before I was able to schedule an interview with them. Their self-termination took place in full accordance with participants' rights and all ethical considerations associated with research and the role of the primary investigator in facilitating the research study. As such, however, I was unfortunately unable to access this participant's perspective for added depth and insight into this research study. At the same time, I understand, respect, and uphold fully the rights any and all participants to opt out of participation within a research study at any given point in time of their choosing, and without question.

Finally, the GLA activity I facilitated as part of this research study functions optimally when participation from the group for which it is intended to serve and provide insight into some aspect thereof, is generous. As such, I anticipated that I would achieve a minimum threshold for participation of approximately 20 teachers from throughout the district. Despite my recruitment efforts, regrettably only two teachers attended. Time and schedule constraints, including state testing, the approaching end of the school year, as well as the pending changeover in district administration leadership at this particular point in my data collection phase in Spring of 2015 made further efforts to organize and implement a follow-up group level assessment activity untenable. At the same time, these two teachers generously contributed their ideas.

Summary

Within this chapter, I have detailed the research methodology I used to conduct my research study into teacher evaluation from the perspectives of teachers. This discussion included ethics, site selection, participant recruitment and sampling, as well as teachers' positionality and its relevance within the context of this research study. Further, I provided an explanation of the each of the four phases that comprise this research study, including my considerations for participant recruitment, as well as with regards to intent, data sources, and procedures during each phase. Finally, I discussed my procedures for performing data analysis as well as with regards to establishing validity and trustworthiness overall with regards to the data I have collected, as well as my findings, based upon this data. This chapter has thus served to provide the framework for this research study of teacher evaluation from the perspectives of teachers.

Chapter 4

Findings

Within this chapter, I discuss the findings that emerged from my analysis of interview data collected during the first two phases of my research study, as well as from the Group Level Assessment (GLA) activity and self-reflection data collected during the third and fourth phases of my research study, respectively. I do so specifically with regards to three themes, Teacher Perceptions of Evaluation, Evaluation Validity and Reliability, and The Role of Mentoring in the Context of Evaluation. I explore each theme by sharing the voices of the teachers that contributed to my understanding of it. Further, I support their voices with links to the research that served to shape my thematic framework, while providing my own analysis of the overall findings and what they mean within the context of this research study.

Teacher Perceptions of Evaluation

Teacher Perceptions of Evaluation is the first of three themes based upon the thematic framework I developed from my review of research literature. As a theme, it specifically addresses aspects of teachers' perceptions of evaluation support concerning their professional needs, as well as the centrality of the evaluator within the process and the need for such persons to be well-trained. The intent here is to gain understanding with regards to how teachers perceive evaluation, and it is supported by the varied contexts of teachers' accounts during the initial and late stages of their careers. Evidence for the Evaluation Perceptions theme was common during the conversations I had with teachers during my first and second phase interviews, and also appeared during my third phase GLA activity.

First and second phase interviews. I inquired about teachers' initial experiences with evaluation both as an ice-breaker question as well as to provide a point of reference for

comparison between past and present experiences for teachers' benefit, as well as for the benefit of this research study. I discovered that teachers' initial experiences with evaluation were somewhat varied between individuals. Teachers' time of initial entry into the profession, as well as their knowledge of the stakes associated with their initial evaluation, or lack thereof, however, influenced the general perception of their initial evaluation experience.

Joyce, a veteran music teacher, first entered the profession at a time when teacher evaluation was much less defined, compared with today. She discusses teacher evaluation as a rudimentary part of teaching, something that took place during her early years within the profession with few specifics regarding what it was supposed to be about, or what actual purpose it served. In this form, teacher evaluation was the exact opposite of the rigorous, accountability-focused models that are in common use within the K-12 education field today. She further describes her perceptions of the lack of focus regarding evaluations, including with regards to the role of the facilitating administrator:

> Nobody really said what it was supposed to do, really… what [was its] intention and purpose… There was no preconference. You really didn't know what people were looking for. Somebody sat and scribed in the back of your room. The… these very normal kids who were there would see [the principal] and he'd be scribbling and [scrabbling] and so forth and so on. And of course, the kids got their heads turned like this. [They tried] to figure out what the deal is. (3a. 87-89, 97-103)

Joyce's early experiences with teacher evaluation did not seem to be helpful or meaningful. She suggests even her students were unclear about the principal's purpose for being there, and she could not provide them with a rational explanation for his presence. None of this appears to be consistent with the research-based concept of teacher evaluation as serving a constructive

purpose, specifically with regards to promoting the growth and development of teachers (Colby et al., 2002; Kimball, 2002; Kyriakides et al., 2007). Teacher evaluation should ideally be an experience in which teachers receive constructive feedback and support that helps them to improve their professional practice.

The initial experiences of Tara, a veteran secondary teacher, are similar to Joyce's. She describes her first evaluation experience as consisting of a simple checklist form, and her evaluator's observations based upon it:

> My very first experience was a checklist… do you see these things, yes or no, and then you were handed the checklist. That was in the… old [NCR] forms. One copy went to the central office, one to the principal, one to me, sign it, and very little summative feedback. (8a. 19-23)

She also describes the frequency of principal visits, and her perceptions of the quality of feedback she received from them:

> I don't recall the principal coming in more than one time. I think it was 15 minutes, and back… it was in the elementary school so I think [it went] checklist, are there bulletin boards, is the room engaging, those kinds of things that you could have blank walls and the kids could be engaged, and you could have beautiful things and the kids be totally [off]… but it was not designed for true feedback for the teacher. It was basically, I think, for evidence that there's something that's happening in the classroom regarding teaching and learning. [LAUGHTER]. But nothing specific. (8a. 41-50)

Tara's experiences, like Joyce's, do not suggest that teacher evaluation is being used as constructively as it could be, such as to provide teachers with professional development (Ovando, 2001). At best, their initial evaluations seem like just another obligation. At worst, it

was a complete waste of time for teachers, who were unclear as to how they were being evaluated and apparently did not receive much if any feedback afterwards.

In addition, similarly, Paul, a veteran secondary social studies teacher, describes his first-year experience with professional evaluation:

> Well, I was a first-year teacher in Cincinnati Public. In terms of the actual process piece, it was very straightforward, here are some things we're going to be looking for, here's your final rating… I… was going through the motions. It was one more hoop I had to jump through. I was mainly focused on my class, my planning, my kids... (2a. 20, 24-26, 32-34)

As with Joyce and Tara's experiences, Paul's perception of his initial evaluation experience was that it was a simple process, one that also apparently did not lead to constructive feedback for him, which would have consisted of professional dialogue between he and his evaluator (Ovando, 2001).

Paul, Tara, and Joyce's early experiences with teacher evaluation each stand in stark contrast with the approach to teacher evaluation in use today within Ohio and a number of other states focused on teacher instructional accountability. In Ohio under OTES, evaluators do not use simple checklist forms when they observe teachers, but rather, a multi-domain rubric (Ohio Department of Education, 2017c). Further, they observe teachers several times during the course of a school year, and are required to provide feedback to teachers based upon these observations. Paul's recollection of his early career experiences with teacher evaluation, combined with Joyce and Tara's early career accounts, collectively illustrate a picture of teacher evaluation as being both unimportant and unhelpful to teacher growth and development in the form that it existed during their early teaching experiences, the opposite of what it can and should be (Colby et al.,

2002; Kimball, 2002; Kyriakides et al., 2007). Perhaps most notable here, however, is the fact that neither Paul, nor Tara, nor Joyce perceived their initial evaluations to be particularly stressful or demanding, a quality of experience that was not universally shared among the teachers I interviewed.

When it was apparent that teachers' initial year of evaluation came with high-stakes attached—the potential to be rehired or dismissed from their positions, as has been the recent trend with teacher evaluations, and OTES specifically—these teachers indicated they perceived their evaluations as being stressful. Stephanie, a mid-career music teacher, describes the professional expectations and pressures she perceived during her initial period of professional evaluation, as a Resident Educator Teacher:

> [S]someone from the [Ohio Department of Education] was to come in at the end of my first-year teaching, which was kind of a provisionary license type situation, and… watch my teaching and observe and look for evidence that I was a competent teacher in the four different domains of the Danielson framework. So needless to say, it was a little bit… stressful knowing that if I don't represent here in this first year that someone may see something within my teaching that would keep me from obtaining a license. (1a. 72-80)

What Stephanie discusses is the opposite extreme end of evaluations—those that actually do matter and are focused on teacher accountability, standing in contrast to those that appear trivial, such as what Joyce, Paul, and Tara experienced. The former type carries high-stakes, not only for one's employment, but also for one's career in education, as Stephanie notes. If her evaluator had not rated her performance as satisfactory, she may have indeed been denied a teaching license, and with it, the opportunity to be a public-school educator. It is therefore little surprise that she would perceive such conditions as being stressful, part of which could also perhaps be attributed

to the fact that her evaluator was an outside person whom she had no relationship with, and accordingly, few, if any, opportunities to seek out professional dialogue and support with them beyond her observations.

Sarah, a veteran elementary school teacher, similarly indicates that first-year evaluation experience as a teacher was a make-or-break, high-stakes scenario that she had the potential to fail. In contrast with Stephanie's experience, however, she suggests that she did receive professional support from her principal, who was also her evaluator. She perceives this support as having made a difference for her and helped her to succeed:

> I was nervous having someone come in to evaluate me, but the experience was [fine]... I was glad that I was asked back for the following year. [LAUGHTER]. I did fine on my evaluation. I just remember being nervous, having someone come in my classroom and watch me teach and kind of write down what I was teaching. But I had a really good first principal, and [both the] principal and assistant principal [were] really [supportive] my first year. So, it was an… overall positive experience. (6a. 25-26, 29-35)

In Sarah's instance, although she experienced stress during the evaluation process, the support she received from her supervising principals helped to mitigate her stress somewhat to help her perceive her evaluation experience as being positive, and ultimately a successful one, more so than perhaps it might have been without their involvement. She did not go into great detail about the type of support she received from her principal, but her perception here regarding what she received and how it impacted her evaluation experience is what is most important. Research does suggest, however, that teachers want and do benefit from positive teacher-principal relationships when it comes to professional evaluation (Kimball, 2002).

In contrast with Sarah's perceptions of her initial evaluation experiences, however, Karen, a veteran elementary teacher, perceived her own initial evaluation experience with her principal as being very demanding and negative:

> I was a first-year teacher and in our district [teacher evaluation] was really rigorous. We had to have 10 hours [of] evaluation as a first-year teacher… So we had a really tough entry-year program… But what I remember most about my first evaluation… You… felt that you were… it just seemed it was so glaringly negative. There wasn't anything positive in the evaluation…. I …wrote a rebuttal. [I] wanted [the principal] to know that there were things… surely there must have been something positive so that [I] could grow from it or feel good about [myself as] a new teacher. And I remember that… Really negative and really harsh. He couldn't have found one thing that [I was] doing well. (5a. 32-40)

What stands out most about Karen's perception of her initial evaluation experience is the abject lack of support she received from her principal, whom she describes as focused on finding fault with regards to her performance while never offering a constructive word of praise and support. Clearly, her experiences were qualitatively worse than what Sarah experienced. Whereas Sarah perceived both feeling under pressure and stress for having to meet what was expected of her through the evaluation process, but ultimately having the constructive support of her principal, Karen perceived her evaluation experiences as being overtly negative and unfair based upon how they were facilitated by her evaluator. This particular contrast and comparison underscores the importance of having principals that are competent evaluators who are capable of working with teachers and can also offer constructive feedback to them (Burns & Badiali, 2015; Zimmerman & Deckert-Pelton, 2003).

After teachers shared their initial evaluation experiences during the first and second phase of interviews, I then turned the focus within each interview conversation toward teachers' recent evaluation experiences and their perceptions of the evaluation process within the current context of their role as teachers within K-12 education. My purpose in shifting our conversation was to better understand teachers' views on evaluation in the present while also building a broader frame of reference for comparison and analysis with respect to existing research findings. While all of the teachers I interviewed are subject to OTES, all of them also shared different perspectives on their relationship with it and how, if at all, it affects them within a professional capacity.

Tara expressed a positive perception with regards to professional evaluation, expressing that she enjoys it, while also embracing it as an opportunity for professional growth:

> [Evaluation is] something that I enjoy. I think it's an opportunity for me to show my skill set while opening the door for opportunity to grow in areas that may be areas that aren't my strength. So …that's not something that intimidates me or that I shy away from. I have an open-door policy in my room. Anybody, parents, other teachers, whatever can come anytime. So it doesn't bother me. (8a. 56-63)

Tara evidences what Shapiro and Stefkovich (2011) would describe as a professional ethic, one in which she is driven to give her best and meet all challenges no matter what they are, in the spirit of service to her constituents, namely her students, their families, and the community itself, as well as her colleagues. She clearly views teacher evaluation as a source for her professional growth and development, consistent with findings suggesting the importance of professional evaluations to teacher growth and development (Kimball, 2002).

Like Tara, Paul expressed a similar perception of professionalism regarding teacher

evaluation and his relationship with the process and current system:

> I've been through the [OTES] training. I feel like I have a… solid understanding of the process, the rubric. I also have a very good working relationship with my administrators. So on a personal level, it has not been a source of stress to me. There are a lot of teachers who are 180 degrees from every piece of that puzzle that I just laid out. [B]ut personally, to me, … I've got little to no beef with it. (2a. 225-231)

Paul completed the OTES evaluator training, which Ohio-licensed teachers also have access to, but are not required to complete, in order to understand the evaluation process and measures by which his evaluators would judge his instructional performance. Further, he expresses that he maintains a positive relationship with his principals. Clearly, like Tara, he embraces Shapiro and Stefkovich's (2011) ethic of profession, seeing the evaluation process as something that is constructive to his professional growth and development, and embracing it as such, also supporting research findings that suggest the importance of professional evaluations to teacher growth and development (Zimmerman & Deckert-Pelton, 2003). He does suggest, however, that his outlook and reception of teacher evaluation within its present form is more than the exception than the norm among his colleagues.

Also similar to Tara and Paul, Angela, a veteran secondary English/Language Arts teacher, expresses a positive perception of teacher evaluation. She embraces the OTES evaluation system because in her view, teachers have power and agency in determining their performance outcomes through it:

> I actually like… OTES. [T]he actual OTES stuff in general I like because I feel like the person being evaluated... has all of the… you have all of the control in… determining what evidence is presented, what evidence is validated. No one can just make up

> something and put it in and you'll have… you can say well, what is the evidence for that? Where did you see that? When did you see that? Is this hearsay? So I like that from the standpoint that it… definitely zeros in on specific things. And then I like the areas that are listed. (7b. 20-27, 29-36)

Angela embraces the focus OTES places on accountability as she refers to the OTES rubric, which consists of multiple domains and categories that serve as "look-fors" for evaluators during their observations of teachers. Also, very similar to Tara and Paul, she expresses a favorable attitude towards the teacher evaluations process and being held accountable for her performance, seeing its focus on specific details as empowering. Further, she seems to share a similar ethic of profession by embracing OTES as an opportunity to demonstrate—and be assessed—for her performance (Shapiro & Stefkovich, 2011).

In short, Tara, Paul, and Angela each perceive teacher evaluation as being beneficial to them and openly embrace it as an essential component of their professional growth and development. Sarah also expresses a similarly positive outlook while looking back on her experience for the current school year:

> I… had a pretty positive experience this year. [M]ine went well. My rating came out skilled, which I'm… happy to be skilled. [LAUGHTER] It's not the highest, but it's… hard to get the highest. So for me, the experience went well. I think evaluating… teachers is important… I think it is important for teachers to get feedback because I'm… I made changes this year that I wouldn't have made… if I wouldn't have been evaluated. So you can always grow as a teacher. (6a. 174-176, 178-182)

Sarah indicates that she was motivated to make changes because of the evaluation she received, embracing feedback on her performance as an opportunity to make constructive changes with

regards to her instructional practice. This is consistent with research suggesting teachers benefit from evaluation feedback (Kimball, 2002; Zimmerman & Deckert-Pelton, 2003).

The perspectives of two teachers, however, stood in contrast with the positive views and professional ethic their colleagues shared towards teacher evaluation, Karen expresses experiencing stress and overt negativity through the process of being evaluated, regardless as to what form it takes, and whether her end rating is satisfactory or not. She describes her general experiences:

> It's stressful. It's… like there shouldn't be such [stress] and I shouldn't carry that weight of feeling so stressed. I don't really have any reason [to feel stressed] except for one [negative] experience over all these years… I still find it, even now, really stressful [to] be evaluated. (8a. 85, 100-103)

Karen emphasizes that she experiences all of her evaluations as stressful and emotionally heavy experiences, based upon one experience that took place earlier in her career (that she shared earlier during our interview conversation regarding her initial evaluation experiences). It is striking that so many years later into her career during the recent period, this long past experience continues to affect her, having seemingly made a permanent impression on her schema of professional evaluation in education. A single instance in which a principal did not have the skills and preparation to serve effectively as an evaluator for a teacher resulted in long-term damage to the same teacher's perception of the evaluation process and its effectiveness, the very opposite of what teachers need and expect from their evaluating principal (Zimmerman & Deckert-Pelton, 2003).

Similarly, Julia, a mid-career secondary social studies teacher who had recently completed an OTES evaluation cycle with a new principal serving as her evaluator, summarizes

her perceptions of OTES as an evaluation system:

> No teacher wants to be labeled as less than an accomplished teacher. So that's just demoralizing. So I mean it's… hard. I really feel like there are other models that are more growth oriented and coaching oriented and this one is truly not. It's… truly not. (4b. 140-144)

Julia refers to the *Accomplished* rating, which is the highest among four possible ratings a teacher can receive within the OTES system, the others within the sequence (from second highest to lowest) being *Skilled*, *Developing*, and *Ineffective* (Ohio Department of Education, 2017c). During the 2016-17 school year, teachers in 94 percent of school districts did not receive less than a rating of Developing through OTES (Kelley, 2017). The concerns Julia expresses regarding OTES, however, are not only with the rating she received, but with her perception that the evaluation system itself does not promote teachers' professional development, essentially providing them with a critique of their performance, but with limited if any accompanying guidance on addressing evaluators' concerns regarding their performance. She further suggests this stands in contrast with Danielson's system of evaluation (Danielson & McGreal, 2000). Julia's perspective regarding what teacher evaluation should be is consistent with her colleagues' perspectives, as well as research findings suggesting the importance of having principals as evaluators that provide teachers with constructive and meaningful feedback, and whose content knowledge background and expertise is a match for their teachers' needs.

