March 2019

Leadership Advocacy, Ethical Negotiations, and Resignations to High-Stakes Assessment: A Pilgrimage

Jennifer Galbraith Canady

*University of South Florida, jncanady@mail.usf.edu*

Follow this and additional works at: https://scholarcommons.usf.edu/etd

Part of the Educational Administration and Supervision Commons, Education Policy Commons, and the Other Education Commons

**Scholar Commons Citation**


This Dissertation is brought to you for free and open access by the Graduate School at Scholar Commons. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of Scholar Commons. For more information, please contact scholarcommons@usf.edu.
Leadership Advocacy, Ethical Negotiations, and Resignations to High-Stakes Assessment: A Pilgrimage

by

Jennifer Galbraith Canady

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Educational Leadership Department of Educational Leadership, Counseling, Adult, Career, and Higher Education College of Education University of South Florida

Major Professor: William R. Black, Ph.D.
Vonzell Agosto, Ph.D.
Zorka Karanxha, Ed.D.
Arthur Bochner, Ph.D.

Date of Approval: March 8, 2019

Keywords: Educational Leadership, High-Stakes Testing, Testing Season, Culture of Measurement, Narrative, Pilgrimage

Copyright © 2019 Jennifer Galbraith Canady
DEDICATION

For Craven and Julia without whom I would be lost. Thank you for your enduring love, sacrifices, and support which made my doctoral journey and this pilgrimage possible.

Julia, you are the best thing that ever happened to me. Thank you choosing me to be your mother. Thank you for being my daughter and a wonderful young woman.

Craven, you are the best husband and friend any human could wish for, thank you for being my rock, listening to me, walking next to me, teaching me math, and holding my hand through thick and thin.

I love you both $\infty + 1$. 
ACKNOWLEDGEMENTS

First, I wish to thank my participants, Michelle, Nora, Cher, and Dave. Your openness and honesty made this pilgrimage possible. I wish to thank my peers who are also on the doctoral journey in Educational Leadership with me, particularly Carrie, Mike, and Drew. Thank you for your critical eyes and critical perspectives. I would be remiss if I did not thank two of the most understanding principals, Van Aryes and David New, for working with me and supporting my dreams over the last six years. I wish to thank my colleagues, Catherine, Antonio, Rory, Chris, Deanna, Dietrich, Kevin, and Brian for picking up my slack over the years. I am grateful to my friends: Kieonna, Barbara, Sheryl, Mary, Kris, Lisa, Jessica, and Gail for listening and checking on my progress throughout coursework and candidacy. Thank you to my committee for patiently walking me through the process of becoming a scholar. Dr. Agosto, you opened my mind to the possibilities of educational leadership and curriculum. Dr. Karanxha, you are gifted at telling it like it is in the most compassionate way possible. Dr. Bochner, I found my writer’s voice in your class. Thank you Dr. Black, other than being the best chair a doctoral student could ask for, I would not be here if not for your support and encouragement. Thank you for pushing me when I needed it, letting me cry when I needed it, and for your illegible scribbles. Most of all I want to thank my family: my parents, Nora and Dave, and my sister and brother-in-law, Erin and Daniel, who have been there for me in many ways during my many years in school, from babysitting to moral support I thank and love you.
# TABLE OF CONTENTS

List of Tables ........................................................................................................................................ vi

Abstract ........................................................................................................................................... vii

Prologue ........................................................................................................................................... 1

Chapter 1: Tale of My Journey’s Beginning ...................................................................................... 4
  Introduction ....................................................................................................................................... 4
  Assumptions I Bring to the Study ...................................................................................................... 7
  Purpose ........................................................................................................................................... 9
  Research Questions .......................................................................................................................... 9

High-Stakes Assessment Policy ...................................................................................................... 10
  High-Stakes Assessment Policy History ....................................................................................... 10
  Florida Context ................................................................................................................................. 12
  Terms Important in this Proposal .................................................................................................... 15

The Path of My Pilgrimage ............................................................................................................. 18

Conclusion ......................................................................................................................................... 21