Group level assessment activity. During the GLA activity phase of this research study, I again prompted teachers to share their perceptions of teacher evaluation. Two relevant comments were recorded. The first one highlights a new teacher's perceptions about their experience:

> I was evaluated three times. The principal spent time pre- and post- the first section. I felt

> that…. little value or oversight was gained in the post-conference on the part of the principal listening to why I did something in a particular way. However, his explanations were insightful and enabled me to modify my teaching. Additionally, knowing where you stand in the category, i.e., skilled or low, medium, high, that matters.

The teacher who made this particular comment apparently did not find much value within the act of reviewing the session observed with their evaluating principal, stating as much. At the same time, however, they did perceive their principal's feedback as being useful to them from the standpoint of supporting their professional development and growth as a teacher by giving them something to build from, including with regards to a performance rating. This latter point echoes the perceptions of teachers during the earlier phases of my study with regards to the value and importance of receiving constructive feedback from an evaluator, and then being able to use this feedback to improve their performance. This is consistent with research suggesting the prevalence of a general desire among teachers for constructive professional growth and development opportunities through evaluations (Kimball, 2002; Zimmerman & Deckert-Pelton, 2003).

Next, the second comment highlights the evaluation perceptions of an experienced teacher:

> I have found that I love the professional conversations about my teaching. My evaluators on the balcony approach forces me to identify, critique, as well as reflect why I did what I did, use the materials I chose and behave as I did with my students. Practice… this practice made me a better teacher. If I don't agree, I express my thoughts and ideas through written reflection and conversation.

This comment clearly embodies the professional ethic and outlook of an established, veteran

teacher, one who embraces evaluations as opportunities for professional reflection, refinement, and growth. There is a certain amount of comfort that appears to be present here, which does not, by any means, suggest complacency on the part of the teacher who wrote this comment. Rather instead, this teacher is at a place within their career where they understand their capabilities but remain open to seeking new opportunities for growth and professional development. In this regard, they perceive—and embrace—evaluation as a fresh, new challenge. This particular teacher's perception of the evaluation process actually is not very different from that of the beginning teacher who authored the first comment. Both teachers see it as a source and opportunity for growth, as is once again consistent with research regarding teachers and professional development (Kimball, 2002; Zimmerman & Deckert-Pelton, 2003). The key difference, however, is their experience and the amount of self-efficacy that accompanies it. The beginning teacher wants to survive and get better, a theme reflected within teachers' perceptions of evaluation during their initial experiences in the classroom. The established veteran teacher, however, perceives evaluation as an opportunity for continuing growth and refinement of their professional practice.

Personal reflection. From my own standpoint as a mid-career educator who, as of the time of this study, is teaching in a traditional classroom setting, my perceptions regarding teacher evaluation are similar to those of the teachers I have interviewed here. Just as they desire to grow and improve professionally as educators, so do I. Similarly, as Joyce, Paul, Tara, and Sarah also do, I embrace the ethic of profession and seek to effectively meet and successfully answer any potential challenge to my legitimacy that may emerge (Shapiro & Stefkovich, 2011). Accordingly, I understand that feedback on my performance in the classroom can serve an important purpose that can prepare me to this end, by helping me to better understand my

strengths, as well as my areas in which I can improve in order to become an overall stronger and more effective teacher.

One thing that is also important for me, however, as several of the teachers I interviewed also indicated, is my professional relationship with my evaluator. As Sarah in particular suggested, having the support of your evaluating principal through a challenging first year in the classroom, or while making the transition into a new school, course, or grade level, can make a world of difference with regards to your survival and successful adaptation in a new role or setting. The contrast with this is an ugly and unfortunate situation similar to what Karen described with regards to her own first year, and one that I am also familiar with. In such a situation, you find yourself without the support of a principal and are forced to seek support, feedback, and encouragement wherever you can to succeed and survive. This is not at all beneficial to teachers, nor is it helpful to the school and its mission to serve students, or to education in general. Simply put, teachers, and particularly new teachers, need support, praise, and encouragement from their evaluating principal. They need to know that not everything is going to work out during their first year, and that this is okay, because it should be okay to try and fail at something so long as you learn something from it and improve on the next try. Within a school or professional community where learning and growth are encouraged throughout, not only for kids, but for adults as well, teachers are able to receive the support for their professional growth and development that they need, where they need it most (Dewey, 1938; Kimball, 2002; Zimmerman & Deckert-Pelton, 2003).

Summary. With regards to the theme of Teacher Perceptions of Evaluation, the teachers I interviewed, as well as the ones that participated within the GLA activity, collectively share a common interest in wanting to improve their professional performance and be successful as

teachers. This held true despite a wide range of perceptions among the teachers regarding the evaluation process. All teachers seemed to perceive the potential for evaluation to support them with regards to these aims, although their individual experiences with teacher evaluation have been both positive and negative. Within the positive instances, teachers seemed to display a professional ethic that enabled them to constructively perceive teacher evaluation as a challenge to prove themselves and their value to their chosen profession (Shapiro and Stefkovich, 2011). Even with regards to negative evaluation experiences, however, teachers perceived the potential for evaluation to serve them and to support their professional needs.

Most pivotal with regards to teachers' perceptions of the quality of their evaluation experience, however, are evaluators themselves, who more often than not are principals with the ability to shape their schools' climate in ways that can either support or undermine professional development opportunities for their teachers. This suggests a need for principals serving as evaluators to be well-trained in the evaluator role so that they can provide constructive feedback and support for teachers within a positive climate (Burns & Badiali, 2015), and one in which evaluations are conducive to the growth and development of teachers (Colby et al., 2002). Findings across all four phases of this research study support teachers' expectations that principals evaluate them who can provide constructive, relevant and meaningful feedback supportive of teachers' professional growth and development (Kimball, 2002; Kyriakides et al., 2007; Ovando, 2001; Zimmerman & Deckert-Pelton, 2003).

Evaluation Validity and Reliability

The theme of Evaluation Validity and Reliability addresses teachers' perspectives regarding the validity of their evaluations with respect to their principals' background instructional knowledge, professional experience, and familiarity with their assigned teachers'

content area. The specific role and contributions of principals within teachers' evaluations featured prominently within each of the interview conversations I held with the eight contributing teachers. This section is divided into two subthemes consisting of Teachers' Impressions of Evaluation Validity, and Teachers' Impressions of Reliability.

Teachers' impressions of evaluation validity: First and second phase interviews. Initially, I asked teachers to share their thoughts on the teacher observations that principals conduct as part of the evaluation process. Paul, a secondary social studies teacher, offered his perspective:

> Well, I think [they're] important. There needs to be accountability. In a perfect world, the evaluation is also an opportunity to help the teacher grow… [I]t's part of the job… [I]n any line of work, somebody is going to evaluate your performance. And in teaching, this is the way it comes about. (2a. 47-51)

Paul expresses an ethic of profession in which accountability in the form of observations is not something to be feared, but rather embraced as an opportunity for growth (Shapiro & Stefkovich, 2011). Even if teachers could somehow avoid principal observations of their practice, from Paul's perspective, it would not be a professional action with respect to the need for accountability, an idea also supported by research (Croft et al., 2016; Ellis, 1986; & Smagorinsky, 2014). Further, willfully avoiding principal observations would also mean missing out opportunities for professional and even personal growth, the essence of education and of being an educator (Dewey, 1938).

Tara, a veteran secondary teacher, offers a similar perspective of principal observations as professional development:

> When done effectively, [teacher evaluation] allows for my growth and continued acquisition of skills, which is what we would all… no matter what our profession is, we want to get better at what we do and find new ways to engage, not only the kids in learning but myself in the process of designing lessons and planning lessons and executing lessons… [W]hen it's used as professional development, which is what I really see the teacher evaluation process as being is ways to develop myself as a better professional… [then] I become better at what I do and everybody benefits from that. (8a. 176-185)

Like Paul, Tara embraces principal observations of teachers as potentially constructive opportunities for professional learning and growth, which is again consistent with the ethic of the profession (Shapiro & Stefkovich, 2011). She expresses that teachers want to improve themselves professionally, not only to be able to better engage their students in learning, but to further develop their competency in all aspects of their professional role.

Curiously, however, Tara also suggests here that teacher evaluation effectiveness, or quality, is a potential concern for teachers with regards to their professional growth and development, if not, perhaps in some particularly challenging instances, their employment. I attribute her concern to the role of principals as the school officials charged with the responsibility of facilitating teacher evaluations, and observations of teachers specifically. Why would principals' observations constitute a concern for teachers? As a thematic lens, Evaluation Validity offers some insights within my interview conversations and GLA activity with teachers, as well as within my self-reflection regarding the role of the principal as evaluator. Further, it addresses the skills and experiences that principals bring to this role, and the challenges and

limitations they face within it. In short, many principals are unprepared to serve as teacher evaluators, thus creating problems with evaluation validity.

Julia, a secondary history teacher, suggests one possible reason for issues with Evaluation Validity is the fact that many principals are teachers with relatively limited classroom experience who left the classroom seeking higher-paying positions:

> In general, most of the people that we have either as administrators or as superintendents have been out a while. So if you are a third, fourth year teacher and you want to make a pretty big [pay] jump, get your administrative degree. Once you get hired, you want to get your master's so you can get a raise. Well, if you want to make… good [money], come to administration. So you could have two, three, four years' experience the classroom and you're hopping out to administration and then you stay in administration forever. What do you know? (4a. 738-740, 743-752)

The question she raises at the very end is a critical one. What indeed do principals and superintendents know about teaching that they can meaningfully share with teachers about instances in which they themselves have limited experience in the role? Also, how relevant is the knowledge that they do have, particularly as they become farther away from fresh involvement within their own instructional practice? The competency of principals as evaluators is a key concern of teachers. Experience, or lack thereof, can influence evaluation validity with regards to the quality of evaluations that teachers receive from administrators.

Concerning why teachers with more years of classroom experience could potentially contribute to stronger evaluation validity, Julia offers an explanation based upon her personal perspective:

> [Why we] don't have the best teachers jumping to administration. Where I am right now, I don't like the pay freeze, but I'm… a master's degree… master's plus 10, I'm frozen at nine years' experience even though this is my 14th year. So for me to be frozen at $65,000, it's not the end of the world. [W]ould I like to be making more? Yes. Do I think I deserve a raise? Yes. I'm a two-income house. My husband makes a nice salary; $65,000 is decent money. If I were to jump to be a principal, they'll probably pay me 70 [thousand], 75 [thousand]. I'll be working all of July. I won't get to go home, ever. I have to work every football game and basketball game, so why would a good middle of the road teacher jump to administration? (4a. 758-770)

As Julia explains, the financial incentive for teachers to become administrators who are responsible for conducting teacher evaluations narrows as teachers gain more experience and the higher pay that accompanies it. It is ironic that better, more experienced teachers have less incentive to become principals (who bear the responsibility for ensuring and improving instructional quality), than teachers with more limited experience and professional knowledge. This obviously contributes to the lack of quality evaluations and teachers' instructional growth, which depends on evaluator experience and the quality of their contributions (Firestone & Pennell, 1993; Papay, 2012).

Tara offers a perspective that, like Julia's, highlights the fact that many administrators who conduct teacher evaluations are not necessarily the most well-prepared for the responsibility. The difference between their perspectives, however, concerns the general competency of principals as evaluators, regardless of experience:

> I think… it becomes evident, much like there are teachers who aren't very good at what they do, there have been administrators who aren't very good at what they do. And

> perhaps they weren't very good teachers and they don't recognize components of good teaching and so there's not really ways for them to… coach and counsel you through areas of growth. And so for me… I've done this long enough to know that I've got a handle on what's going on in my classroom. I don't need somebody to say I think you know what you're doing. I know what I'm doing. But I would love to know this is what I think would take you to the next step. And when you don't get that, because the person who's evaluating you is not a skilled teacher, that's annoying. That's happened. (8a. 69-80)

Tara suggests that administrator competency to conduct teacher evaluations is a concern that transcends instructional experience as well as administrator readiness and ability to serve as an evaluator. This contrasts with Julia's perspective suggesting that amount of time spent within the classroom prior to becoming an administrator is the key factor in evaluator competency. Where both teachers agree, however, is that administrator competency is essential, which supports strong evaluation validity within teacher evaluations (Firestone & Pennell, 1993).

Tara further offers additional perspective on evaluator competency, based upon her experiences with multiple administrators who have served as her evaluators over the years:

> So I think it really depends on who your evaluator is and how successful they were in the classroom, or your perception of their success in the classroom, or [with-it-ness] as far as classroom necessity. …I've never gotten a bad evaluation. I've always had good evaluations. But they've varied as far as how much it was helpful to me. So for example, this year I ended up with an accomplished rating, both for my evaluator and my SLOs, but I don't think I learned a damn thing. (4a. 275-282)

In Tara's view, the ratings she received on her evaluations have never been a real concern for her. She was concerned, however, about the actual quality and helpfulness of the evaluations she received from her administrators. While some of the evaluations she has received have been relevant and more helpful to her with regards to improving her instructional practice, she expressed that the evaluation she received this year from her administrator was unhelpful, if not downright useless to her.

As a veteran teacher, Tara clearly expects, and rightfully so, that the evaluations she receives can still help her with regards to improving her practice, reflecting the same ethic of profession that Paul also expressed (Shapiro & Stefkovich, 2011). She suggests, however, that the quality of match between administrator and teacher is an important, if not critical, aspect of evaluation quality. This is a point further reinforced by research findings suggesting the importance of evaluator-teacher match, with similarities in content knowledge and instructional experience being key areas of concern (Firestone & Pennell, 1993).

Angela, a veteran English/Language Arts teacher, echoes Paul and Tara's ethic of profession, as well as Julia and Tara's concerns about evaluation validity. She also raises her own concerns about the instructional background of the administrators who conduct teacher evaluations. In particular, she questions how they can effectively support teachers licensed in a different content area than the one they themselves taught prior to becoming administrators:

> I actually have no problems proving that I'm accountable for what it is I'm supposed to be doing. What I have a problem with… and this is not speaking to my evaluation this year or my evaluator… I have a problem with being evaluated by people who weren't able to prove that they could do the job that I'm doing. And I think that's the biggest sentiment among most teachers that a lot of our current administrators have… were not in the

> classroom. They might have been Special Ed people or they were in the classroom when we were in the classroom and it's a completely different system so they don't understand the constraints of a) not being able to teach anymore, to teach to a test, to teach to these evaluative categories. And it's not… those of us who have been around a long time know that there are some valuable things… I want to be evaluated by someone who is competent in the area that… I'm in and I know that they're competent in the area that I'm in. They're just not in a position because someone moved them there. (7a. 424-440)

Angela, like Tara, also suggests that the match between evaluator and teacher is essential. She wants to receive a meaningful, relevant evaluation that supports her professional development and growth as a veteran teacher. This is something that an evaluator lacking content knowledge and instructional experience relevant to her instructional practice cannot provide to her. She further calls into question the effectiveness of principals serving as teacher evaluators who have had no prior experience whatsoever as classroom teachers. Despite their lack of experience or perspective relevant to the role, they were still assigned to the task of evaluating teachers for the apparent purpose of supporting their professional development and growth. The result is concern on the part of teachers for evaluation validity both regards to their individual evaluations and also with regards to the entire evaluation process. This also supports questions regarding the validity of evaluations based upon principal observations that have been raised within research into this subject (Croft et al., 2016; Harris et al, 2014; Firestone, 2014; Peterson, 1987; Smagorinsky, 2014).

Angela also shares her perspective about the minimum instructional experience that administrators are typically, but apparently are not always, required to have. This echoes Julia's sentiment on evaluator background experience:

> [T]hree years in the classroom and they're still developing in the profession and have no idea what they're doing. Because they didn't spend long enough. So we're going to say you can have a minimum of three years in the classroom. You haven't… you're just mastering that skill and I'm going to take you out and put you someplace else. [T]here's no correlation between an effective teacher and an effective administrator. They're two completely different jobs. (7a. 453-455, 460-464)

Angela is sharply critical of the lack of systemic awareness regarding the preparedness of administrators to serve as instructional leaders. She, like Julia, emphasizes the lack of instructional knowledge and expertise they obtain within the classroom as teachers, prior to moving into such roles, suggesting that three years is nowhere near enough time to develop instructional mastery. This is, of course, a legitimate concern concerning evaluation validity as previously established (Firestone & Pennell, 1993). Building on this particular point, however, she challenges the underlying correlation between effective instructional practice and effective supervision of teachers, a concern that also has support from research findings challenging the correlation between principal observations of teachers and teachers' quality of instructional performance (Harris et al., 2014).

The research and teachers' perspectives explored thus far suggests that many principals are poorly prepared to serve as effective teacher evaluators based upon their own instructional background. Further, they are considerably restrained within their role concerning their ability to improve their preparation. A question emerges here regarding possible solutions for principals who find themselves in this situation, and who also desire to rectify it.

Joyce, a veteran music teacher, also has previous experience serving as a building principal. This has given her a unique insight into the role and the responsibilities that principals

carry on their shoulders. Joyce shares her knowledge of the primary objective for building principals during a previous era when teacher evaluation was not nearly as high-ranking of a priority in education:

> You're a stranger. Even though you're in the building… your purpose then was not to be an educational leader. Your purpose then was to be a personnel manager. So [my principal] was managing. He was managing. (3a. 105-108).

Early in Joyce's career as an educator, principals were indeed mostly expected to be school managers more than anything else. I can certainly recall even now from my own personal experiences as student, seeing my principal in his office holding disciplinary hearings throughout the day, and monitoring student behavior in the lunchroom as well as during recess when he was anywhere else in the building or on school grounds. Rarely did he appear in a teacher's classroom, and the few times he did step into the room, we immediately knew that someone was in deep trouble and that they were about to embark on a nervous walk back to his office. The concept of the principal observing teachers and serving as leaders of instructional practice was all but nonexistent in my elementary school, to the best of my knowledge, much as it was something that Joyce did not see her principal engaged in when she first began teaching. She shares a notable exception, however, in the form of a building principal she worked under named "Paula":

> Paula… was the first administrator… who really began to construct the purpose of evaluation. She'd say I'm looking for this, this, and this and I want to know how you're tying this in with instruction. I started teaching in '78, so we're talking about very, very early in my career. She wanted evidence to… back up, to say that you've done what you

said you did. That was… a very new agenda. And the woman was brilliant. She was very much so ahead of her time. (3a. 126-134, 136)

Again, the common expectation of principals then, and for many years, was to manage and maintain order and discipline within their buildings. Indeed, throughout the modern history of public education, up until the very recent period, principals were not expected to closely monitor, coach, and evaluate the instructional practices of the teachers within their buildings (Sergiovanni & Starratt, 2007).