Chapter 2: The Tale of the Literature Review ................................................................................... 22
  Introduction ....................................................................................................................................... 22
  Management Versus Leadership ....................................................................................................... 23
    Ethics and Closely Related Terms .................................................................................................. 24
  Ethics and Educational Leadership ................................................................................................. 25
    Ethic of Justice ............................................................................................................................... 30
    Ethic of the Profession .................................................................................................................. 31
    Ethic of Care .................................................................................................................................. 33
    Ethic of Critique ............................................................................................................................ 35
    Ethic of the Community ................................................................................................................. 37

High-Stakes Assessment Policy Implementation .......................................................................... 38
  The Culture of Measurement ......................................................................................................... 38
  The Narrowing of the Curriculum ................................................................................................... 39
  Testing Season ................................................................................................................................. 40
  Student-Level Performance ............................................................................................................ 40
    The culture of measurement ......................................................................................................... 40
    The narrowing of the curriculum ................................................................................................. 42
    The loss of instructional time ...................................................................................................... 45

The Intersection of Ethics, Educational Leadership, and High-Stakes Assessment .................. 46
  Destructive Impacts ....................................................................................................................... 46
  Neutral Impacts ............................................................................................................................. 49
| Chapter 4: Nora’s Tale: Here’s the Crappy Thing, We Have to Test, but the Good News is | 88 |
| Introduction | 88 |
| Culture of Measurement | 89 |
| Testing Season | 96 |
| Ethics of Implementation | 107 |
| Chapter 3: The Tale of Methods | 63 |
| Introduction | 63 |
| Epistemology | 64 |
| Historical Roots | 65 |
| Central Arguments of Interpretivism | 66 |
| Methodology | 68 |
| Rationale for Narrative Inquiry | 68 |
| Methods | 71 |
| Participant Recruitment and Selection | 71 |
| Nora | 72 |
| Cher | 72 |
| Michelle | 73 |
| Dave | 73 |
| People, places, and motivations | 73 |
| Story Collection | 74 |
| Focus group interviews | 74 |
| Electronic diaries | 75 |
| One on one interviews | 76 |
| Transcription | 76 |
| Story Selection | 79 |
| Story Analysis | 80 |
| Presentation | 81 |
| Reflections on limitations | 84 |
| Attending to the relationship | 84 |
| Internal review boards and consent | 86 |
| Conclusion | 87 |
## Chapter 6: Michelle’s Tale: It’s a Fiscal Testing Year

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>141</td>
</tr>
<tr>
<td>Culture of Measurement</td>
<td>142</td>
</tr>
<tr>
<td>Testing Season</td>
<td>150</td>
</tr>
<tr>
<td>Ethics of Implementation</td>
<td>155</td>
</tr>
<tr>
<td>Conclusion</td>
<td>165</td>
</tr>
<tr>
<td>Vignette</td>
<td>166</td>
</tr>
</tbody>
</table>

## Chapter 7: Dave’s Tale: The CYA Mentality Just Rolls Downhill

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>169</td>
</tr>
<tr>
<td>Culture of Measurement</td>
<td>171</td>
</tr>
<tr>
<td>Testing Season</td>
<td>178</td>
</tr>
<tr>
<td>Ethics of Implementation</td>
<td>189</td>
</tr>
<tr>
<td>Conclusion</td>
<td>194</td>
</tr>
</tbody>
</table>

## Chapter 8: Tale of My Reflection on the Narrative Collective

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>196</td>
</tr>
<tr>
<td>Purpose</td>
<td>196</td>
</tr>
<tr>
<td>Research Questions</td>
<td>197</td>
</tr>
<tr>
<td>The Implicated Advocate</td>
<td>197</td>
</tr>
<tr>
<td>Culture of Measurement</td>
<td>198</td>
</tr>
<tr>
<td>Accountability</td>
<td>198</td>
</tr>
<tr>
<td>High-Stakes Assessment is Big Business</td>
<td>200</td>
</tr>
<tr>
<td>Narrowing of the Curriculum and Loss of Instructional Time</td>
<td>200</td>
</tr>
<tr>
<td>Labeling</td>
<td>202</td>
</tr>
<tr>
<td>Resegregation</td>
<td>202</td>
</tr>
<tr>
<td>Experiencing the Testing Season</td>
<td>203</td>
</tr>
<tr>
<td>School Administrator Performance</td>
<td>203</td>
</tr>
<tr>
<td>Emotional labor</td>
<td>204</td>
</tr>
<tr>
<td>Compliance and agency</td>
<td>205</td>
</tr>
<tr>
<td>Resignation</td>
<td>207</td>
</tr>
<tr>
<td>Resistance</td>
<td>207</td>
</tr>
<tr>
<td>Obligation versus inspiration</td>
<td>208</td>
</tr>
<tr>
<td>Student Performance</td>
<td>208</td>
</tr>
<tr>
<td>Pressure to perform</td>
<td>208</td>
</tr>
</tbody>
</table>
Appendix B

Appendix

Chapter 9: The Tale of Narrative Possibilities

Introduction .......................................................................................................................... 221

Narrative Possibilities for Practice .................................................................................. 223
  Opt Out, Come Back! ........................................................................................................ 223
  De-Tracking ..................................................................................................................... 224
  Paper and Pencils .......................................................................................................... 225
  Experimental Schools .................................................................................................... 225
  Bring Back the Local ....................................................................................................... 224
  Give Students a Purpose ................................................................................................. 226
  Small Acts of Resistance ............................................................................................... 226
  Social Justice Advocacy Leadership ............................................................................... 227

Narrative Possibilities for Scholarship and School Leader Preparation ......................... 227
  Allocation of Resources ................................................................................................. 228
  Researching and Teaching Advocacy and Social Justice Leadership .......................... 229
  Students with Disabilities ............................................................................................... 229
  Students Who Are English Language Learners .......................................................... 230
  Scholarship in Ethical Educational Leadership ............................................................. 230
  Ethic of Care ................................................................................................................... 230
  My Implications for Future Research .......................................................................... 231

Conclusion ......................................................................................................................... 231

Epilogue: Inquiry as Pilgrimage .......................................................................................... 234
  The Journey .................................................................................................................. 234
  The Motivations ............................................................................................................ 235
  The People and Places ................................................................................................. 237
  Implicated Advocates ................................................................................................. 237
  The Sacred Destination ............................................................................................... 238

Conclusion ......................................................................................................................... 238

References .......................................................................................................................... 240

Appendix A: Interview Questions ...................................................................................... 283

Appendix B: University IRB Approval Letter ................................................................... 285
Appendix C: University Continuing Review Approval Letter .......................................................... 287
Appendix D: Informed Consent Form .......................................................................................... 289
LIST OF TABLES

Figure 1: Professional Standards for Educational Leaders .................................................32

Figure 2: Strategies to Attend to the Relationship ...............................................................85
ABSTRACT

The purpose of my study is to explore the stories of the ethical tensions K-12 educational administrators navigate when implementing high-stakes assessment policy in a culture of measurement during testing season. Some educational leaders, in particular K-12 school and district administrators, struggle with the tensions existing between their own personal belief systems, organizational dilemmas, and the requirements of enacting high-stakes assessment policies. Using narrative inquiry as method, I collected and analyzed four school administrators selected who expressed frustration with enacting high-stakes assessment policies. The participants include a middle principal, a middle school assistant principal, a high school assistant principal, and a middle school testing administrator. Through their stories, I raise questions about the purpose of high-stakes assessment and the impact of policies at the implementation level and the impact of policies on the daily practices of school administrators. I illustrate how school administrators negotiate these tensions or have resigned themselves to accept what they cannot change. The findings include the reallocation of capital and human resources during testing season and the abundant loss of instructional time. Additional findings comprise of the juxtaposition between compliance and agency school administrators’ experience, and the nuanced ways schools and districts work to game the system of accountability. Findings also involve how educational leaders work within the boundaries of high-stakes assessment, and at times, find small spaces to resist high-stakes assessment implementation. The study shines light on how they accept the differences between their own personal ethics and the requirements of their jobs. Implications include the need for more scholarship surrounding the allocation and
reallocate of resources in public schools during testing season, and the impact high-stakes assessment implementation has on vulnerable populations of students, especially students with disabilities, and students who are English language learners. The participants' stories revealed aspects of high-stakes assessment policy implementation, which impact the lives of students and educators that have not been explored in great depth. I argue for centering ethical leadership and the need for training and socializing school leaders to be social justice advocates for their students even while they are also implicated in systems of accountability. Finally, I also present the inquiry as a pilgrimage metaphor as journey toward not only understanding how school leaders grappled with ethical dilemmas associated with implementing high-stakes assessment in a culture of measurement during testing season, but also a journey to understand my place, as a school administrator, in this ethical conundrum.
PROLOGUE