Not only this, but as Joyce further shares, they simply were not able to satisfactorily balance building management and teacher evaluations, particularly within schools based in urban settings:

> I was a principal at [Name Removed] School… And the challenges that are going on in that particular building now were the very same challenges that I had to go on then. So … here's your question. Here's your pedagogical question as a building principal. What do I do? Do I make sure that we have soap in the bathroom and toilet paper on the rolls, because… that's part of your job. Or do I spend 95.5 percent of my time in the classroom where you want to be with instruction when you're in an urban building… that is [all the way live]. [CLAPPING NOISE]. Those are questions that you come up with every day and it never fails. (3a. 245-247, 253-260)

Joyce's question is a critical one. How can principals even begin to work towards completing meaningful, relevant evaluations and also achieve validity for the evaluation process if they have so many other demands on their time and resources to which they must also give their attention?

Despite having to grapple with this challenging situation, however, she suggests from her own perspective that she would have enjoyed spending more time being an instructional leader—

working more closely with the teachers under her supervision and supporting their professional development—while she served as a building principal:

> [T]he part that you love the most is the thing that you get to do the least. Because… you're worried about is Mrs. Smith's class going to have a sub, or Jonathan's got a rash and there's nobody coming to pick him up and it's contagious, or Sally Mae has thrown up in the middle of the floor… and Ms. Smith over here is having a breakdown because she didn't do what she needed to do in the beginning and her kids are having [problems]… this goes on every single day. So you don't become the best evaluator because you don't get the chance to know your people. (3a. 261-272)

Clearly, as a building principal, Joyce wanted to work with and support her teachers. She expressed concern about her capacity to be able to effectively complete quality evaluations, evidencing a personal standard necessary among principals in order for them to achieve evaluation validity (Papay, 2012).

She continues to express her concerns about the present expectations for principals to conduct evaluations for an entire teaching staff within one building. Pointing to her current building and principal as examples, she also worries about the principal's capacity to do so in a manner that has quality and validity:

> [I]n a building like this… [Name Removed] has 52 people that he has to evaluate… every year eat least two times a year. There are a hundred and eighty-three days in the instructional calendar, a hundred and eighty-six days… that we are supposed to be on contract. A hundred and eighty-three days of instruction whereby he can see a teacher, not only an evaluator who's having hard times, going to see somebody on Monday and Friday. That dog ain't going to [hunt]. That just ain't going to happen. Although you

> could see a whole lot at the beginning and the end of the week… Where do you get the opportunity to build and grow your schema, your understanding of context, your understanding of content, your understanding… of people knowledge? When do you get to do that? When do you get to do that? (3a. 377-390)

Referring to her own professional background knowledge and experiences as a principal, Joyce suggests that her current principal is likely just as overwhelmed as she was, only even more so now because of the current policy emphasis requiring principals to prioritize their roles as teacher evaluators. Her insights into the principalship broadens the picture of evaluation validity and issues surrounding it to encompass a more detailed understanding of the challenges principals face beyond the teacher-centered perspectives that Julia, Tara, and Angela express. Further, Joyce's insights present a fundamental challenge to the research on evaluation validity: How can evaluation validity be a realistic aim if principals do not have adequate time and space to conduct evaluations?

Further, Joyce also described the potential for problems that can emerge regarding the validity of evaluations when principals already challenged for time among their other responsibilities conduct them with bias:

> Any… system where an observer is recording evidence and then applying the evidence to the rubric, the door is wide open for cherry picking. If I were an administrator, could come into a teacher's classroom. Oh, I really like this teacher. I have a lot of respect for the teacher. I already know that what they're doing is great and so I'm going to select the pieces of evidence that will lend themselves to an accomplished rating. Oh, this teacher over here, I know they have discipline problems in their class. I'm walking in the door highly critical. I'm going to focus on the pieces of evidence that reinforce my existing

> convictions. So, there's always the possibility of confirmation bias anytime an administrator gets to pick evidence. (3b. 160-173)

Thus, not only are principals challenged to conduct a high number of evaluations as Joyce describes, but in the context of balancing all of their other responsibilities, duties, and demands, principal bias may weigh substantially over the outcome of many teacher evaluations. It is not at all difficult to see how this could happen when principals are simply trying to meet deadlines and demands on multiple fronts. Bias can essentially become a short cut in effect, particularly when a principal already has an established working relationship with a teacher and is familiar with their instructional performance and overall professional track record. Confirmation bias, as Joyce describes it, is a commonplace phenomenon. This, however, supports research findings suggesting the need for evaluations that contain multiple measures in order to serve as a check against bias in any one measure (Baker et al., 2010; Darling-Hammond et al., 2012).

In addition, Angela described the perspective one administrator shared with a colleague of hers regarding the administrator-teacher relationship in general:

> One of my friends was an administrator and she was taking administrative classes at [Name Removed] and he was a speaker. And the first thing… she said I'll never forget this… She said that he came in and said don't trust your teachers. Have students that you can use as [spies] on teachers. (7b. 628-633)

Even Angela herself was taken aback by the boldness of such a comment, but it nonetheless left a lasting impression on her as well as the colleague who shared it with her. Such a comment coming from a school administrator in a position to train other administrators certainly calls into question the validity of evaluations by principals who hold such a view of their teachers. Similarly, it should also raise concerns about evaluation validity when a principal has developed

a close, positive working relationship with their teachers. This finding once again supports research which recommend multiple measure evaluations in order to serve as a check against bias in any one measure, in this case principal evaluations of teachers (Baker et al., 2010; Darling-Hammond et al., 2012)

Karen, a veteran elementary teacher, contributes an additional perspective towards our understanding of evaluation validity and the issues surrounding it. In particular, she raises concerns about the declining quality of the evaluations that she has received. She shares the apparent short cut techniques that administrators use to meet their completion deadlines:

> [I]t seems like… most recently, they have been so overwhelmed that you can tell they're cutting and pasting and they're doing all this and you're reading this… that wasn't the lesson… all those things, those aren't applicable to me. That must be somebody else. And it just doesn't seem to be valid like it once was. It… seemed more valid early on, and now it's just… an exercise in trying to get through it… because the demands are really high. High for them, high for us. And it's just …it's just too much for one individual to do. I mean for the principal, for that principal to do. They're just totally overwhelmed and bogged down that they can't do it in a way that's truly helpful anymore. (5a. 143-154)

Karen's concerns echo Joyce's regarding the responsibilities principals must deal with in addition to conducting teacher evaluations, and the difficult situation this places them in when they cannot do everything well. As Karen suggests through her own experiences as a teacher being evaluated that the end result of the process often reflects a considerable amount of missed detail such as in a "cut and paste" job. Literally, principals are cutting and pasting elements, and even boilerplate language, from one teacher's evaluation to others, in order to streamline and expedite the process.

This, of course, does not support evaluation validity when the specific processes and procedures that evaluators are using are poorly applied (Papay, 2012). Extending this point, Karen expresses additional concern with regards to the quality of evaluations she has received:

> [T]here's things in there that are not even about you and there are things that are not… you know doggone well that that was not your lesson and there's comments in there that are not… that shouldn't be in there. But at this point in time, it's like they didn't worry about it and you're going forward and they're not putting any… recommendations. (5a. 156-161)

Based upon the account she provides, the integrity of the evaluation process itself is very much in question. If principals serving as teacher evaluators are not taking the overall process seriously, and in fact, have predetermined how they are going to evaluate and rate teachers' performance, what is the point of the entire process? Such carelessness fosters bias within evaluations, which in turn gives credence to teachers' concerns about favoritism (Firestone & Pennell, 1993).

Along with Joyce and Karen, Sarah, a veteran elementary school teacher expresses her concerns regarding the demands that teacher evaluation in its present form places on principals. She suggests how and why this is affecting the quality and validity of the evaluations they complete:

> [T]ruly [in] my opinion, [teacher evaluation] puts too much stress on the principals. They… don't have time to do the other jobs that they are responsible to do… I know my principal who is new to our building this year, she… came from out of state, and its gets down to crunch time and—this wasn't really her fault—but it got down to crunch time and the people she did at the very end, I don't feel like they got the same thought as I did.

> I signed up early. I like to get that kind of stuff over with. And there were people at that last week… because you have to have the first cycle done by a certain time… and so she had to do all these people at the very end, and… it's almost impossible for you to put as much time in with those people (6a. 319-331)

She suggests that principals are simply not able to devote as much time into observing and providing feedback to each teacher as the evaluator role demands of them. Sarah's principal is showing evidence of stress—of being at or beyond her maximum capacity—by providing significantly more attention and feedback to the teachers she meets with early on during the evaluation cycle than she provides to the teachers she observes and meets with closer to the end of the same cycle.

Recognizing this situation, and also wanting to manage her own stress levels as a teacher—which being evaluated can certainly elevate—Sarah volunteers to have her principal evaluate her as soon as possible. The benefit of doing so is that she likely receives a better-quality evaluation that has more validity to her as a snapshot of her performance. Her principal has put more time and effort into completing it (Papay, 2012). Her concerns regarding her principal's propensity to sacrifice the quality and validity of the process in order to get it done, bear similarities to Karen's account of principals cutting and pasting language between evaluations to expedite the process, but at the same time reducing its validity.

Group level assessment activity. Two concerns were raised by teachers participating in the GLA activity with regards to Evaluation Validity during this portion of my research study. The first concern focused on Student Learning Objectives, or SLOs, which constitutes 50-percent of teachers' summative rating in the school district where my research took place (the other half of this being principals' rating of teachers' performance based upon observations). The second

concern centered on the overall evaluation process. What both concerns have in common is the role of students in the evaluation process. The first respondent discussed SLOs and student accountability, or rather, the lack thereof:

> The pre- and post- SLOs are ineffective, not in how they are designed, but in their evaluation. The teacher is held accountable, while the student has no accountability. SLOs should be weighted and graded as exams, both affecting the student and the teacher. The students [would be] given 100 percent if they meet their growth, and 75 percent if some growth. I actually heard students laughing about purposely failing a test so they could get a teacher fired, and this would not affect their grade. In the business world, both employee and manager are evaluated for their progress.

What the first respondent describes is an entirely plausible and frightening scenario, one in which a group of students colludes to sabotage a teacher's SLO rating, and therefore the entire evaluation process, by intentionally failing a key assessment covered by the SLO. The potential ramifications for the teacher in such an instance can be severe, ranging from professional sanctions, to dismissal and loss of livelihood. This would be based on a non-valid summative evaluation rating based on student interference within the process. As the teacher respondent suggests, such a scenario is not commonplace in the business world, where employees, as the equivalent to students, and teachers, as the equivalent to managers, are both evaluated for their performance and progress achieved, thus holding both groups accountable. Indeed, in some instances today, managers and employees may interchangeably rate each-others' performance within their respective roles.

The second GLA activity respondent provided support for the first teacher's perspective with their own perspective:

> In all the stuff that I've read about evaluation, that's the one piece I have not read, and that is the misnomer from students. Students are hearing whether or not their teacher goes or stays based upon their performance. And therefore, they are given the power to misperform (sic) on the context that it will change the employment status of the teacher. That's deep. That's really something that I think when people came up with this they weren't thinking about. I've never read that before. If kids think that they are empowered to predict the destiny of their classroom teachers, whether they continue working or not, where is that piece within the paradigm about improving learning coming in? This is punitive against the teacher.

The second GLA activity respondent, like the first one, raises concerns about the validity of OTES, but also goes further to question the validity of all teacher evaluation systems in which there is no control for student accountability. The implications of such an evaluation system design flaw are seemingly obvious inasmuch that students can willfully invalidate evaluation processes and experience no repercussions for doing so themselves, all the while potentially causing great harm to their teachers. This particular issue of validity should be of great concern to educators, as well as evaluation system designers. The validity of using assessments to summatively evaluate teachers is further called into question by research findings suggesting that such measures cannot fully control for all variables, particularly concerning student demographics, home environment (Baker et al, 2010) and the school environment itself (Darling-Hammond et al., 2012).

Self-reflection. From a personal perspective, I cannot say that all of my evaluations have been valid from the standpoint of having an evaluator who had deep familiarity with my content area. Further, I cannot attest that my evaluators maintained extensive and fresh operational

knowledge of what it takes to facilitate and maintain a productive classroom learning environment. During my early years as a teacher, evaluations were brief "one and done" conferences that took place annually with the school principal, and that may or may not have also involved a formal observation of teachers' instruction. In my case, I did not have a single principal observation of my instructional practice during my early career. It was not until my second year of teaching in my current school district that I had a formal principal observation. Prior to this, I had an assigned "probationary" evaluator for my first year. In each of these later instances, both my probationary evaluator and the school principal evaluating me my second year were licensed Social Studies teachers with extensive content and professional knowledge, as well as classroom experience. Meanwhile, my current evaluating principal has an extended teaching background, but is not licensed in my content area.

From my own standpoint, the most helpful—and valid—evaluations I have received during my career, from the standpoint of supporting my professional growth and development, have been from two individuals who have also taught in my subject area and who appeared to understand what it means to be a good Social Studies teacher. They could convey this understanding to me in a way that was both constructive and beneficial to my own professional growth and development. I believe having a common frame of reference is essential to the teacher-evaluator working relationship, and I have seen its value within my own professional practice, during those times when I was working with such an evaluator who knew my content area and could convey a practical understanding of it, and of teaching Social Studies at the Secondary level. This is the precise type of evaluation experience that I see as being valid, one that I believe Angela and Tara also suggest as being valid to them professionally, one that is

further supported by research findings (Firestone & Pennell, 1993). Angela and Tara each sought evaluations that were relevant to their practice and ongoing professional development.

Teachers' impressions of evaluation reliability: First and second phase interviews. I asked the teachers I interviewed about their conceptualization of reliable evaluations. Paul shares his perceptions of this issue:

> If you are getting solid feedback from an experienced educator… in an administrative position, I would say that the kind of solid feedback you can get from a good evaluation system can help you grow professionally. It can help keep you on point. It can help remind you of things that maybe you've forgotten because you haven't been evaluated recently. But in the real world… I think amidst [OTES] and SLOs and the new common core test… I think a lot of teachers see it as okay, I'm already stretched very thin, now here's one more thing I've got to do. Even… if it's not a complicated thing… it's one more time sucker. (2a. 104-114)

His comments somewhat highlight the earlier concerns Julia and Angela expressed about administrator experience and preparation for the role of teacher evaluator, suggesting that an experienced, well-prepared administrator can indeed make the overall process valid, constructive and worthwhile for teachers (Firestone & Pennell, 1993). Also, they echo Sarah's concerns concerning teachers' own growing workload and the stress that they themselves must manage in order to meet all that is expected of them, which can also have an impact on evaluation outcomes (Papay, 2012).

Further, Paul directly addresses inter-rater reliability and the concerns his colleagues have expressed about the quality and validity of their evaluations:

> I can point to several different evaluators where if I have evaluator A, I know I'm going to coast. If I have evaluator B, it's going to be a fantastic growth experience. If I have evaluator C, well, maybe that person's already determined from the beginning well, nobody's going to be at that accomplished level. So everybody's going to be skilled or developing. And so… I think until you get that calibration down, it's going to be a very big deal who is my evaluator… and in an ideal situation, it shouldn't make a difference. Good teaching or bad teaching should be perceived the same way by any evaluator who's looking at the same rubric. (2a. 178-188)

Paul's comments provide a rationale for supporting and strengthening inter-rater reliability among evaluators. Doing so would support evaluation quality and teachers' perceptions thereof (Harris et al, 2014). His perspective on this subject, however, contrasts with Julia and Angela's expressed concerns about the instructional knowledge and experience that principals bring to the table as evaluators, and how well instructional background matches up with that of the teachers that they evaluate, which they believe that it should do so. In contrast, Paul believes that good instructional evaluations focus on identifying good teaching rather than content knowledge, negating the need for principals to be content knowledge masters within a particular area and to only assign themselves to specific teachers. As ideal and desirable as it would be for each teacher to have an evaluator who is an expert in their content area in addition to being an instructional expert (Firestone & Pennell, 1993), Paul's perspective emphasizes a strong preference for inter-rater reliability among evaluators, one that also acknowledges the time and resource constraints of administrators and teachers alike and supports more accurate findings (Papay, 2012).

A few other teachers contribute their views about how teacher evaluation might be improved from a reliability standpoint. Stephanie, an elementary music teacher offers her perspective:

> Well… I think just in wanting to be an effective teacher and also wanting to give students a well-rounded music education, I would hope that… in the future, the teacher evaluation system is one that can encourage teachers to be at the top of their game while also benefiting the students… There's so much more to… student learning than just… can you master these and these things and just spit them out the same way I told [them to do]. Because when you're asking students to do that, you're taking away a lot of their creative abilities. (1a. 520-528)

Stephanie, like Paul, Tara, and Angela, sees teacher evaluation primarily as an opportunity for professional growth and development and wants it to function well towards this end. What she offers in contrast to her peers, however, is insight into how teacher evaluation in its present form negatively affects students, particularly concerning their creativity. She suggests that the process is so focused on assessing teachers for specific standardized outcomes that it detracts from their ability to support creativity, which is a critical part of learning and growth and consistent with the educational values of Dewey (1938). Stephanie's perspective suggests a need for a constructive approach to evaluation that supports teachers' professional development, but also one that incorporates a broader view of the learning process and teachers' contributions to it, consistent with the suggestions of Musella (1970), who recommended granting teachers the ability to self-evaluate while also receiving principal input and feedback.

Along similar lines, Joyce, a music teacher as well, offers her perspective concerning teacher evaluation in its present form:

> I think in general, John… it seemed that it's supposed to be a help but it's not. [I]t's a performance. It's a dog and pony show. I still don't feel as though that we've done enough work bridging the gap between the evaluator and the teacher to really, clearly know what happens on a day-to-day basis within that classroom. (3a. 232-239)

Joyce suggests that teacher evaluation in its present form has not lived up to its original promise of helping teachers and schools to improve their performance. Worse, it has merely become a burden for administrators and teachers to get it over and done with by any means necessary, and without consideration for quality. She further suggests that there exists a major opportunity for strengthening the evaluation process in ways that make it far more relevant to teachers and less burdensome for principals. This could be made possible through mutual collaboration between teachers and principals, as Musella (1970) suggests, or otherwise by expanding evaluations to incorporate multiple measures, as suggested by Croft et al. (2016), Peterson (1987), and Smagorinsky (2014).

Julia, however, expands upon her previous observations as well as Angela's concerning the importance of administrator experience and content area expertise with regards to teacher evaluations:

> I don't know that you can fill your administrative team with people that teach non-tested subjects. I don't think there's anything wrong with a principal that was a music teacher or a health teacher or a gym teacher or a Driver's Ed teacher. I think those people are all important. But I think until you understand the pressures of being a tested… and I don't just mean SLOs, because now we're all tested teachers… I mean until you know what it's like, until you have the state shove a test down your throat and throw standards behind that, English, math, science, social studies, I think that those are the people that need to

> be at the helm, especially when you're evaluating a poor teacher because it's a very different kind of pressure when you don't do well. (4a. 1072-1086)

Essentially, Julia suggests that school districts should consider whether administrator candidates have experience as teachers in subject areas measured by state standardized-tests before bringing them onboard and placing them in charge of evaluating teachers in tested content areas. This may even be more important in her view than matching up evaluators with teachers based upon specific content knowledge alignments. Expanding this suggestion to incorporate her earlier comments, there could be a priority selection process, in which administrators are matched up with teachers their evaluators based upon administrator experience teaching in a tested subject, followed by content area alignment.