Hyre Bygynneth the pilgrymage:

Whan that Aprille with his shoures soote,

The droghte of March hath perced to the roote,

And bathed every veyne in swich licour

Of which vertú engendred is the flour;

Whan Zephirus eek with his swete breeth

Inspired hath in every holt and heeth

The tendre croppes, and the yonge sonne

Hath in the Ram his halfe cours y-ronne,

And smale foweles maken melodye,

That slepen al the nyght with open ye,

So priketh hem Natúre in hir corages,

Thanne longen folk to goon on pilgrimages,

And palmeres for to seken straunge strondes,

To ferne halwes, kowthe in sondry londes;

And specially, from every shires ende

Of Engelond, to Caunterbury they wende,

The hooly blisful martir for to seke,

That hem hath holpen whan that they were seeke. (Chaucer, 1987, p.23)
So begins Chaucer’s famous medieval poem filled with the tales of old told by pilgrims on their way to Canterbury to visit the gravesite of St. Thomas Becket. I find this introduction inspirational for its literary endurance, its ability to situate the reader in place and time, and finally because I adore the metaphor of life as pilgrimage. Stoddard (1997) suggests pilgrimages are journeys or movements to sacred destinations by people with various motivations. Further, Diaz (2010) suggests pilgrimages, especially stories of pilgrimages, are quest narratives wherein the protagonist searches for spiritual awakening or knowledge. In this study, my pilgrimage is motivated by my desire to understand ethical dilemmas associated with the implementation of high-stakes assessments. My journey or movement to my sacred destination, or spiritual awakening, takes me down a path with nine stops. At each stop I gather knowledge, I learn from existing literature and methods and then I engage school leaders in conversations revealing their experiences implementing high-stakes assessments and connect these experiences with knowledge collected by previous travelers. My participants are the people and places on my journey. It is their motivation moving me forward on the path toward my sacred destination. My sacred destination is my attainment and creation of knowledge, to be enlightened by my findings and to share the story of my pilgrimage with others in hopes of creating new ways of understanding, much like a spiritual awakening, culminating with the completion of this dissertation.

This study contains nine pilgrimage tales. First, to begin my journey, I will begin by introducing my pilgrimage toward ethical consciousness as it concerns high-stakes assessment. Second, I introduce the literature informing my journey. Next, I tell the tale of the methods chapter, which outlines the methodology guiding how I proceeded to learn how school administrators storied their experiences working through ethical dilemmas related to high-stakes
assessments policy. I then make four more stops on my journey to collect and present the narrative motivations of each of participants in chapters four through seven. My journey continues in chapter eight with my reflection on the collective narratives. I end my journey, but perhaps never the pilgrimage, in chapter nine with narrative possibilities. By studying, collecting, and presenting these tales it is my hope aspiring and veteran school and district administrators re-examine their role in the praxis of high-stakes assessment as ethical practitioners.
CHAPTER 1: THE TALE OF THE JOURNEY’S BEGINNING

“For a minute there
I lost myself, I lost myself
Phew, for a minute there
I lost myself, I lost myself“ (Yorke, Greenwood, O’Brien, Greenwood, & Selway, 1997)

**********

Of the 42 days of school between April 2 and May 30, 2014 I proctored approximately 35 high school-level assessments, including the Florida Comprehensive Assessment Test (FCAT) for Reading, and FCAT Reading Retakes, End of Course (EOC) exams for biology, algebra one, geometry, and world history, Advanced Placement (AP) exams for environmental science, calculus, and world history, Florida’s Postsecondary Readiness Test (PERT), the Comprehensive English Language Learning Assessment (CELLA), and various district course exams. This does not include other proctoring I did during the year like administering the Florida Assessment in Reading (FAIR) in September and December/ January, FCAT Reading Retakes in October, algebra and geometry EOC retakes in November and December, district course exams in January, FCAT Writes in February, and the Scholastic Assessment Test (SAT) in March.