Several teachers suggested that teacher evaluations would benefit from inter-rater reliability improvements, the lack of which makes evaluations more subjective. Stephanie contributes her perspective:

> Well, there is some subjectivity within… the evaluation process. The one evaluator can come in and look for something very specific. For example, smart goals or where you have to have percentages of the students achieving a goal. And of course, proof that each student is achieving the goal. But even within that… a student earning a three or showing that… they're at a skill level at their grade level… someone else come in and argue that it's …they're approaching the grade level and not quite at the grade level. So if that's the case, then… the data will look different [from different evaluators' perspectives] as far as what percent of the students in that class reach the goal or didn't reach the goal, or mastered the goal. (1b. 144-158)

Stephanie describes the perfect example of a scenario that undermines evaluation validity and the amount of trust that teachers can place within the evaluation process, if not the role of administrators within the process. When administrators see the same outcome differently, the validity of the evaluation process suffers (Harris et al., 2014).

Contributing to this conceptualization of the inter-rater reliability as a deficiency within the current approach to teacher evaluation, Sarah offers her own insights, from the perspective of having had two principals in as many years:

> So really… I've had two different evaluators now. I had [Name Removed] last time and then this was a… new principal this year. And so both experiences… were fine. I didn't have any issues and they were able to give some insight into the lesson. So in terms of validity, I know it can be subjective in terms of who's evaluating whom, different… even though it's not supposed to be that subjective because I went through the training for it. It still is. I mean just talking to people and what principals look at and the evidence that they look at. So it still is subjective. (6b. 61-73)

Sarah suggests that she did not receive a negative evaluation or have any specific problems with either of her evaluators. Nevertheless, she understands that evaluations do have the potential to be subjective. As an outstanding veteran teacher, she may never encounter an instance during the remainder of her career in which evaluation subjectivity might pose a problem for her with respect to her professional needs or ongoing development. An inexperienced teacher, however, could very well find themselves within the scenario that Stephanie described and come away from it with less trust, and possibility a weakened sense of self-efficacy (Harris et al., 2014).

Further building on Stephanie and Sarah's perceptions with regards to teacher evaluation and its inter-rater reliability, Julia offers a summary:

> [I]t continues to be flawed and there continue to be concerns raised in addition to the administrator fidelity. Even though they did the training, you get one administrator that would rate you one way, another administrator rates you another way. Teachers talk and they're not having [the] same experience. Even within a building. Never mind within a district or within the state. And concerns raised about validity of evaluations, like was this exactly 30 minutes. [D]id they actually do their walk-through? Because they don't come in and say this is your walk-through. And we've had concerns in our district and I think we'll continue to have them across the state. We've had examples of a post-conference [that] was missed. Well, what does that mean overall? Do you throw that whole evaluation out? And if you do, what happens to that teacher's rating? (4b. 72-87)

Julia suggests that not only is there potential for there to be a range of outcomes for a specific result, based upon which administrator observes the result, but that it is indeed happening because of inconsistent approaches towards facilitating teacher observations. This is, in fact, what weakens evaluation validity, as well as teacher trust in the process as a fair and constructive means of supporting their professional growth and development. Reducing the personal biases of evaluators with regards to teachers, while also isolating other factors that can lead to differences in judgment with regards to the quality of instruction observed is necessary, supports greater inter-rater reliability and the validity of teacher observations (Firestone & Pennell, 1993; Harris et al., 2014, Papay, 2012).

Considering the additional demand that teacher evaluation in its current form has placed upon administrators, and the problems that have resulted from it, from administrators seeking shortcuts, to them not being well-suited to provide teachers with quality evaluations that suit their professional development needs, one wonders if perhaps administrators should even be

conducting teacher evaluations? Joyce, who once served as a building principal, challenges this suggestion, however:

> I really believe that from my experience as a principal and as an… evaluator. I wish it was something within the paradigm whereby the educational leader/principal could be just that. The innovator, the encourager, the assistant, the support, the researcher… the manager of learning in a real pure way. I know that the lights have to come on in a school building and I know that there's got to be toilet paper and soap in the bathroom. I understand that. But that cannot be our evaluators/principals/educational leaders, that can't be their only job. (3b. 135-149)

Joyce expresses here that she feels being an instructional leader who evaluates teachers is an important, even necessary part of being a school administrator. From her perspective, one grounded by the experiences she has gained as a principal as well as a teacher, these roles are not mutually exclusive, nor can they be separated and regarded as such.

Extending the conversation on improving the validity of teacher evaluation to include teachers themselves, Tara offers the following perspective:

> You're putting a lot of burden on the evaluator to find evidence. …[I]f the educator doesn't bring the evidence and it wasn't necessarily seen, for example, there's no way for my evaluator to know what my community involvement is unless she's following me around. And if I don't bring a calendar that says this is when I went to a PTA meeting, this is the ballgame that I went to, then it's assumed that I have not done those things. …[I]t falls back on the educator once again to… provide that evidence… There will be evidence of this. So unless or until I produce that evidence, it's …the assumption is that it doesn't exist. (8b. 153-170)

Tara suggests here that teachers actually have a substantial role within the evaluation process, but that it is one requiring them to provide evidence of their performance and accomplishments to their evaluator. While evaluators do carry the burden to find evidence of teachers' performance within a given component of the rubric, teachers can assist them to some degree. This suggests that a space exists for collaboration between teachers and administrators, even if it is not a well-developed one. Were this to be expanded and play a more formal role within the evaluation process, it could make teacher evaluations more collaborative, constructive experiences overall for teachers as well as administrators (Musella, 1970).

Group level assessment activity. With regards to Evaluation Reliability, one teacher participant within the GLA activity shared a personal experience with respect to how the contrasting perspective of an evaluator affected the outcome of a candidacy for a position that they were seeking:

> I was going through a [Teacher Leader] certification process, worked in an innovative school that believed in project-based learning. we were a middle school, and we introduced projects… I had gone to the library and brought them books… And so created this... set of resources. It was collaborating with this social studies teacher. And the evaluator came in. And we had taught conflict management skills, all of this. And so the kids were picking some books and two students were arguing over a book. I didn't intervene. They resolved their argument peacefully between the two of them as we had taught them, and the class continued on. It wasn't a major disruption. It was, you know… they went back to working. And I got to the end of my lead teacher evaluation and was not granted lead teacher evaluation because the [Name Removed, evaluating] teacher who came into my [Name Removed] classroom saw something different than what I saw.

The teacher described a situation in which they and their school had taken proactive steps to prepare students to handle and resolve conflicts between themselves. Ideally, as they believed, such positive handling of a situation by students without the involvement of the teacher should be well-looked upon by anyone observing it. This, however, was not the case for this particular teacher as the evaluator judged them negatively for not intervening, even when there was no obvious need to do so, because the students had successfully resolved the issue themselves. As a result, the teacher was denied the position they were seeking, all because of the summative judgment of an evaluator who did not grasp what was actually taking place in one particular instance, or perhaps chose not to do so. The end result is a teacher who was seemingly punished for being proactive in nurturing personal responsibility in the form of conflict resolution skills among their students. This underscores the concerns about the reliability of evaluations suggested by Sarah, Stephanie, and Julia. Further, it provides support for the suggestion made by Croft et al. (2016), Peterson (1987), and Smagorinsky (2014), regarding the need for evaluations that contain multiple measures in order to gather a more accurate picture regarding a teacher's performance and their needs.

Self-reflection. From my own personal perspective, I developed my understanding of inter-rater reliability during my experience as an Instructional Supervisor at a previous school for which I worked. In this capacity, I was charged with evaluating teachers, much as a principal would, by conducting observations of their instructional practice, and then conferencing with them afterwards to provide them with feedback on their performance, as well as suggestions for improvement. Periodically, Instructional Supervisors would meet for our own professional development, which often included sessions in which we would all "practice evaluate" a teacher whose lesson on a given day had been recorded.

During these inter-rater reliability calibration sessions, the Instructional Supervisors would observe the recorded lesson and take notes of what we saw. We would then pull out copies of the evaluation system rubric which we would then use independently to check our notes and determine the quality of the instruction that we observed according to the available performance categories (which were similar to OTES' "Accomplished," "Skilled," "Developing," and "Ineffective" ratings). After a process that generally took anywhere from 30 to 60 minutes depending upon the length of the lesson and time we needed as individuals to review our notes and match them with the rubric categories, we would then convene as a group and discuss our ratings with each other. Generally, we were either all in agreement or otherwise pretty close to each other regarding the categorical ratings and the overall summative rating we assigned to the teacher for the lesson we observed. The idea behind participating in this exercise on a regular basis, was as our director explained it, to establish inter-rater reliability between all of us, so that we could be consistent possible as evaluators across the board, no matter what classroom we observed, even if efforts to achieve such consistency would always be ongoing, according to him.

Granted, the director still assigned us to evaluate teachers whose content area of practice matched up with our own whenever possible. Nevertheless, I took away from my entire experience as an Instructional Supervisor both the need to be a content area expert as an evaluator, as well as consistent in rating teacher performance based upon my observations and evaluation rubric alignment. This, to me, is what constitutes evaluation reliability, which in turn also supports evaluation validity. It further recognizes that achieving inter-rater reliability is possible to some degree, and while it may not ever be perfect, it can at least partially address the concerns about variations raised by Papay (2012).

Summary. To summarize, the theme of Evaluation Validity and Reliability encompasses teachers' perspectives regarding the validity and reliability of evaluations they receive from their principals. With regards to validity, teachers expressed concerns about the role of the principal as evaluator, which include the demands placed upon them to facilitate evaluations as well as run the school buildings to which they are assigned. Many teachers expressed concern about principals serving as evaluators and their relative state of removal from the classroom, as well as their lack of extended successful experience as teachers themselves within the classroom. Others described the pressures and challenges facing their principals and suggested how these detract from the quality of the evaluations that they receive. One particular problem this appears to create, as several teachers suggested, is the incentive for principals to take short cuts when evaluating teachers. This may consist of principals relying on working knowledge of teachers' practices, as well as principals' conformational biases based upon positive or negative relationships with specific teachers. Two teachers also raised concerns with regards to the potentially damaging role that unaccountable students may play within evaluations, raising additional questions with regards to validity.

Evaluation reliability was also a concern that teachers expressed, noting that each evaluator observes instruction differently. This apparent subjectivity leads teachers to mistrust evaluations, particularly when and where it can make a difference regarding their retention and promotion as a GLA activity participant noted. In general, teachers did seem supportive of being evaluated provided that such evaluations can be performed in ways that are both valid and reliable. To this end, consistent with these observations from teachers, the research suggests evaluations that are comprehensive, making use of multiple measures (Croft et al., 2016; Peterson, 1987; & Smagorinsky, 2014), and that are more participatory from teachers' standpoint

and formative, rather than summative (Croft et al., 2016; Ellis, 1986; Harris et al., 2014; Musella, 1970; Smagorinsky, 2014). What becomes clear from my discussions with teachers, as well as through my own self-reflection, is that great care should be taken to make evaluations as valid and reliable as possible, and that caution is warranted when this is not possible, a fact supported by multiple studies (Baker et al., 2010; Darling-Hammond et al., 2012; Firestone, 2014; Firestone & Pennell, 2013; Papay, 2012).

The Role of Mentoring in the Context of Evaluation

The third and final theme, The Role of Mentoring in the Context of Evaluation, encompasses the concept of professional mentoring and similar coaching-style supports that teachers receive in the classroom for the purpose of their professional growth, development, and retention. It also addresses the role and actions of specific and non-specific individuals that serve in various professional capacities as teacher mentors, supporting the development and retention of teachers. Most importantly, however, mentoring is increasingly relevant to teacher evaluation based upon policy trends and research (Danielson, 2017; Marzano, 2017) increasingly concerned with teacher professional development and school improvement. Further, research has established mentoring as an effective means to facilitate teacher growth and improvement (Feiman-Nemser, 1998), hence its inclusion as a theme within a research study focused on teachers' perspectives on evaluation and potential methods for its improvement.

In education, beginning teachers often receive mentors. Mentoring may happen informally, by way of a supportive, engaging learning community of teachers eager to welcome new members into its fold. It can also happen simply by way of a veteran teacher either across or down the hall who frequently checks in on new arrivals between classes during the first several weeks of school, bringing a handful of supplies and a mouthful of survival strategies and

encouragement. Mentoring, which is also occasionally known as coaching, can also take place between established peer teachers when one of them needs informal support in a particular area, particularly one that is part of the formal evaluation process (Marzano, 2017; Marzano & Simms; 2012). Further, mentoring can be comprehensive and proactive. Feiman-Nemser (1998) argued for an intentional *educative mentoring* embedded in practice that would lend itself to teachers' professional development. Such an approach would consist of "observation, co-planning, co-teaching, joint inquiry, critical conversation and reflection" (p. 73)

Increasingly, mentoring occurs in a formal manner, as is currently the case for new teachers in Ohio who hold provisional licenses through the Resident Educator program and must complete a multi-year mentored program in order to obtain a full professional license. This process also incorporates evaluation of professional performance, which teachers must pass in order to receive their full professional license. Teachers enrolled in this and similar programs are assigned mentors who work with and also walk them through the process until they have completed it. Mentoring can be regarded as a form of professional development that is a critical link to teachers' success within professional evaluations (Danielson, 1999; The Danielson Group, 2017).

First and second phase interviews. I spoke with several of the teachers about their own experiences with mentoring and its influence on their professional performance and development. Although a direct connection with teacher evaluation was not always evident, their discussions of mentoring nevertheless provided insight into a form of support that is increasingly relevant to it.

Stephanie, an elementary music teacher, describes the mentoring she received during her first year of teaching, which took place within a program her school offered similar to the Ohio Resident Educator program:

> I was… fortunate my first year to have a mentor. That was… required at that time. And that helped me a great deal. I was very fortunate that it was also a veteran music teacher. So it was in my own content area and someone who had been… in this district for quite some time. Yes, that person had been… in the same building that I was teaching at but was no longer in that building… they were in a different building my first year but they had been in the building that I was at, if that makes any sense. [LAUGHTER]. (1a. 86-90, 92-93, 95-97)

Stephanie expresses gratitude here for having a mentor as a new teacher. She also seemed to find encouragement within the fact that her mentor was an experienced veteran music teacher with previous experience working within her school, presumably also possessing knowledge of students and the school community she is to serve. She goes on to describe her first year of teaching, as well as her perspective on teachers' first year in general:

> I think the first year of teaching is always the hardest. It was for me and most people gave me that advice going into it. Of course, I thought I was going to be different… [LAUGHTER]… well, surely not me, I'm going to be amazing first thing out… Little did I know that… no, it indeed turned out to be my… toughest year just because I wanted to appear confident and wanted to appear that I knew exactly how things were going to work even though because it was my first year I honestly didn't know. (1a. 112-120)

Here Stephanie acknowledges that while she was initially confident upon entering the classroom for the first time, during her first year, she soon found that teaching was far more challenging

and difficult than she knew. This acknowledgment and the reflection that followed it reinforce the value she placed on the mentor relationship she previously described:

> I think part of becoming acclimated to the teaching profession and education profession is just saying I can do that and then figuring out how to do it… from day-to-day. But I think the willingness… is just as or more important than actually being able to be competent [in] every single thing you're asked to do in the first year… [LAUGHTER]. (1a. 120-126)

Going back to the earlier point in our conversation, it was evident to me that having a mentor helped Stephanie to not only make it through her first year, but also set her on a successful trajectory within her teaching career. Because she was able to see where she was heading, and had a tangible model of what her own successful future could be, she was able to persevere and overcome the challenges and obstacles that many first-year teachers face, a positive outcome that mentoring makes more likely to happen, and is supported by research (Holloway, 2001; Ingersoll & Strong; Smith & Ingersoll, 2004).

Julia, a mid-career social studies teacher, describes her own mentoring experience, an unofficial program provided by the district:

> I have not gone through the Resident Educator process. I went through mentoring way back before it was [formal], and it was just a [district] program. My department head [Name Removed] … he wasn't a department head at the time… he was my mentor. But having him helped me grow but it wasn't official. There was no rubric. There was no paper. It was like hey, okay, your organization systems are good, your discipline systems are good. Here's an idea to manage tardies and here's an idea to manage entering grades or… here's a spreadsheet I use to enter my grades so I don't have to do it all by hand,

> because when I first started we were doing it in a grade book… the old-fashioned way. (4a. 603-616)

Although they participated in different mentoring programs at different times, Julia, like Stephanie, also seemed to benefit from the mentoring program she participated in, one that the school district facilitated during her time of entry into the district before Resident Educator was implemented statewide by the Ohio Department of Education. Like Stephanie, Julia expressed gratitude both for the program and for her mentor in aiding her growth as well as setting her on the path to becoming a successful teacher, a positive outcome of mentoring that supports the retention and successful professional development of teachers, one that is further supported by research (Holloway, 2001; Ingersoll & Strong; Smith & Ingersoll, 2004).

Julia was also familiar with the Resident Educator program as it has been facilitated within her district, and was able to offer her perspective on it:

> [O]ur first and second year teachers get a mentor that has to be mentor trained. I'm not mentor trained so I don't know all that's involved in it, but I know that part of what they have to do is they have to sit and meet and they have to observe them and coach them on management and on academics… You're supposed to ideally have somebody that's in the same genre as you so usually a history with history, or maybe English and history. But it shouldn't be like a PE teacher and science teacher. Typically, it's somebody that shares a discipline with you. (4a. 725-740)

As with the less formal mentoring program that she received support from when she was a first-year teacher, new teachers within the district, including Stephanie, and also throughout the state, currently receive extensive support through the Resident Educator Program. This support is intended to help them succeed and thrive as teachers, the essential purpose of mentoring

(Danielson, 1999; Marzano, 2017). In addition, the Resident Educator program helps districts address issues that, when unaddressed, typically lead to teacher turnover. Such issues include the isolation that teachers experience, as well as the frustration that can result from newer teachers realizing they have limited authority to address the concerns that newer teachers experience within school organizations (Johnson & Birkeland, 2003). It is clear then that having a mentor can make all the difference for a new teacher by supporting them through the process of adjusting to their new role and its demands, a fact that is also supported by multiple studies (Danielson, 1999; Holloway, 2001; Ingersoll & Strong, 2011; Smith & Ingersoll, 2004).

It is noteworthy that mentoring has become an integral part of the teacher evaluation process, particularly for new teachers participating within the Ohio Resident Educator Program. As she reflected on her own first year, Stephanie reaffirms her gratitude for having a mentor, and also suggests that mentors are essential for new teachers to succeed under the new evaluation system:

> I was very glad to have a mentor my first year teaching. And I know [the] state requires new teachers to have a mentor… at the current time. But that… [new] evaluation is actually a bit more intense than mine was. Mine was a one-year program with a mentor, and I believe now [the resident] educator program [is]… four years long. So that seems to me to be pretty intense. [LAUGHTER]. But yeah, I would highly recommend continuing the mentoring program for new teachers, especially having a mentor in your own subject matter, whatever that happens to be. (1b. 61-67, 69-73)

Stephanie is clearly in favor of mentoring as a means of helping teachers succeed, not just in becoming acclimated to their profession, but also with regards to their evaluations. While mentors within the Ohio Resident Educator Program do not serve as teacher evaluators within

OTES, they can support Resident Educator teachers with regards to understanding the qualities of successful teaching that their school administrator will assess them on within OTES (Ohio Department of Education, 2017a). In this regard, they are well-positioned to support new teachers with their evaluations, contributing to their retention and professional development within multiple areas that OTES addresses, and also with which their evaluating administrator may be concerned, which Johnson and Birkeland (2003) highlighted.

Two teachers discussed an alternate form of mentoring specifically within the context of teacher evaluation, offering contrasting perspectives. Karen, a veteran elementary school teacher, suggests that mentoring could take place through collaborative practice, or peer evaluations, as an alternative to the current teacher evaluation system:

> I would change it to make it more that teachers would have more input and the way they have before where teachers evaluating other teachers, helping them and [COUGHING] collaboration. Those are important things because we learn from one another and we… maybe we could have helped [the] teachers not walk away after two years of teaching.

What Karen describes here is a form of mentoring in which teachers observe each other's instructional practices and provide feedback for the benefit of their individual and collective professional growth and development. It is a potential holistic alternative to the current evaluation system in which principals observe and rate teachers on their performance. It further addresses concerns that have been expressed about the validity and reliability of the current evaluation system by teachers as well as by researchers (Croft et al., 2016; Ellis, 1986; Peterson, 1987; Smagorinsky, 2014) by involving multiple raters in a formative, rather than summative, collegial process.

Sarah, a veteran elementary school teacher, offers an alternative perspective on peer evaluations, however:

> I think we've talked about here at [Name Removed] years ago having [maybe] the coaches come in …the first round. Like if we did some type of peer evaluations, having …one of the coaches come in, the learning coaches. And I would be okay with that. But …I'm personally not for having the person two doors down come in, just because …I've worked with teachers that …I wouldn't respect their opinion and other teachers I would. But I'm not a big fan of peer evaluators… Maybe a coach. [O]ur coaches have a good rapport with the teachers here and I would be okay with maybe a learning coach coming in. (6a. 48-56, 58-60)

In contrast with Karen, Sarah expresses reservation about having random peers serve as either her mentors or her evaluators. She states that she does not trust every teacher's instructional perspective based upon her prior experiences working with several individuals that cause her to question how what they have to offer could benefit her practice. At the same time, however, she expresses that she could conceive of having a teacher coach serve as both a mentor and an evaluator for her because they have clearly established themselves as instructional experts and understand what constitutes good teaching. In turn, they would provide constructive support and professional development for teachers, and further support their retention.

Sarah further elaborates on her rationale for preferring coaches as evaluating mentors over random peers serving within the same capacity:

> I would be okay with the coaches… I definitely would have been okay with them coming in because they have been in the district. I feel like they had a lot of knowledge to share. I just have trouble sometimes with people that you really maybe don't know very well or

> maybe don't have very much experience in the classroom and they're a coach and maybe come in and tell me something, and I've been teaching for 20-some years. So maybe if I had… if people had a choice… a list of maybe coaches that they could choose who they wanted to come in, I would be okay with that. But them just assigning somebody who... And maybe doesn't teach… or has never taught the grade they're going to come in and observe or evaluate. That type of thing [is a concern]. (6b. 80-81, 90-92, 94, 96-98, 102-115, 117-119)

Karen and Sarah offer contrasting perspectives on mentoring as a form of teacher evaluation. Karen favors mentoring in the form of peer evaluations, placing her trust in the perspectives of her peers and the collegial relationships that develop over time. Sarah, in contrast, expressed a lack of trust in the perspectives and opinions of some of her peers and suggested that she would not want to rely upon them to judge her performance. She did suggest, however, that she might possibly be okay with mentorship limited to the context of coaching, rather than as part of the formal evaluation process. Perhaps this indicates the need for choice and flexibility within the options that are available to teachers with regards to professional evaluation?

Karen meanwhile describes the benefit of informal collegiality as a form of mentoring, which took place during her first years as a teacher:

> I had a… new teacher meeting… every four weeks and then for three years following that I had not-so-new. That's a lot of training for… a new teacher. And was it tough going to those meetings? Yeah, but you learned… you were close-knit with all the teachers and you learned. You knew what was going on at [Name Removed] or [Name Removed] or [Name Removed] and you… [traded] things, you knew those people and there was a…

> networking with all those other young teachers… Around the district. And that… was really powerful [then] because we really learned a lot. (5a. 478-488, 492-502)

There is certainly something to be said for the informal collegiality as mentoring that Karen describes here, which, for her, was apparently was a very important part of her acclimation to teaching, and in her own words, may have been just as important as the meetings she attended during her first years, if not more important than them.

Overall and most important to point out here, however, is the fact that two teachers are suggesting how mentoring and evaluation may be combined in ways that constructively support their professional growth and development, while supporting the retention of their less experienced peers. These considerations, supported by research, suggest potential avenues for improving teacher evaluation through mentoring, and in turn, the professional development, growth, and retention of teachers (Borko, 2004; Johnson & Birkeland, 2003).

With regards to the role that principals maintain as teacher evaluators and how they might also serve as mentors, Stephanie offers her perspective with regards to what she would like to receive more support on from her evaluator:

> Tell me what I need to do to get the three out of four [rating] because we're conscientious.... A lot of our evaluation is okay, prove that you can do this. It's not just well, overall, it seems like they're a pretty good teacher. And so we have to state things that are part of our job that we just normally do that we don't even really think about it anymore because… we've never really been asked to show that we do it. Of course I call parents and document that. Of course I send emails or follow up with behaviors or speak to the principal about an incident that happened here or there… okay, I can do that. If that's the hoop I need to jump through to get there, that's fine. I just want to

> understand… [LAUGHTER] …what tricks you want me to do while going through the hoop. Stuff like that. Or at least to say I see that I got a three in this particular category. Can you explain to me why I wasn't a four or what I can do to make it a four? (1b. 256-257, 263-277, 279-281)

Stephanie suggests that not only does she expect her administrator to help her better understand what they want from her when providing her with her evaluation results, she is looking for her administrator to be a coach, or in essence, a mentor. A more traditional evaluator is someone who simply checks off boxes on a form and then reports what they observed to the person who they observed. A mentor, however, is someone who helps a person they observe grow and become more effective within their role by explaining what the data they have collected with regards to their performance means, and how they can use it to constructive effect (Smith & Ingersoll, 2004; Holloway, 2001).

Group level assessment activity. When discussing mentoring, two GLA activity participants shared their perspectives on its relevance to their growth, development, and success as early teachers. The first respondent was a participant within the Ohio Resident Educator program:

> As a new teacher, part of the resident educator program, I did have a mentor. She provided wonderful feedback, constructive criticism, and resources. As I developed, having a mentor is not so important, as it takes both mentor and mentee away from their classes. And once a teacher has taught for a while, the expectation would be for collaboration. So, I believe this depends on where a teacher is in their career.

The first teacher respondent apparently found their mentor to be helpful at first, particularly during their initial professional acclimation phase. It is interesting to note, however, that their

view of the mentor changed as the teacher gained more experience and apparently a greater degree of comfort with regards to their professional practice, viewing it as one more thing to do aside from actually teaching. Further, this teacher suggests that the mentoring relationship should evolve over time, becoming more collaborative and collegial, as well as flexible based upon need, a suggestion also supported in the context of teacher evaluation by Danielson (1999, 2017) as well as Marzano (2017).

The second GLA activity teacher respondent offered an alternate though similarly constructive view of mentoring, from the perspective of being a mentor teacher:

> To me, this is the key to any effective teacher growth and development. The validity of the mentor teacher relationship can't be more vital in a new teacher's development. As a mentor teacher for many years working with other new teachers, it helps my professional practice become strong. It does take a village to help develop one into a strong, vibrant, professional educator.

Several points stand out within the second teacher respondent's post. First, mentoring is essential to the growth and development, and arguably the success, of a new teacher, a point also suggested by several research studies (Danielson, 1999; Holloway, 2001; Ingersoll & Strong, 2011; Smith & Ingersoll, 2004). Second, the validity of the relationship between mentor and mentee is critical, perhaps just as much if not more so than the relationship between teachers and their evaluating principals. Third, the teacher expresses that they themselves have benefitted professionally within their own practice by being a mentor to new teachers, supporting a widely established concept within the profession that teaching is synonymous with learning. Finally, the teacher expresses that it takes a community to support the professional growth and development of a new educator who is seeking to become established within the profession, an idea also

suggested by Karen with regards to the informal mentoring that takes place among teachers and their peers.

To summarize, both participating GLA activity teachers establish that mentoring either is or has been an essential component of their professional practice. Both would also seem to agree that it is vital to the retention and growth of new teachers. Although they each offered contrasting perspectives regarding the relevancy of mentoring to them in their current career positions, with regards to each teacher's perspective, a case could be made for an approach to mentoring that evolves based upon the need of each individual teacher, one that also supports their growth and development with respect to the teacher evaluation process, and is further supported by research (Danielson, 2017, Marzano, 2017).

Self-reflection. As a new teacher, I did not have a formal mentor, and I entered the profession well before the advent of the Resident Educator program here in Ohio. At the same time, I did have professional colleagues who informally mentored me, showing me the ropes and offering me support, guidance, and encouragement when and where I needed it, which in all honesty was quite often. I survived to reach the point where I am today within my professional career because of these wonderful people, several of whom I still am in contact with, despite the fact that we no longer teach within the same school organization. From my own perspective, I understand the point Karen made regarding the importance of collegial relationships, as well as the perspective of the GLA activity teacher respondent who stated that "It does take a village to help develop one into a strong, vibrant, professional educator." I am a product of such a village, and it is clear to me from where I stand that every new teacher should enjoy a similar benefit.

At the same time, I can also see the benefit of more formal mentoring programs like Resident Educator, particularly with regards to preparing new teachers for the performance

expectations they face according to the current evaluation system and the overarching policy focus on accountability that serves as its rationale as well as a general backdrop for our times. Although experience has exempted me from Resident Educator, when I accepted a position within my current district, I was assigned a mentor teacher who also served as my evaluator during my first year. To say the least, this mentor was extremely demanding and critical of my performance from day one. Yet as challenging and even brutal as it was, I learned so much from this experience and my mentor that I feel I can handle pretty much anything that comes my way regarding a formal evaluation. Since that time, I have been well prepared for any professional evaluation. In all honesty, however, it was my most recent set of informal mentors that also helped me to survive the experience with my formal mentor. Therefore, I see the benefit and importance of both formal and informal mentoring for teachers, and also how these can benefit teachers in the context of professional evaluation.

Summary. Concerning the theme of the Role of Mentoring in the Context of Evaluation, beyond the Resident Educator program in Ohio, the evaluation process as designed within OTES does not incorporate mentoring as a formal process. Teachers may individually seek out mentoring from peer teachers for support in an area of need at any time, but they are not required to do so by the evaluation process itself once they have completed the Resident Educator process. Several of the teachers I interviewed, including those who shared their experiences through the GLA activity, described constructive mentoring experiences that took place early within their careers either formally or informally with peer teachers. These peer teachers supported the professional growth and success of these teachers by helping them acclimate to the teaching profession through the feedback, suggestions, and resources they provided. In turn, teachers suggested that mentoring could support them with the evaluation process, either by

supplementing it, or else supplanting it altogether through alternative forms such as peer evaluations.

In a basic sense, principals serving as evaluators could be considered mentors, much as peer teachers may receive this designation. Principals provide teachers with feedback on their observed performance, and they may also offer suggestions for refinement of practice. Where the comparison between peer teachers and principals as mentors breaks down is concerning the fact that principals must ultimately make summative assessments about teachers' performance. Such judgments can affect teachers' career outlook and employment status. In contrast, peer teachers serving as mentors have no power over their colleagues. Accordingly, they may only make formative assessments of their colleagues' performance. Teachers are thus more likely to confide in their peers when seeking out support for improving their performance. If, however, teacher evaluation as presently designed is intended to fulfill the role of improving teacher performance, the contrast between evaluator and mentor suggests an opportunity for transforming the process to better serve teachers.

Within all instances where mentoring takes place, but particularly with regards to new teachers, it is apparent that teachers benefit from mentoring as a means of support that also encourages their retention and professional growth (Feiman-Nemser, 1998; Holloway, 2001; Ingersoll & Strong, 2011). Without access to mentors that can provide them with professional support, instructional resources, room for growth, and reassurance that they can succeed within the teaching profession, it is clear that many new teachers can and do leave the profession well before they have the opportunity to become established in the field (Johnson & Birkeland, 2003; Smith & Ingersoll, 2004). At the same time, my interviews and GLA activity interactions with teachers, as well as my own self-reflection, collectively provide insight into the benefits new and

veteran teachers alike may potentially receive from collegial interactions, as well as through professional evaluations that are conducted more as informal, collaborate coaching sessions, rather than as more traditional processes of accountability (Marzano, 2017; Marzano & Simms, 2017).

Chapter Summary

Within this chapter, I discussed my findings for each of the three primary themes, Teacher Perceptions of Evaluation, Evaluation Validity and Reliability, and The Role of Mentoring in the Context of Evaluation. I used these three themes to analyze the teacher interview data I collected, as well as the GLA activity data collected, and my own personal reflection, seeking to understand their perspectives with regards to teacher evaluation in each of the three thematic categories. My findings within each theme suggest areas in which teacher evaluation in its present form either does or does not meet the needs of teachers, their schools, students, and administrators. Also, however, my analysis also revealed opportunities for improvement regarding teacher evaluation processes, and how they can best support the professional growth and development of teachers, as well as the needs of their students and schools.

With regards to Teacher Perceptions of Evaluation, teachers held a wide range of views with regards to professional evaluation and their general engagement with it, from negative to positive, as well as from stressful to irrelevant. The one perception teachers universally held in common, however, was that teacher evaluation has the potential to support their professional growth, development, and success as teachers. According to teachers, principals have the most pivotal role to play regarding both their perception of evaluation and its actual effectiveness or lack thereof with regards to meeting teachers needs and expectations. As such, principals must be

well-prepared to serve as evaluators that provide their teachers with constructive, relevant and meaningful feedback. Further, principals serving as evaluators must provide support that promotes their teachers' professional growth.

With regards to Evaluation Validity and Reliability, teachers delved more deeply into the role of the principal as evaluator, particularly regarding principals' instructional knowledge and experience. Not all teachers indicated principal matchup with teachers' content knowledge and experience as a concern. Some were more worried for their principals and the demanding expectations of the job that they are charged. In turn, some teachers suggested that principals facing such pressures may have added incentives to use shortcuts, which reduces the validity of the evaluation process. Teachers also discussed the reliability of evaluations, suggesting that unique differences between evaluators and their approaches to facilitating evaluations undermine the reliability of the process, resulting in mistrust.

With regards to The Role of Mentoring in the Context of Evaluation, an opportunity exists for the role of evaluator as presently designed to become less concerned with summative results, and much more focused on formative suggestions. The teachers I interviewed and interacted with through the GLA activity suggested how mentoring can support their retention, professional growth, development, and success within evaluations. All of the teachers who contributed to my understanding of this theme reported positive early mentoring experiences, both formal and informal, that helped them to successfully acclimate to the profession. Some teachers discussed the informal mentoring they received from colleagues who provided them with ongoing encouragement, support, and guidance during their formative years. Other teachers described experiences with formal mentors who more closely worked with, and in some instances evaluated them as part of their initial development and residency period. Finally, some

teachers argued for mentoring as a possible alternative approach to principal-facilitated evaluations of teachers.

Within the next chapter, I will discuss my recommendations with regards to each of these three themes, based upon my analysis and findings as discussed within this chapter. I will also discuss how these findings relate back to the themes as they are established within research literature. In addition, I will also discuss the implications of my findings for practice, limitations, implications for future research, and finally, the significance of this research study.

Chapter 5

Conclusions

The purpose of this research study has been to gain understanding of the perspective of K-12 public school classroom teachers regarding teacher evaluation. By conducting this research study, I have sought to answer the following three central questions:

- What are teachers' perceptions of teacher evaluation?
- What are teachers' concerns regarding the validity and reliability of teacher evaluation?
- What are teachers' views regarding the role of mentoring in the context of evaluation?

The theoretical basis of this research study and its design were informed by the feminist and critical theories serving as the basis of my theoretical stance. Feminist theory in particular seeks to grant voice to those who silenced or otherwise go unheard within a dominant paradigm in which objectivity and empathetic detachment are prioritized over intentional listening and care (Gilligan, 2003; Gilligan, 2011; Griffiths, 1995; Noddings, 1984). Critical theory meanwhile focuses on uncovering and challenging inequities within society and its social constructs that are detrimental to those impacted by them. Within an educational context, students, educators and society itself are affected (Anyon, 2011; Aronowitz, 2009; Dewey, 1938; Freire, 2000; Giroux, 2009).

Each of these theories served to guide my review of research literature into teacher evaluation, and further supported me with regards to assembling the research literature I reviewed into three thematic categories—Teacher Perceptions of Evaluation, Evaluation Validity and Reliability, and The Role of Mentoring in the Context of Evaluation. In turn, these themes

supported my analysis of the data I collected from teachers during Phase I and II Interviews as well as during the Phase III Group Level Assessment (GLA) activity I conducted, and it further served to guide the practitioner self-reflections I conducted during Phase IV this study.

Within the next section, I will summarize my findings corresponding with each of the three themes, Teacher Perceptions of Evaluation, Evaluation Validity and Reliability, and The Role of Mentoring in the Context of Evaluation. I will also discuss how I interpret the meaning of these themes in light of my theoretical stance, based in feminist and critical theory.

Summary of Major Findings

This section serves to summarize the major findings of my research study. Also, I use these findings, corresponding with each of my three research themes, to address my three research questions.