Many times I saw the same students over and over again in the various testing sessions I proctored, especially struggling students. For example, I saw one student for two FAIR assessments in the early fall and late winter, the FCAT Reading Retake, and the algebra EOC in the fall. Then in the spring, I saw the same student for exams in January, FCAT Writes, SAT, FCAT Reading retakes (again), algebra and geometry EOCs (again), American History EOC and
PERT. This means I saw this one student take a minimum of 11 high-stakes assessments over the course of one school year, and this is only taking into account one student and only the tests I proctored.

I distinctly remember one day in particular when I was standing in the media center behind a bank of fifty computers watching students take PERT. PERT was administered to 11th grade students who have not demonstrated they are college and career ready via other assessments, such as SAT, ACT, or FSA. I was policing the room to make sure students were not cheating, and I was in a position to see all the rows of students from left. As I observed the 25 students testing, I was struck by their body language. These young, beautiful people with their whole lives ahead of them looked defeated. Shoulders were drooping and backs were either hunched over or slouched down in their chairs. They had taken so many tests and were worn down by being told over and over again they were not good enough because of their assessment scores. I often see or even do things I do not agree with, but this scene replays in my mind when I think of ethics surrounding high-stakes assessment in my former role as a teacher leader as well as my current role as an assistant principal.

In retrospect, I realize this was the beginning of my journey on my pilgrimage or my, “call to adventure,” because, “destiny had summoned,” me to this quest (Campbell, 2008, p. 48). According to Campbell (2008), the quest, “may begin as a mere blunder…one may be only casually strolling, when some passing phenomenon catching the wandering eye and lures one away from the frequented paths of man (sic)” (p. 48). My quest or pilgrimage to gain an understanding how we, school administrators, enact policies we do not agree with personally and even disagree with ethically is rooted in the moment I saw those students’ shoulders slumped over in defeat. My journey, albeit a metaphoric one, started then and has led me on a path of
inquiry represented in this study. This dissertation is an inquiry as pilgrimage, embodying my metaphorical journey to understanding how school leaders carry out aspects of their jobs they do not believe are right for students.

Although I was already uncomfortable with high-stakes assessment and aware students were taking too many, this experience left me personally conflicted. In my teacher leadership role of reading coach, my job was dependent on high-stakes assessments, as are certain aspects of my current role as an assistant principal, but as educator, a scholar, and human I think the amount and frequency of assessments and the rewards and consequences associated with mandated high-stakes assessments for students, teachers, administrations, schools and districts include elements which are unethical. Thus, I wondered if other school-level administrators, principals and assistant principals, have similar ethical concerns about high-stakes assessment policies. It is my responsibility as a scholar of educational leadership to find ways to shed light on the ways the harmful aspects of high-stakes accountability policies currently in place impact students’ educational experiences negatively.

Educational administrators often face dilemmas and tensions due to incongruence between their personal belief systems and the requirements or demands of school leadership positions (Larson, 2014; McGhee & Neslon, 2005; Muzaliwa & Gardiner, 2014; Vernaza, 2012). In particular, ethical issues faced by school and district administrators related to the demands of high-stakes assessment are not uncommon in the literature. For example, McGhee and Nelson (2005), both former school principals, expressed living with the dilemmas associated with high-stakes assessments and meeting the needs of their students. Likewise, Vernaza (2012) called for educators to have a voice concerning high-stakes assessment policies. Larson (2014) argued for “in-depth ethnographic and case study research” to begin to understand the nature of unjust
institutional systems of achievement and argues for a “capability” approach to transform education (p. 163). Additionally, Muzaliwa and Gardiner (2014) call for the use of narrative inquiry as a method that can “advance” educational leadership (p. 184). My journey to my sacred destination aims to address these calls in the literature.

Assumptions I Bring to the Study

As an educational leader, I have wrestled with doing what I feel is best for students and enacting high-stakes assessment practices I find harmful to students but were required as part of my job. Perhaps, there are educational leaders who do not struggle ethically with high-stakes assessment policies and may, in fact, support high-stakes assessment policies. Yet, there are some educational administrators who do wrestle with these same ethical dilemmas. Further, there are likely educational administrators who have not considered high-stakes assessments policies at all.