Research question #1: What are teachers' perceptions of teacher evaluation? The teachers I interviewed and interacted with during the GLA activity expressed a wide range of views with regards to their perceptions of professional evaluation and their general engagement with it. Their perceptions ranged from negative to positive, and also highlighted, in several instances, the stress they experienced during or in anticipation of evaluations. Some teachers, however, expressed feelings and concerns that the evaluation process was irrelevant to them, specifically with regards to their professional needs.

The one perception teachers universally held in common about evaluation, however, was that it holds the potential to support their professional growth and development, which all seemed to agree is one of the most fundamental and important aspects of being a professional educator, one that reflects a commonplace adherence to an ethic of profession (Shapiro & Stefkovich, 2011). The component of evaluations that most strongly influenced teachers' perceptions of the process, and of their individual experiences with it, is the role of the principal.

My interviews and interactions with many teachers suggested principals may have the most pivotal role to play regarding both their perception of evaluation and its actual effectiveness or lack thereof with regards to meeting teachers needs and expectations. The quality of the teacher-principal relationship, as these teachers suggested, may be best determined by the quality of feedback and support principals provide to teachers as part of the evaluation process.

Receiving such constructive feedback from principals is indeed beneficial to teachers, as I was able to attest through my own professional experience and perceptions of evaluation, both of which highlighted within my self-reflection for this theme. Collectively, my own self-reflection, combined with the research, and teachers' perspectives as expressed during interviews and the GLA activity highlight the particular importance of principal feedback and support provided to teachers through the evaluation process as a central factor with regards to teachers' perceptions of evaluation.

Research question #2: What are teachers' concerns regarding the validity and reliability of teacher evaluation? The data I collected from interviews, and the GLA activity interactions I had with teachers, with regards to this theme, provided deeper insights into the role of the principal as evaluator and their contribution to the validity of evaluations. This was particularly relevant with regards to principals' instructional knowledge and their professional experience, and how well or poorly these align with the instructional knowledge and experience of the teachers they evaluate. Several teachers indicated the importance of having a principal for an evaluator who also had content knowledge expertise, professional experience, and instructional knowledge beyond their teachers that they could share with them. This would result in evaluations being meaningful and relevant to these teachers, rather than merely a formality or

otherwise a rehash of what the teachers themselves already know. When principals as evaluators meet such criteria, these teachers see their evaluations as being valid.

Not all teachers indicated that principal matchup with teachers' content knowledge and experience, based upon their own professional background, was necessary for evaluation validity to be achieved. One such teacher suggested that the evaluation process should be generalizable to any teacher based upon observable evidence of effective or ineffective instruction, which in turn can be validated through the use of the evaluation system rubric. Other teachers, however, expressed concern regarding the extensive duties and workload principals carry in addition to the responsibility they have as teacher evaluators. These same teachers suggested that principals, when pressed to meet multiple demands and deadlines at once, may opt to cut corners by relying on biases when it comes to completing evaluations, thereby reducing if not completely eliminating the validity of the evaluation process. Additionally, some teachers raised concerns with regards to student accountability, or the lack thereof, in relation to formal assessments used to evaluate teacher performance, providing an example of unaccountable students who purposely performed poorly on one such assessment in order to damage the rating of the teacher associated with it.

Finally, several teachers reflected upon the reliability of evaluations, and suggested that the most relevant issue with regards to evaluation reliability is the unique differences that exist between evaluators and their approaches to facilitating evaluations. These teachers expressed that such differences undermine the reliability of the process, and some teachers among this group suggested that such circumstances result in teacher mistrust with regards to the teacher evaluation process. One teacher, however, suggested that it is possible to achieve inter-rater reliability among evaluators that adhere to common practices, supported by the official

evaluation rubric. This particular observation was supported by my own self-reflection in which I reflected upon the calibration practices I participated in as an Instructional Supervisor and their contributions to strengthening the inter-rater reliability of myself and my colleagues as teacher evaluators.

Collectively, my pertinent self-reflection, along with GLA activity and interview data collected, emphasize the importance of establishing and maintaining inter-rater reliability with regards to addressing teachers' concerns relevant to the reliability of evaluations. Also, the interview data from teachers and the GLA activity, in addition to my self-reflection, suggested evaluator match with teachers based upon commonly shared professional knowledge, content area knowledge, and experience, as well as adherence to common evaluation practices around the use of an observation rubric, are key concerns teachers have with regards to evaluation validity. Further, some teachers also expressed concern about principal workload and student accountability as potential threats to evaluation validity.

Research question #3: What are teachers' views regarding the role of mentoring in the context of evaluation? The teachers I interviewed, and also those I engaged with through the GLA activity process, collectively suggested how mentoring can serve to support teacher retention, as well as teachers' professional growth, development, and success with regards to evaluations. Notably, all of the teachers, including myself through my self-reflection, suggested that their early mentoring experiences were both constructive and critical with regards to supporting their acclimation to the teaching role, and their general successful careers as teachers, as determined by their retention and evaluation ratings.

Some teachers in particular highlighted and also emphasized the importance of the informal mentoring they received, discussing the value of the mentoring experiences they shared

with colleagues who offered them the support, guidance, and encouragement they needed to receive as novice teachers. Other teachers discussed their experiences with the formal mentors assigned to work with them through a developmental period, and in some instances, notably as part of the Ohio Resident Educator process, also evaluated their instructional practice as part of their licensure qualification requirement. Further, some teachers suggested that peer evaluations as a form of mentoring might serve as a legitimate potential alternative to principal-directed teacher evaluations, and these teachers considered the potential positives and negatives of a system in which their colleagues could observe their classroom practice and rate their performance.

In addition, some teachers offered the suggestion that mentoring has the capacity to grow and change along with the professional needs of teachers, including with regards to teachers' professional evaluation needs. Collectively, the interview and GLA activity data I collected, along with my self-reflection on the role of mentoring, suggests that mentoring does indeed play an essential role with regards to teacher professional growth and development, as well as with regards to supporting teachers

Implications for the School District

The following recommendations are based upon findings from this research study and, as such, are intended to be specific to the school district in which this research study was facilitated. These recommendations are not intended to be generalizable to any other school district within the State of Ohio or elsewhere within the United States or the world.

Providing additional balance and support for the principal as evaluator and instructional leader. As was highlighted within my research study analysis, principals serve many roles within their schools. While this has traditionally been the case, more recent policy

trends that have prioritized instructional accountability with regards to school and teacher performance has made evaluation a top priority for principals, requiring them to devote substantially greater amounts of their time and resources to facilitating teacher observations and completing reports associated with these observations (Darling-Hammond et al., 2012). Principals are required in the school district by the Ohio Department of Education (2016) to complete a series of observations and reports, including short-form and long-form, brief and extended versions, for their assigned rosters of teachers, all within state-specified periods of reporting accompanied by mandatory deadlines. They are accountable for ensuring that evaluations are successfully completed for each teacher, yet principals must also maintain all of their other assigned duties, including building and facilities management, security, student discipline, community relations, and other aspects of the role that typically require ample resources and attention, often of an immediate nature in a number of instances.

None of the teachers I interviewed for this research study seemed to believe that principals should completely give up their role as evaluators. Nevertheless, there does seem to be some room for balance and improvement within their capacity as evaluators. The school district itself, as I learned during my facilitation of this research study, has responded to the increasing demands being placed upon principals by OTES and other recent trends in education policy by splitting up general aspects of the position into designated, specific roles, such as operations and instructional supervision. In doing so, the district has attempted to support its principals by balancing their workload and responsibilities in such a way that enables them to focus and improve their performance both individually as well as collectively as a team. Perhaps further adjustments and considerations are needed, however, and may benefit through a process of teacher and community input. The school district featured within this study, and others, may also

benefit from exploring and perhaps sharing best practices among school districts, understanding however that each solution is likely to be unique based upon the needs and interests of each school district.

Improving alignment between evaluators' professional backgrounds and the teachers they evaluate. Lack of principal experience as classroom teachers, content area expertise, and success in the classroom were concerns voiced by several teachers during my conversations with them and raised teachers' concerns about the relevancy and validity of the evaluations they received by their assigned principals. These teachers wondered how a principal with little or no professional experience in a particular content area could reasonably be expected to judge the performance of a teacher with substantially greater experience teaching their subject. They also expressed concern about principals too long removed from professional practice to fully comprehend recent changes in education as well as newer developments resulting from ongoing trends. The teachers participating in this study that expressed such concerns do not necessarily see how their principal can help and support them through it when they themselves are not well-prepared or best suited for the role of a professional evaluator.

Although the school district has made recent attempts to balance the role of its principals, particularly at the secondary level, it was not immediately clear what additional changes could be made to the role or allocation of principals within the district from a practical standpoint, in order to support teachers with respect to their concerns about principal experience and content knowledge. One possibility, however, might be to appoint experienced teachers to "Instructional Supervisor" non-instructional special assignment roles and grant them the means and authority to be able to facilitate teacher evaluations. Teachers could be selected for Instructional Supervisor roles to conduct evaluations of practicing teachers based upon demonstrated records of success

and expertise within their respective content areas, as well as across multiple areas of instructional and professional practice.

Perhaps teachers could be appointed to serve in the Instructional Supervisor role on a limited basis, perhaps ranging from one to three years, so that they would not lose the relevancy of their own practitioner knowledge and skills. Of course, considerations would need to be made for allocating funding to compensate members of a new class of administrator, as well as for justifying the creation of new administrative positions before the community. Also, there may need to be a discussion between the union and district administration regarding any changes that may be needed within the collective bargaining agreement in order to accommodate a new, temporary role for teachers.

Incorporating mentoring into evaluations. Mentoring plays a limited role in teacher evaluations today. Within Ohio, new teachers in the Resident Educator program receive a mentor who assists them in successfully completing the requirements of the program, which incorporates performance evaluations. Once graduating from the program and obtaining a professional license, teachers may or may not be assigned a new mentor. In situations where teachers do receive a mentor, post-Resident Educator, the mentor is not formally involved in any part of the OTES evaluations that teachers receive throughout their careers. Considering that the teachers I interviewed were unanimous in their support of mentors, mentoring, and support as a means of encouraging the professional growth and development of teachers, this needs to change.

My recommendation is for an overhaul of OTES that formally incorporates mentors into the evaluation process in a capacity superior to that of principals, whose role would henceforth be limited to formatively assessing teachers' professional contributions. OTES mentors would support full-licensed teachers with meeting the standards upon which they are assessed during

evaluations. Further, these mentors could support teachers' in achieving mutually agreed-upon professional growth and development goals. This would be similar to how Resident Educator mentors support new teachers in meeting the requirements of their induction program. Such mentors would only formatively evaluate teachers, observing their practice and then providing them with constructive feedback about one or more aspects of their performance during a given period. Any feedback that mentors provide to teachers about their performance would only be for the purpose of fostering teachers' professional growth and development, and not for summative decisions about employment. Accordingly, evaluations would no longer be about summative employment decisions.

Holistic evaluations. The current evaluation system is designed primarily with two specific expectations in mind. First, teachers will be observed several times during the course of the school year by an evaluator, usually a principal, who will at the end of the year assign them with a quantifiable rating of their performance based upon what they have observed during several brief instances relative to a year's worth of instructional time. Second, teachers will demonstrate the quality of their instructional practice in a quantifiable manner, using test scores from an assessment of either their own design, or of the school's making. The sum of these two measures at the end of the year produce an overall summative rating of the teacher's performance.

Such a system is fundamentally antithetical to a holistic system of evaluation in which many measures of a given teacher's performance are used to formatively assess and provide opportunities for constructive reflection and feedback about how well the teacher is performing. The purpose of such holistic evaluations, however, is not to produce a quantifiable, summative rating for each teacher that can then be used to rationalize retention, promotion, or firing thereof.

Rather, it is about promoting personal and professional growth for their own sake, a seemingly novel but long-established humanistic and educational concept (Dewey, 1938; Ovando, 2001; Peterson, 1987; Smagorinsky, 2014). This particular concept is quite relevant within a view of education as an ongoing lifelong experience—one lived by most teachers in accordance with the license renewal cycle and the professional development it requires at the very least—as opposed to one in which education is framed as a limited opportunity to acquire essential knowledge and facts before entering adult life and the working world.

A holistic system of teacher evaluation is therefore precisely what ought to replace OTES. Two or, in some instances, three measures alone cannot constitute a valid or reliable assessment of any teacher's performance when there are so many other variables to consider. Such variables include students' ability to focus within the classroom, and random disruptions to instructional continuity because of public address system announcements, among other things, simply are beyond teachers' ability to control. Accordingly, it is both unfair and unethical to assign teachers a rating based upon such measures that could impact their professional employment, even when the outcome, on the surface, appears to be a positive one, such as when a teacher receives a commendation or a promotion for high performance. Further, a haunting question that emerges here, is one of whether such a commendation or a promotion were truly warranted or deserved, particularly if the teacher receiving it enjoyed a uniquely fortunate and non-replicable contributing set of circumstances.

Within the ideal holistic system of teacher evaluation, teachers' ongoing employment and any administrative decisions therein would not be contingent on summative evaluation outcomes or based upon any part of the evaluation process itself. In fact, no part of teachers' holistic evaluations would be summative. Multiple measures of a teacher's performance, such as parent

surveys, teachers' own professional self-assessments, student portfolios, administrator assessments of teachers' professional contributions, and additional items, would be collected at multiple intervals throughout the course of the school year. Such "snapshots" would also still include observations of teacher instruction, but they would be facilitated by peer teachers, either alone or in small groups, serving as critical yet supportive evaluators who would provide the peer teacher being observed with constructive, formative feedback.

Collectively, these multiple measures of teachers' professional performance would serve to provide teachers with what would essentially be a panoramic image of their professional practice. Teachers would be encouraged to take in all of the information it provides to them, from the fine details to the bigger picture in its entirety, reflect upon it, and then seek out professional development, including perhaps mentoring, in a specific area of need. This approach would genuinely benefit and support the professional development and growth of all teachers, regardless of their level of experience, professional skill or knowledge interests, or their developmental needs.

In essence, from a policy implementation standpoint, a holistic evaluation system would make teachers wholly accountable for their own professional growth and development, based upon the constructive, formative feedback they receive from their school and learning community. It would accomplish this by removing the responsibility for enforcement from principals and other administrators within schools, districts, and state government itself, and place this responsibility within the hands of teachers. The act of doing so would not only serve to empower teachers and entrust them with full responsibility for their own professional growth and development, but it would also both foster and reflect a shift in the political and social climate. The result of such a shift in approach to teacher evaluation is a culture in which teachers are

valued and trusted as highly autonomous professionals who both represent and uphold an ethos of lifelong learning in the field of education, as well as within society. In this new paradigm, teachers, and not business leaders or efficiency advocates, become the well-respected experts within their own profession.

In addition, implementing a holistic system of teacher evaluation would return balance to other roles currently given high priority within the current system, and place them in their proper context. As a result, principals are no longer burdened with the excessive responsibility of facilitating teacher observations, and instead have more freedom to handle their administrative responsibilities while still contributing to teachers' holistic evaluations from a more authentic perspective. Assessment data is also just one more potential piece within a formative puzzle about a given teacher's strengths, talents, skills, and areas of need, instead of an all-important yet deeply-flawed, unreliable and invalid arbiter of a teacher's effectiveness.

Clearly the benefits of such a holistic system of teacher evaluation being implemented for teachers would be numerous, beyond the points highlighted here. Most critically, however, is the fact that it would support a genuine integration of Dewey's (1938) perspective of learning as being holistic and integrated with the lived experiences of students, teachers, and the communities in which they live. Public education facilitated in this manner would be constructively transformed, and it would also serve as a constructive, transformative agent within the lives of our students. It would further benefit the professional growth and development of teachers, and also support the health and well-being of our communities. Accordingly, I strongly recommend to state policymakers that OTES be abolished immediately and replaced with an entirely new, holistic system of teacher evaluation based upon my theoretical considerations presented herein.

Implications for Ohio State Government and Education Policy Makers

Based upon the perspectives that teachers within this study have shared, and also supported by previous research, it is apparent that teacher evaluation, as presently constituted here in Ohio, falls short of its potential to serve as a valid and reliable system supporting the professional growth and development of teachers. A research-based restructuring of the role principals maintain within evaluations should be a priority. Also given consideration within this restructuring should be the demands principals face to fulfill many of their responsibilities—including teacher evaluations—that several teachers discussed within this research study. Further, the concern teachers within this study expressed regarding the ability of principals to conduct evaluations of teachers that are valid and without bias under ongoing pressure to keep up with their responsibilities should also be considered.

In addition, some teachers have questioned the use of assessment data within teacher evaluations based upon the fact that students can intentionally elect to fail such assessments in order to inflict professional harm upon a teacher they do not like, and face no negative consequences for such willfully malevolent behavior. The observations provided by these teachers support findings within other research studies that have similarly called into question assessment measures as a form of teacher evaluation when such variables exist that are beyond teachers' ability to control (Baker et al., 2010; Darling-Hammond et al., 2012; Harris et al, 2014). Such circumstances make teacher evaluation reform not merely necessary, but urgently so. To this end, I provide the following recommendations to Ohio state government and education policy makers:

Reform OTES to prioritize formative evaluations while also expanding the number of measures of performance used. An evaluation system that has the potential to produce

unreliable and invalid results must be overhauled in order to achieve to the greatest extent possible, a system that not only produces reliable and valid outcomes, but ones that are also fair and equitable in the eyes of the professionals being evaluated within it. Indeed, a majority of the teachers participating in this research study suggest that the current system needs to be reformed to this end, and these teachers are further supported by other research suggesting a need for evaluation systems that are formative, rather than summative (Croft et al., 2016; Ellis, 1986; & Smagorinsky, 2014), and that also feature diverse measures of performance (Croft et al., 2016; Peterson, 1987; & Smagorinsky, 2014).

To this end, I recommend that OTES be overhauled to eliminate summative ratings that are based upon principal observations and student assessment data. In its place, a formative system of evaluation should be instituted that incorporates comprehensive measures of teachers' performance in order to create a broad view of instructional strengths and areas of opportunity. Such a system could also provide substantial incentive for professional dialogue and mentoring, items of interest to many teachers who participated within this research study, and that are supported as legitimate professional practice within the literature (Danielson, 2017; Marzano, 2017).

Implications for Teachers

It is my personal opinion, albeit one supported by the lens of critical theory (Anyon, 2011; Aronowitz, 2009), that as a group, teachers must continue to advocate for our professional concerns and needs in all matters of education, but specifically with regards to policies of accountability that are an integral component of the dominant structural paradigm. Further, we need to collectively raise our voices and express them with regard to an evaluation system that does not appear to have been designed to support our professional development and growth. It is

indeed possible to conceive of a better evaluation system, one that affirms a constructive and collective ethic of profession (Shapiro & Stefkovich, 2011), learning as a democratic right (Strauss, 2017), and an overall holistic perspective of education for which Dewey (1938) advocated not only for our field, but for all aspects of our lives. Furthermore, Counts (1932) argued that teachers carry the responsibility of being active participants in creating a better society and strengthening the democratic tradition upon which it must be based.

Implications for Future Research

As a qualitative research study of teacher evaluation from the perspectives of teachers within a public-school district, this work contributes towards filling a gap within the body of research relevant to teachers' perspectives on teacher evaluation. As of the time this study was conducted, no other relevant research studies were accessible that addressed the central research questions that I have sought to address here. Accordingly, I undertook this study to examine the central issue for the benefit of the education field, as well as to fill the gap I identified, and further to grant a voice to teachers where none previously appeared to exist.

While I was unable to reference previous studies into teacher evaluation from the perspectives of teachers to build on for this research study, I have referenced a number of other professional and research works for their general contribution to my knowledge of teacher evaluation and relevant practices, most notably among them Danielson (1999; 2017) and Marzano (2012; 2017). This study builds most strongly on their work and contributes the perspectives of a group of teachers from an Ohio-based school district to a body of research concerned with evaluation practices, including teachers' perspectives on evaluation, evaluation validity and relevancy, and also mentoring in the context of evaluation.

Future research studies may consider expanding upon my work to address teacher perspectives on evaluation more broadly by incorporating perspectives across in-state school districts, as well as nationally, across multiple states and evaluation systems. Further, such studies may seek to build upon the use of critical theory and feminist theory as means of both critically uncovering inequitable situations for teachers (Anyon, 2011; Aronowitz, 2009; Giroux, 2009), as well as for granting them voice where it has been denied, upholding an Ethic of Care in the process of doing so (Noddings, 2009).

Significance

This study, focused on teacher evaluation as designed and implemented within Ohio during the era of high-stakes accountability in schools, is a contribution towards the gap in research concerning the quality of teacher evaluation from the perspectives of teachers. In the same context, it serves to provide teachers with a voice in the body of research concerning a critical component of their role, professional evaluation. Such a perspective has not been well-evidenced previously within previous research. Accordingly, this research study and its findings in this study suggests several possibilities for future research.

Such possibilities include broadening this particular line of research to include the voices of teachers within other school districts within Ohio. Additional possibilities include studying teacher evaluation systems and teacher perspectives on these within other states. Furthermore, the potential exists to study teacher evaluation within non-traditional school settings, such as charter schools and private schools. Still further, the possibility exists for studying teacher evaluation from the perspective of other members of the school community, including principals, as well as parents, for the purpose of gaining insight into their unique perceptions of the phenomenon.

Ultimately, it is my expectation that the body of research on this subject will expand substantially, and that it will guide educational policy and activism in the direction of deep systemic reform. The underlying principles of education during the post-modern era must be replaced with holistic concepts of learning and growth that acknowledge the human condition and seek to elevate it to its highest potential. While such concepts must guide the learning experiences of our students, they must also serve to guide teacher induction, professional development, and growth.

References

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

Anyon, J. (2011). *Marx and education*. New York: Routledge.

Aronowitz, S. (2009). *Against schooling: Education and social class*. In A. Darder, M. P. Baltodano, & R. D. Torres, (Eds.), The critical pedagogy reader. (2nd ed.). New York: Routledge.

Attick, D., & Boyles, D. (2016). Pearson learning and the ongoing corporatization of public education. *Journal of Thought*, *50*(1-2), 5-19. Retrieved from http://go.galegroup.com.proxy.libraries.uc.edu/ps/i.do?p=EAIM&sw=w&u=ucinc_main&v=2.1&it=r&id=GALE%7CA459985358&sid=summon&asid=44959b798a5d7a706c2ac8b93c371f79

Baker, B. D., Oluwole, J. O., & Green, P. C. (2013). The legal consequences of mandating high stakes decisions based on low quality information: Teacher evaluation in the Race-to-the-Top era. *Education Policy Analysis Archives*, *21*(5), 1–71.

Baker, E. L., Barton, P. E., Darling-Hammond, L., Haertel, E., Ladd, H. F., Linn, R. L., Ravitch, D., Rothstein, R., Shavelson, R. J., & Shepard, L. A. (2010). *Problems with the use of student test scores to evaluate teachers* (EPI Briefing paper 278). Washington, DC: Economic Policy Institute.

Blakely, J. (2017, Apr 17). *How school choice turns education into a commodity*. Retrieved from https://www.theatlantic.com/education/archive/2017/04/is-school-choice-really-a-form-of-freedom/523089/

Boas, T. C. & Gans-Morse, J. (2009). Neoliberalism: From new liberal philosophy to anti-liberal slogan". *Studies in Comparative International Development*, *44*(2), 137–161.

Booher-Jennings, J. (2005). Below the bubble: "Educational triage" and the Texas accountability system. *American Educational Research Journal*, *42*(2), 231-268.

Borko, H. (2004). Professional development and teacher learning: Mapping the terrain. *Educational Researcher*.

Brandt, C., Mathers, C., Olivia, M., Brown-Sims, M., & Hess, J. (2007). *Examining district guidance to schools on teacher evaluation policies in the Midwest Region* (Issues & Answers Report, REL 2007-No. 030). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Midwest. Retrieved from http://ies.ed.gov/ncee/edlabs/regions/midwest/pdf/REL_2007030.pdf

Brown, A. B., & Clift, J. W. (2010). The unequal effect of adequate yearly progress: Evidence from school visits. American Educational Research Journal, *47*(4), 774-798.

Burns, R. W., & Badiali, B. J. (2015). When supervision is conflated with evaluation: Teacher candidates' perceptions of their novice supervisor. *Action in Teacher Education*, *37*(4), 418-437.

Colby, S. A., Bradshaw, L. K., & Joyner, R. L. (2002, April 4). Perceptions of teacher evaluation systems and their impact on school improvement, professional development and student learning. *Paper presented at the 2002 Annual Meeting of the American Educational Research Association.* Retrieved from http://files.eric.ed.gov/fulltext/ED464916.pdf

Cook, T. D., Habib, F. N., Phillips, M., Settersten, R. A., Shagle, S. C., & Degirmencioglu, S. M. (1999). Comer's school development program in Prince George's County, Maryland: A theory-based evaluation. *American Educational Research Journal*, *36*(3), 543-597.

Counts, G.S. (1932). *Dare the school build a new social order?* New York: John Day Company.

Craig, C. J. & Ross, V. (2008). Cultivating the image of teachers as curriculum makers. In F. M. Connelly, M. F. He, & J. A. Phillion (Eds.), *The SAGE handbook of curriculum and instruction* (282-305). Los Angeles: Sage.

Croft, S. J., Roberts, M. A., & Stenhouse, V. L. (2016). The perfect storm of education reform: High-stakes testing and teacher evaluation. *Social Justice*, *42*(1), 70-92,147.

Danielson, C. (1999). Mentoring beginning teachers: The case for mentoring. *Teaching and Change*, *6*(3), 251-57.

Danielson, C. (2009). *Talk about teaching: Leading professional conversations*. Thousand Oaks, CA: Corwin.

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

Darling-Hammond, L., Amrein-Beardsley, A., Haertel, E., & Rothstein, J. (2012). Evaluating teacher evaluation. *Phi Delta Kappan*, *93*(6), 8-15.

Dewey, J. (1938). *Experience and education*. New York: Touchstone.

Ellis, T. I. (1986). Teacher evaluation. *Research roundup*, 2(2), 1-6. ERIC Clearinghouse on Educational Management.

Fang, L. (2014). Venture capitalists are poised to 'disrupt' everything about the education market. *The Nation*. Retrieved from https://www.thenation.com/article/venture-capitalists-are-poised-disrupt-everything-about-education-market/

Feiman-Nemser, S. (1998). Teachers as teacher educators. *European Journal of Teacher Education*, *21*(1), 63-74.

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

Firestone, W. A., & Pennell, J. R. (1993). Teacher commitment, working conditions, and differential incentive policies. *Review of Educational Research*, *63*(4), 489-525.

Firestone, W. A., & Rosenblum, S. (1988). Building commitment in urban high schools. *Educational Evaluation and Policy Analysis*, *10*(4), 285-299.

Freire, P. (2000). *Pedagogy of the oppressed*. New York: Bloomsbury.

Frye, B., Button, L., & Kelly, C. (2010). Preservice teachers' self-perceptions and attitudes toward culturally responsive teaching toward culturally responsive teaching. *Journal of Praxis in Multicultural Education*, *5*(1), 5–22.

Garet, M. S., Porter, A. C., Desimone, L., Birman, B. F., & Yoon, K. S. (2001). What makes professional development effective? Results from a national sample of teachers. *American Educational Research Journal*, *38*, 915–946.

Giroux, H. A. (2009). Critical theory and educational practice. In A. Darder, M. P. Baltodano, & R. D. Torres, (Eds.), *The critical pedagogy reader*. (2nd ed.). New York: Routledge.

Goe, L., Bell, C., & Little, O. (2008, June). Approaches to evaluating teacher effectiveness: A research synthesis. *National Comprehensive Center for Teacher Quality*. Retrieved from http://dev-tqsource.airws.org/publications/EvaluatingTeachEffectiveness.pdf

Guarino, C. M., Reckase, M. D., & Wooldridge, J. M. (2014). Can value-added measures of teacher performance be trusted? *Education Finance and Policy*, *10*(1), 117–156.

Hanushek, E. A., & Rivkin, S. G. (2010). Generalizations about using value-added measures of teacher quality. *American Economic Review: Papers & Proceedings, 100*, 267-271.

Harris, D. N., Ingle, W. K., & Rutledge, S. A. (2014). How teacher evaluation methods matter for accountability: A comparative analysis of teacher effectiveness ratings by principals and teacher value-added measures. *American Educational Research Journal*, *51*(1), 73-112.

Hatch, J. A. (2002). *Doing qualitative research in education settings*. Albany, NY: SUNY.

Hill, H. C., Kapitula, L., & Umland, K. (2011). A validity argument approach to evaluating teacher value-added scores. *American Educational Research Journal*.

Holloway, J. (2001) The benefits of mentoring. *Educational Leadership*, *58*(8), 85–86.

Holloway-Libell, J., & Collins, C. (2014). VAM-based teacher evaluation policies: Ideological foundations, policy mechanisms, and implications. *InterActions: UCLA Journal of Education and Information Studies*, *10*(1), 1-23.

Huang, F. L., & Moon, T. R. (2009). Is experience the best teacher? A multilevel analysis of teacher characteristics and student achievement in low performing schools. *Educational Assessment Evaluation and Accountability*, *21*(3), 209–234.

Hughes, G. D. (2012). Teacher retention: Teacher characteristics, school characteristics, organizational characteristics, and teacher efficacy. *Journal of Educational Research, 105*(4), 245-255.

Hunter, A and Brewer, J (2015). Designing Multimethod Research In: Hesse-Biber, S and Johnson, RB eds. *The Oxford Handbook of Multimethod and Mixed Methods Research Inquiry*. London: Oxford University Press, p. 185–205.

Ingersoll, R. M. & Strong, M. (2011). The impact of induction and mentoring programs for beginning teachers: A critical review of the research. *Review of Educational Research*, *81*(2), 201-233.

Jennings, J. L. (2012). The effects of accountability system design on teachers' use of test score data. *Teachers College Record*, *114*, 1–23.

Jimenez-Silva, M. & Olson, K. (2012). A community of practice in teacher education: Insights and perceptions. *International Journal of Teaching and Learning in Higher Education*, *24*(3), 335-348.

Johnson, S. M., & Birkeland, S. E. (2003). Pursuing a "sense of success": New teachers explain their career decisions. *American Educational Research Journal*, *40*(3), 581-617.

Jones, C., Parker, M., Bos, R. T. (2005). *For business ethics*. Abingdon, UK: Routledge.

Kelley, J. P. (2017, Sep 29). *Schools rate almost no teachers ineffective*. Retrieved from http://www.mydaytondailynews.com/news/schools-rate-almost-teachers-ineffective/4vq2RAUfoPdn0NrQKCrKmJ/

Kimball, S. M. (2002). Analysis of feedback, enabling conditions and fairness perceptions of teachers in three school districts with new standards-based evaluation systems. *Journal of Personnel Evaluation in Education, 16,* 4, 241-268.

Kushman, J. W. (1992). The organizational dynamics of teacher workplace commitment: A study of urban elementary and middle schools. *Educational Administration Quarterly*.

Kyriakides, L., & Demetriou, D. (2007). Introducing a teacher evaluation system based on teacher effectiveness research: An investigation of stakeholders' perceptions. *Journal of Personnel Evaluation in Education, 20,* 43-64.

Marzano, R. J. (2011). *Art and science of teaching: Teacher evaluation model: Domain 4: Collegiality and professionalism*. Retrieved from http://www.marzanoevaluation.com/files/domain4/Marzano_AST_Domain4_LongForm.pdf

Marzano, R. J. (2012). *Marzano teacher evaluation*. Retrieved from http://www.marzanoevaluation.com/

Marzano, R. J. (2017). *Tips from Dr. Marzano*. Retrieved from https://www.marzanoresearch.com/resources/tips/cci_tips_archive

Miles, M. B. & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd ed.). Thousand Oaks, CA: Sage.

Musella, D. (1970). Improving teacher evaluation. *Journal of Teacher Education*, 21(1), 15-21.

Noddings, N. (1984). *Caring: A feminine approach to ethics and moral education*. Berkley, CA: University of California Press.

Noddings, N. (2007). *When school reform goes wrong*. New York: Teachers College Press.

Ohio Department of Education. (2016). Educator evaluation overview. Retrieved from http://education.ohio.gov/Topics/Teaching/Educator-Evaluation-System/Educator-Evaluation-Overview

Ohio Department of Education. (2017a). Resident educator FAQs. Retrieved from http://education.ohio.gov/Topics/Teaching/Resident-Educator-Program/Resident-Educator-FAQs#FAQ572

Ohio Department of Education. (2017b). The mentoring years. Retrieved from http://education.ohio.gov/Topics/Teaching/Resident-Educator-Program/The-Mentoring-Years

Ohio Department of Education. (2017c). Teacher performance. Retrieved from http://education.ohio.gov/Topics/Teaching/Educator-Evaluation-System/Ohio-s-Teacher-Evaluation-System/Teacher-Performance-Ratings

Ohio Department of Education. (2017d). Value-added student growth measure. Retrieved from http://education.ohio.gov/Topics/Teaching/Educator-Evaluation-System/Ohio-s-Teacher-Evaluation-System/Student-Growth-Measures/Value-Added-Student-Growth-Measure

Orfield, G. & Lee, C. (2007). *Historic reversals, accelerating resegregation, and the need for new integration strategies*. Retrieved from http://civilrightsproject.ucla.edu/research/k-12-education/integration-and-diversity/historic-reversals-accelerating-resegregation-and-the-need-for-new-integration-strategies-1/orfield-historic-reversals-accelerating.pdf

Ovando, M. N. (2001). Teachers' perceptions of a learner-centered teacher evaluation system. *Journal of Personnel Evaluation in Education, 15*(3), 213-231.

Papay, J. P. (2012). Refocusing the debate: Assessing the purposes and tools of teacher evaluation. *Harvard Educational Review*, *82*(1), 123–142.

Peterson, K. D. (1987). Teacher evaluation with multiple and variable lines of evidence. *American Educational Research Journal*, *24*(2), 311-317.

Prendergast, C. (2008). Intrinsic motivation and incentives. *American Economic Review* (Vol. 98, pp. 201–205).

Pring, R. (2000). *Philosophy of educational research* (2nd ed.). London: Continuum.

Rivkin, S. G., & Hanushek, E. A. (2005). Teachers, schools, and academic achievement. *Econometrica*, *73*(2), 417–458.

Rubin, H. J., & Rubin, I. S. (2012). *Qualitative interviewing: The art of hearing data* (*3rd ed*.). Thousand Oaks, CA: Sage.

Sanders, W. L., & Horn, S. P. (1994). The Tennessee value-added assessment system (TVAAS): Mixed-model methodology in educational assessment. *Journal of Personnel Evaluation in Education*, *8*(3), 299–311.

Sergiovanni, T. J. & Starratt, R. J. (2007). *Supervision: A redefinition*. New York: McGraw-Hill.

Shapiro, J.P. & Stefkovich, J.A. (2011). *Ethical leadership and decision making in education: Applying theoretical perspectives to complex dilemmas* (3rd ed.). New York: Routledge.

Smagorinsky, P. (2014). Authentic teacher evaluation: A two-tiered proposal for formative and summative assessment. *English Education*, *46*(2), 165-185.

Smith, T., & Ingersoll, R. (2004). What are the effects of induction and mentoring on beginning teacher turnover? *American Educational Research Journal*, *41*, 687–714.

Strauss, V. (2017). There is no constitutional right to a high-quality public education. Should there be? Retrieved from https://www.washingtonpost.com/news/answer-sheet/wp/2017/09/17/there-is-no-constitutional-right-to-a-high-quality-public-education-should-there-be/?utm_term=.b13a6a6b999f

Stronge, J. H., Ward, T. J., & Grant, L. W. (2011). What makes good teachers good? A cross-case analysis of the connection between teacher effectiveness and student achievement. *Journal of Teacher Education*, 62(4), 339-355. American Association of Colleges for Teacher Education.

Taylor, S. S., Rudolph, J. W., & Foldy, E. G. (2008). Teaching reflective practice in the action science/action inquiry tradition: Key stages, concepts and practices. In P. Reason, & H. Bradbury (Eds.), *The SAGE handbook of action research: Participative inquiry and practice* (656-668). Thousand Oaks, CA: Sage.

The Danielson Group. (2017). *The Framework*. Retrieved from http://www.danielsongroup.org/framework/

Tschannen-Moran, M., Woolfolk Hoy, A., & Hoy, W. K. (1998). Teacher efficacy: Its meaning and measure. *Review of Educational Research*, *68*, 202–248.

U.S. Department of Education. (2016). Race to the top fund. Retrieved from https://www2.ed.gov/programs/racetothetop/index.html

University of Cincinnati. (2005). Human research protection program policy: Protection of confidential data in human subjects research. Retrieved from http://ahc-sharepoint.uc.edu/hrp_policies/HRP%20Policies/IV%20-%20INVESTIGATOR%20RESPONSIBILITIES/Policy%20IV.03%20Human%20Research%20Protection%20Training-Researchers.pdf

Wayman, J. C., & Jimerson, J. B. (2014). Teacher needs for data-related professional learning. *Studies in Educational Evaluation*, *42*, 25–34.

Wiarda, H. (1978). Corporatist theory and ideology: A Latin American development paradigm. *Journal of Church and State,20*(1), 29-56.

Zimmerman, S., & Deckert-Pelton, M. (2003). Evaluating the evaluators: teachers' perceptions of the principal's role in professional evaluation. *NASSP Bulletin*, *87*(636), 28-37.