My assumption is, by sharing the ethical struggles associated with implementing high-stakes assessment, others who are leading education may be encouraged toward deep reflection and consideration of belief systems, which has the potential to encourage dialogue about the purpose and nature of high-stakes assessment. For those who accept accountability at face value, it is more important to expose them to narratives about those who struggle than it is to share the narrative with those who are like-minded.

As a school administrator, I have long struggled with aspects of my job that required me to implement accountability policies, especially high-stakes assessment policies. I chose to become an educator because I wanted to help struggling students succeed because I was a struggling student. I did not read until I was in third grade, and today I would have likely been retained and forced to repeat third grade. I hated school so much I spent much of my time in
high school skipping classes. I missed so much school I took five years to complete high school rather than four. I was lucky I was in school before federal and state systems of accountability existed. Today, I would have been encouraged to go to adult school or a charter school by the end of my sophomore year. If my life’s trajectory were based solely on my success on high-stakes assessments, I would have been written off before I learned to read. I hoped by becoming an educator I could help students like me, but instead, I find myself implicated in the very systems of accountability that would have negatively impacted me had they been in place when I was a student. Thus, this inquiry as pilgrimage is more than a journey to understand how school leaders implement high-stakes assessment policies they do not always agree with, but it is also a metaphorical journey to help me understand the contradictions of wanting to advocate for students like me, but also how I deal with being implicated in systems of accountability as an agent of the state.

Using a narrative inquiry approach to analyzing the ways in which educational leaders deal with the dilemmas between accountability and a sense of ethics they bring to their day-to-day jobs may be beneficial to those aspiring or veteran educational leaders struggling with similar dilemmas. Furthermore, using the metaphor of a pilgrimage, to organize the findings into pedagogical tales can be useful in the preparation of educational administrators.

Since I took Dr. Bochner’s *Narrative Inquiry* course several years ago, I have been passionate about narrative. In his class, I found a narrative as a means of sharing human experiences through storytelling. Narratives can be found in all facets of our everyday lives, such as in conversations with others where we tell stories of things that happened to ourselves or others, and on the radio, television, movies, social media, and news sources tell stories of current and historical events (Bochner, 2014). My undergraduate degree in English literature confirms
these ideas. *The Odyssey* (Homer, 2009) survived centuries in the oral tradition, Shakespeare’s plays are still read and performed 400 years after his demise, and literature written in response to classic works of English literature from the perspective of the antagonist, like in the novel *Wide Sargasso Sea* (Rhys, 1966), are popular for a reason: humans respond to stories. Narratives encourage readers to engage in the moral, intellectual, and emotional experiences of other humans (Bochner, 2014). Inspired by the story of a pilgrimage written 600 years ago, my journey is presented here as a collection and presentation of narratives concerning the ethical experiences and motivations of four school administrators.

**Purpose**

The purpose of this study is to explore the stories of the ethical tensions K-12 educational administrators navigate when implementing high-stakes assessment policy in a culture of measurement during testing season. Some educational leaders, in particular K-12 school and district administrators, struggle with the tensions existing between their own personal belief systems, organizational dilemmas, and the requirements of enacting high-stakes assessment policies.

**Research Questions**

My research questions are: 1) In what ways do K-12 educational administrators navigate potential ethical dilemmas between personal ethical obligations to students versus organizational directives associated with the implementation of high-stakes assessment policy? 2) How do K-12 educational administrators experience and live the testing season? 3) How does the culture of measurement play out in the stories of K-12 educational administrators?

**High-Stakes Assessment Policy**
High-stakes assessment refers to any state-mandated test students are required to take which are tied to school ratings, teacher performance ratings, and whether or not a student is deemed “proficient” in reading, math and/or science. Beginning in the 1990s in states like Texas and Florida, and then with the No Child Left Behind (NCLB) Act of 2001 and until December of 2015, state and federal governments mandated students be proficient in reading and math. Although students did not have to demonstrate proficiency in each and every grade, students had to be proficient in benchmark years. However, NCLB did mandate students take a state assessment annually. In December 2015, President Obama signed the Every Student Succeeds Act (ESSA, 2015) into law. Taking the place of NCLB, the new federal education law made some major changes but the requirement of an annual assessment remained. In many states, such as Florida, Texas, and North Carolina, students who are not proficient on state reading assessments may be retained in third grade, may be required to take additional reading courses, and must demonstrate proficiency by graduation on the state assessment or another assessment the state has determined an acceptable concordance assessment, such as ACT or SAT, or the student receives a certificate of completion rather than a standard diploma (The Florida K-20 Education Code, 1995).

**High-Stakes Assessment Policy History**

The testing of students as a means to control schools and schooling is not a new concept in the United States. Since the turn of the twentieth century, the idea of making schools more efficient based on the ideas of Frederick Taylor’s system of scientific management took hold (Callahan, 1962). Making schools efficient included measuring students. Callahan (1962) describes a school superintendent in a large school district in Michigan hiring substitutes to give over 50,000 examinations to students in 1916. The superintendent did this, “to locate the strong
and weak points in the schools” (Callahan, 1962, p. 103). The tests were used to measure whether or not the school district’s budget was being appropriately spent. According to Bowles and Gintis (1976), “by the end of the 1920s, the relationship between social background and a child’s chances of promotion or tracking assignments would be disguised—though not mitigated much—by another reform: objective educational testing” (p. 195). The use of efficiency policy via testing as a means to manipulate the allocation of resources based on test scores continues to this day.

Over the course of the next 80 years or so, various forms of testing were implemented on state and local levels in many parts of the country. However, the “Texas miracle” was the foundation of what would become federal education policy, NCLB, dictating all children would be tested (Haney, 2000; Smith, 2004; Stobart, 2008). Texas was a model for education policy reform due to reported improved test scores on the state’s high-stakes Texas Assessment of Academic Skills (TAAS) and decreased dropout rates (Haney, 2000). Even though the miracle would prove to be smoke and mirrors created through “arbitrary and discriminatory” cut scores on tests, the retention of students in grade nine as to not impact grade 10 TAAS scores, and the apparent fraudulent reporting of dropout rates by school districts (Haney, 2000, p. 23). In 1999, Sacks (1999) argued the rhetoric of accountability was, “highly effective, though remarkably little good evidence exists that there is any educational substance behind the accountability and testing movement” (p. 155). Nonetheless, the model for national school reform policy was naturally precipitated by the election of Texas governor George W. Bush to the presidency (Vasquez Helig & Darling-Hammond, 2008). The model included the public rating of schools and school districts by the state board of education, rewards for schools and districts performing well on TAAS, and sanctions for the schools and districts not performing well on TAAS.
With the No Child Left Behind Act of 2001, the federal government mandated students be proficient in reading and math in benchmark years. NCLB mandated students take a state assessment annually even in non-benchmark years. Each state would create and maintain its own assessments as well as determine what “proficiency” would be mean for the students in the individual state (Cuban, 2004). Like its model in Texas, the federal legislation also encouraged states rate and rank schools and school districts and create systems of rewards and sanctions, in addition to monetary sanctions states would face from the federal government in the form of Title 1 funding. In 2015, ESSA replaced NCLB and some of the requirements of NCLB were loosened, but ESSA still requires annual testing of students. In most cases, little has changed at the state level. States still rank and publish the test data for schools and most states still have systems of rewards and sanctions in place.

**Florida Context**

My journey takes place in two large school districts in the state of Florida; thus, it is essential to have a sense of the Florida context when it comes to public schools and accountability policies in the state. Perhaps most important when considering historical aspects of Florida’s K-12 public education are segregation and desegregation. Indeed, any research into Florida education policy should acknowledge the dual system of education in place in the state until the 1970s and the connections desegregation has to accountability policy in Florida today.

In fact, Florida passed laws allowing for segregated schools immediately after the Civil War (Cobb-Roberts & Shircliffe, 2007). Cobb-Roberts and Shircliffe (2007) explain, not only was Florida one of the first states to enact legislation creating a dual system of education, the state created ensuing laws making it illegal for White and Black children to attend school together, laws requiring White teachers to only teach White students and Black teachers to only
teach Black children, and laws for a dual tax system, which perpetuated subsidy inequalities between the two schools systems. In one of the districts participants from this study work in, the “per capita school expenditures” were approximately $22.00 for each White student versus approximately $9.00 per Black student for the school year 1909-1910 (Cobb-Roberts & Shircliffe, 2007, p. 23). A similar disparity existed between the salaries for White teachers and Black teachers, with Black teachers making approximately half the monthly salary as their White counterparts (Cobb-Roberts & Shircliffe, 2007). This remained the status quo through the Brown v. Board of Education decision in 1954 and beyond.