Appendix A

Superintendent Letter

January 15, 2014

Dear Superintendent:

I am a doctoral student at the University of Cincinnati seeking to conduct a research study in order to gain understanding from the perspective of teachers as to how their input into the Ohio Teacher Evaluation System (OTES) may best support performance evaluation reform. I am currently seeking teacher volunteers to participate within interviews. Below, I have outlined the criteria for potential volunteers:

Participants may be included in this study if they are all of the following:

- A licensed K-12 public school teacher in Ohio or Kentucky actively assigned to a classroom.

It is not anticipated that participants will receive any direct benefit because of being in this study, although they will be contributing toward research on teacher evaluation and a greater understanding of teacher performance evaluation design. Also, it should be noted that the risks for participation are minimal, and extensive measures will ensure confidentiality of participants and any information shared.

I would like to set up a time to discuss this study with you in additional detail. If you are interested in assisting with this study, please contact me at stegalje@mail.uc.edu.

Sincerely,

John E. Stegall Jr.

Appendix B

Adult Consent Form for Research—Phase I

IRB #: 2014-0434

Approved: 3/19/2014
Do Not Use After: 3/18/2015

Adult Consent Form for Research
University of Cincinnati
Department: College of Education, Criminal Justice, and Human Services
Principal Investigator: John E. Stegall Jr.
Faculty Advisor: Dr. Mary Brydon-Miller

Title of Study: Teacher Evaluation from the Perspectives of Teachers

Introduction:
You are being asked to take part in a research study. Please read this paper carefully and ask questions about anything that you do not understand.

Who is doing this research study?
The person in charge of this research study is John E. Stegall Jr. of the University of Cincinnati (UC) College of Education, Criminal Justice, and Human Services. He is the sole person conducting this research study and is being guided by Dr. Mary Brydon-Miller.

What is the purpose of this research study?
The purpose of this research study is to gain understanding of the perspective of K-12 public school classroom teachers regarding the Ohio Teacher Evaluation System (OTES). These findings may be generalizable to the population of Ohio teachers subject to the OTES.

Who will be in this research study?
About 100 to 200 people will take part in this study. You may be in this study if you are all of the following:

- A licensed K-12 public school teacher in Ohio actively assigned to a classroom.
- Age 21 and older.

What if you are an employee where the research study is done?
Taking part in this research study is not part of your job. Refusing to be in the study will not affect your job. You will not be offered any special work-related benefits if you take part in this study.

What will you be asked to do in this research study, and how long will it take?
You will be asked to respond to a series of 9 profession-related questions posed within an audiotaped, interview format. The interview will require approximately 45 to 60 minutes. The

interview will take place on site within your school in a designated private room or otherwise unattended space. The information being collected in the interview may include your personal and professional perspectives on your work, training, and licensure as an educator.

Are there any risks to being in this research study?
It is not expected that you will be exposed to any risk by allowing your interview responses to be used in this research study. Some questions may make you uncomfortable. You can refuse to answer any questions that you don't want to answer. If you want to talk to someone because this research made you feel upset, the researchers can give you information about people who may be able to help you.

Are there any benefits from being in this research study?
You will not get any benefit because of being in this study. But, being in this study may help educators understand how teacher evaluation may support licensure.

What will you get because of being in this research study?
You will not be paid to take part in this study.

Do you have choices about taking part in this research study?
If you do not want to take part in this research study you may simply not participate. You will not be treated any differently. Audiotaping the interview is important to ensure accurate analysis of your comments. Therefore if you do not want to be audiotaped you should choose not to be in this research study.

How will your research information be kept confidential?
Information about you will be kept private by use of the following:

- A study ID number instead of the participant's name on the research forms.
- Keeping the master list of names and study ID numbers in a separate location from the research forms.
- Limiting access to research data to the research team.
- Not including the participant's name on the typed transcript.
- Erasing audiotapes as soon as they are transcribed.
- Keeping research data on a password-protected computer.

Your information will be kept on written paper and/or digital file (jump drives) for up to three years within a locked cabinet in the faculty researcher's campus office. After three years, it will be destroyed by shredding and/or deletion. Signed consent documents and master lists of participant names and ID numbers will be stored in a separate locked cabinet in the faculty researcher's campus office. After three years, these items will be destroyed by shredding and/or deletion. NOTE: identifiers such as name, birth date, etc. will be deleted as soon as possible. NOTE: federal regulations require that signed consent documents must be kept for a minimum of three years after the study is closed.

The data from this research study may be published; but you will not be identified by name.

Agents of the University of Cincinnati may inspect study records for audit or quality assurance

purposes.

Your identity and information will be kept confidential unless the authorities have to be notified about abuse or immediate harm that may come to you or others.

What are your legal rights in this research study?
Nothing in this consent form waives any legal rights you may have. This consent form also does not release the investigator, the institution, or its agents from liability for negligence.

What if you have questions about this research study?
If you have any questions or concerns about this research study, you should contact John E. Stegall Jr. at (XXX) XXX-XXXX or stegalje@mail.uc.edu.

The UC Institutional Review Board reviews all research projects that involve human participants to be sure the rights and welfare of participants are protected.

If you have questions about your rights as a participant or complaints about the study, you may contact the UC IRB at (513) 558-5259. Or, you may call the UC Research Compliance Hotline at (800) 889-1547, or write to the IRB, 300 University Hall, ML 0567, 51 Goodman Drive, Cincinnati, OH 45221-0567, or email the IRB office at irb@ucmail.uc.edu.

Do you HAVE to take part in this research study?
No one has to be in this research study. Refusing to take part will NOT cause any penalty or loss of benefits that you would otherwise have. You may skip any questions that you don't want to answer.

You may start and then change your mind and stop at any time. To stop being in the study, you should tell John E. Stegall Jr. at (XXX) XXX-XXXX or stegalje@mail.uc.edu.

Agreement:
I have read this information and have received answers to any questions I asked. I give my consent to participate in this research study. I will receive a copy of this signed and dated consent form to keep.

Participant Name (please print) _____

Participant Signature _____ Date _____

Signature of Person Obtaining Consent _____ Date _____

Appendix C

Adult Consent Form for Research—Phase II

IRB #: 2014-0434

Approved: 3/20/2015
Do Not Use After: 3/19/2016

Adult Consent Form for Research
University of Cincinnati
Department: College of Education, Criminal Justice, and Human Services
Principal Investigator: John E. Stegall Jr.
Faculty Advisor: Dr. Mary Brydon-Miller

Title of Study: Teacher Evaluation from the Perspectives of Teachers

Introduction:
You are being asked to take part in a research study. Please read this paper carefully and ask questions about anything that you do not understand.

Who is doing this research study?
The person in charge of this research study is John E. Stegall Jr. of the University of Cincinnati (UC) College of Education, Criminal Justice, and Human Services. He is the sole person conducting this research study and is being guided by Dr. Mary Brydon-Miller.

What is the purpose of this research study?
The purpose of this research study is to gain understanding of the perspective of K-12 public school classroom teachers regarding the Ohio Teacher Evaluation System (OTES). These findings may be generalizable to the population of Ohio teachers subject to the OTES.

Who will be in this research study?
About 100 to 200 people will take part in this study. You may be in this study if you are all of the following:

- A licensed K-12 public school teacher in Ohio actively assigned to a classroom.
- Age 21 and older.

What if you are an employee where the research study is done?
Taking part in this research study is not part of your job. Refusing to be in the study will not affect your job. You will not be offered any special work-related benefits if you take part in this study.

What will you be asked to do in this research study, and how long will it take?
You will be asked to respond to a series of 9 profession-related questions posed within an audiotaped, interview format. The interview will require approximately 45 to 60 minutes. The

interview will take place on site within your school in a designated private room or otherwise unattended space. The information being collected in the interview may include your personal and professional perspectives on your work, training, and licensure as an educator.

Are there any risks to being in this research study?
It is not expected that you will be exposed to any risk by allowing your interview responses to be used in this research study. Some questions may make you uncomfortable. You can refuse to answer any questions that you don't want to answer. If you want to talk to someone because this research made you feel upset, the researchers can give you information about people who may be able to help you.

Are there any benefits from being in this research study?
You will not get any benefit because of being in this study. But, being in this study may help educators understand how teacher evaluation may support licensure.

What will you get because of being in this research study?
You will not be paid to take part in this study.

Do you have choices about taking part in this research study?
If you do not want to take part in this research study you may simply not participate. You will not be treated any differently. Audiotaping the interview is important to ensure accurate analysis of your comments. Therefore if you do not want to be audiotaped you should choose not to be in this research study.

How will your research information be kept confidential?
Information about you will be kept private by use of the following:

- A study ID number instead of the participant's name on the research forms.
- Keeping the master list of names and study ID numbers in a separate location from the research forms.
- Limiting access to research data to the research team.
- Not including the participant's name on the typed transcript.
- Erasing audiotapes as soon as they are transcribed.
- Keeping research data on a password-protected computer.

Your information will be kept on written paper and/or digital file (jump drives) for up to three years within a locked cabinet in the faculty researcher's campus office. After three years, it will be destroyed by shredding and/or deletion. Signed consent documents and master lists of participant names and ID numbers will be stored in a separate locked cabinet in the faculty researcher's campus office. After three years, these items will be destroyed by shredding and/or deletion. NOTE: identifiers such as name, birth date, etc. will be deleted as soon as possible. NOTE: federal regulations require that signed consent documents must be kept for a minimum of three years after the study is closed.

The data from this research study may be published; but you will not be identified by name.

Agents of the University of Cincinnati may inspect study records for audit or quality assurance

purposes.

Your identity and information will be kept confidential unless the authorities have to be notified about abuse or immediate harm that may come to you or others.

What are your legal rights in this research study?
Nothing in this consent form waives any legal rights you may have. This consent form also does not release the investigator, the institution, or its agents from liability for negligence.

What if you have questions about this research study?
If you have any questions or concerns about this research study, you should contact John E. Stegall Jr. at stegalje@mail.uc.edu.

The UC Institutional Review Board reviews all research projects that involve human participants to be sure the rights and welfare of participants are protected.

If you have questions about your rights as a participant or complaints about the study, you may contact the UC IRB at (513) 558-5259. Or, you may call the UC Research Compliance Hotline at (800) 889-1547, or write to the IRB, 300 University Hall, ML 0567, 51 Goodman Drive, Cincinnati, OH 45221-0567, or email the IRB office at irb@ucmail.uc.edu.

Do you HAVE to take part in this research study?
No one has to be in this research study. Refusing to take part will NOT cause any penalty or loss of benefits that you would otherwise have. You may skip any questions that you don't want to answer.

You may start and then change your mind and stop at any time. To stop being in the study, you should tell John E. Stegall Jr. at stegalje@mail.uc.edu.

Agreement:
I have read this information and have received answers to any questions I asked. I give my consent to participate in this research study. I will receive a copy of this signed and dated consent form to keep.

Participant Name (please print) _____

Participant Signature _____ Date _____

Signature of Person Obtaining Consent _____ Date _____

Appendix D

Information Sheet—Phase III

IRB #: 2014-0434

Approved: 4/24/2015
Do Not Use After: 3/19/2016

Adult Consent Form for Research
University of Cincinnati
Department: College of Education, Criminal Justice, and Human Services
Principal Investigator: John E. Stegall Jr.
Faculty Advisor: Dr. Mary Brydon-Miller

Title of Study: Teacher Evaluation from the Perspectives of Teachers

Introduction:
You are being asked to take part in a research study. Please read this paper carefully and ask questions about anything that you do not understand.

Who is doing this research study?
The person in charge of this research study is John E. Stegall Jr. of the University of Cincinnati (UC) College of Education, Criminal Justice, and Human Services. He is the sole person conducting this research study and is being guided by Dr. Mary Brydon-Miller.

What is the purpose of this research study?
The purpose of this research study is to gain understanding of the perspective of K-12 public school classroom teachers regarding teacher evaluation.

Who will be in this research study?
About 100 to 200 people will take part in this study. You may be in this study if you are all of the following:

- A licensed K-12 public school teacher in Ohio actively assigned to a classroom.
- Age 21 and older.

What if you are an employee where the research study is done?
Taking part in this research study is not part of your job. Refusing to be in the study will not affect your job. You will not be offered any special work-related benefits if you take part in this study.

What will you be asked to do in this research study, and how long will it take?
You will initially be asked to respond to a series of 9 profession-related questions posed within an audiotaped, interview format. This interview will require approximately 45 to 60 minutes.

Subsequently, you may also be asked to take part in an additional, follow-up interview in which emergent themes coded within the transcript of your original interview will be shared for discussion and validation. This follow-up interview will require approximately 15 to 20 minutes. All interviews will take place on site within your school in a designated private room or otherwise unattended space. The information being collected in the interview may include your personal and professional perspectives on your work, training, and licensure as an educator.

Are there any risks to being in this research study?
It is not expected that you will be exposed to any risk by allowing your interview responses to be used in this research study. Some questions may make you uncomfortable. You can refuse to answer any questions that you don't want to answer. If you want to talk to someone because this research made you feel upset, the researchers can give you information about people who may be able to help you.

Are there any benefits from being in this research study?
You will not get any benefit because of being in this study. But, being in this study may help educators understand how teacher evaluation may support licensure.

What will you get because of being in this research study?
You will not be paid to take part in this study.

Do you have choices about taking part in this research study?
If you do not want to take part in this research study you may simply not participate. You will not be treated any differently. Audiotaping the interview is important to ensure accurate analysis of your comments. Therefore if you do not want to be audiotaped you should choose not to be in this research study.

How will your research information be kept confidential?
Information about you will be kept private by use of the following:

- A study ID number instead of the participant's name on the research forms.
- Keeping the master list of names and study ID numbers in a separate location from the research forms.
- Limiting access to research data to the research team.
- Not including the participant's name on the typed transcript.
- Erasing audiotapes as soon as they are transcribed.
- Keeping research data on a password-protected computer.

Your information will be kept on written paper and/or digital file (jump drives) for up to three years within a locked cabinet in the faculty researcher's campus office. After three years, it will be destroyed by shredding and/or deletion. Signed consent documents and master lists of participant names and ID numbers will be stored in a separate locked cabinet in the faculty researcher's campus office. After three years, these items will be destroyed by shredding and/or deletion. NOTE: identifiers such as name, birth date, etc. will be deleted as soon as possible. NOTE: federal regulations require that signed consent documents must be kept for a minimum of three years after the study is closed.

The data from this research study may be published; but you will not be identified by name.

Agents of the University of Cincinnati may inspect study records for audit or quality assurance purposes.

Your identity and information will be kept confidential unless the authorities have to be notified about abuse or immediate harm that may come to you or others.

What are your legal rights in this research study?
Nothing in this consent form waives any legal rights you may have. This consent form also does not release the investigator, the institution, or its agents from liability for negligence.

What if you have questions about this research study?
If you have any questions or concerns about this research study, you should contact John E. Stegall Jr. at stegalje@mail.uc.edu.

The UC Institutional Review Board reviews all research projects that involve human participants to be sure the rights and welfare of participants are protected.

If you have questions about your rights as a participant or complaints about the study, you may contact the UC IRB at (513) 558-5259. Or, you may call the UC Research Compliance Hotline at (800) 889-1547, or write to the IRB, 300 University Hall, ML 0567, 51 Goodman Drive, Cincinnati, OH 45221-0567, or email the IRB office at irb@ucmail.uc.edu.

Do you HAVE to take part in this research study?
No one has to be in this research study. Refusing to take part will NOT cause any penalty or loss of benefits that you would otherwise have. You may skip any questions that you don't want to answer.

You may start and then change your mind and stop at any time. To stop being in the study, you should tell John E. Stegall Jr. at (614) 795-8548 or stegalje@mail.uc.edu.

Agreement:
I have read this information and have received answers to any questions I asked. I give my consent to participate in this research study. I will receive a copy of this signed and dated consent form to keep.

Participant Name (please print) _____

Participant Signature _____ Date _____

Signature of Person Obtaining Consent _____ Date _____

Appendix E

Interview Protocol

Interview Questions Protocol (Notes)

1. (Icebreaker) What was it that inspired you to become a teacher?

2. Tell me about your first experience with teacher evaluation. (Probing: What was it like? How did it go?)

3. How do you feel about being professionally evaluated?

4. How do you feel teacher evaluation affects the teaching profession in general?

5. How do you feel teacher evaluation affects you professionally?

6. How do you feel teacher evaluation affects you personally?

7. How might you change teacher evaluation? (Probing: Describe. Can you be more specific? How would you accomplish this?)

8. What are your thoughts about the Ohio Teacher Evaluation System?

9. Is there anything else you would like to share or discuss?

Thank you!

Appendix F

Second Interview Protocol

Interview Questions Protocol (Notes)

1. (Icebreaker) What was it that inspired you to become a teacher?

2. Tell me about your first experience with teacher evaluation. (Probing: What was it like? How did it go?)

3. How do you feel about being professionally evaluated?

4. How do you feel teacher evaluation affects the teaching profession in general?

5. How do you feel teacher evaluation affects you professionally?

6. How do you feel teacher evaluation affects you personally?

7. How might you change teacher evaluation? (Probing: Describe. Can you be more specific? How would you accomplish this?)

8. What are your thoughts about the Ohio Teacher Evaluation System?

9. Is there anything else you would like to share or discuss?

Thank you!

Appendix G

Group Level Assessment Activity Prompts—Phase III

Opening Script: Each of the following six topics has been written onto one of the giant sheets of paper posted around this room. During this activity, within the allotted time, you will have the opportunity to write your thoughts about each topic on its designated giant sheet. You may write down anything that comes to mind about a particular topic. To provide ample space for all participants to share their thoughts, each topic appears on more than one giant sheet. At the end of the allotted time, we will collectively discuss what has been shared.

- Politics & Power
- Professional Outlook, Philosophical Approach, Working Environment, Culture & Change
- Evaluation Relevancy, Validity & Mentoring
- Evaluation Experiences, Impressions & Expectations
- Inequality
- Anything Else?

Appendix H

Self-Reflection Protocol—Phase IV

Overview: This self-reflection protocol served as a means for me to journal my personal thoughts regarding each theme within this research study. I then referred to these notes as a means to support my personal self-reflection analysis as part of this research study.

Theme:

Reflection—Connection to Professional Practice or Experience:

Appendix I

Codes

Teacher Perceptions of Evaluation

- Inspiration/Motivation
- Evaluation Experience & Impressions
- Educational Change

Evaluation Validity and Reliability

- Evaluation Validity & Reliability
- Politics/Politicians
- Privilege/Racism/Sexism

The Role of Mentoring

- Mentoring
- Professional Outlook, Support, Culture & Working Environment
- Perceptions of Teaching & Preparedness