Much of the power and decisions regarding K-12 education are guided or decided at the state level with the Florida Department of Education. Local power began to be usurped in the 1970s when the state legislature crafted the Florida Education Finance Program (FEFP), which was created to equalize school funding across school districts in the state (Dorn & Michael, 2007). The new law addressed the funding imbalances, which set the stage for accountability due to the change in monetary power structures. According to Dorn and Michael (2007), originally local accountability to the state was focused on local education spending, but eventually the state shifted its focus from how local districts were spending to the educational results districts were getting. Dorn and Michael (2007) also note that Florida has concentrated control at the state level, but the responsibility for educational outcomes was and is still diffused to the local levels. This means the state imposes consequences on districts, schools, teachers, and students but has no obligation to provide resources, such as capital resources, at any level.

Accountability focused on educational outcomes began in Florida when, then Governor, Jeb Bush’s A+ plan went into effect in 1999 (Lee, Borman, & Tyson, 2007). Although under the previous Governor, Lawton Chiles, state standards were created (Lee et al., 2007). The A+ plan
created a system of accountability whereby schools and districts would receive a school grade based on student performances on state-mandated assessments, the Florida Comprehensive Assessment Test (FCAT) in reading and math. Schools that did well on the state grading scale were rewarded with A+ money, which is still in effect today. Initially, failing schools received no additional funding or resources, but over the years the state has begun to provide supports to assist these schools.

In addition to standardized testing as a means of grading schools, the A+ plan introduced merit pay for teachers, vouchers for students with disabilities and for students at schools with two years of grades of F on the state school grading system, as well as school choice for students at schools with two years of grades of F on the state school grading system, and higher requirements and tests for teacher licensing (Borman & Dorn, 2007). Bush touted his A+ Plan as the plan to close the achievement gap between White and Black students by holding schools and districts accountable for the performance of all students on high-stakes assessments, but at the same time Bush also denied the impact segregation had on school performance or school law policy or any “correlation between poverty or minority rates and test scores” (Eitle, 2007, p. 118). Systematically grading schools and labeling teachers and students as failures via high-stakes assessments has merely served to perpetuate achievement gaps, with Black and Brown students and students living in poverty not performing as well on high-stakes assessments as their more affluent and White peers.

Denying the impact poverty has on learning and on test scores denies a well-researched reality (Au, 2009a, 2009b; Dell’Angelo, 2016; Hair, Hanson, Wolfe, & Pollak, 2015, Noguera,

---

1 The A+ plan also introduced changes not outlined here to Florida’s higher education system. See Borman and Dorn
Darling-Hammond, & Friedlaender, 2015; Orfield & Lee, 2005; Rodriguez, Amador, & Tarango, 2016; Tate IV, 2008). Since Bush, Governors Charlie Crist and Rick Scott have done little to lessen the impact of Bush’s legacy on educational policy in the state. Indeed, over the last twenty years, the amount of high-stakes assessments administered to Florida students increased to include writing every year from grades 3-10, a civics end of course exam, a biology end of course exam, an United States history end of course exam, a geometry end of course exam, and an algebra two end of course exam. The algebra two exam was eventually done away with under Governor Scott to slightly reduce the amount of high-stakes assessments a public school student is subjected to in the state of Florida. Additionally, Florida initiated computer-based testing (CBT) in the early 2010s for high-stakes assessments administered to students in middle and high schools.

**Terms Important in this Dissertation**

**Ethics.** A term essential to this journey is ethics. Shapiro and Gross (2008) delineate the concept of ethics as having:

Numerous meanings over time. Initially, it came from the Greek word ethos, which meant customs or usages, especially belonging to one group as distinguished from another. Later, ethics came to mean disposition or character, customs, and approved ways of acting. (p. 20)

According to Strike (2007), ethics is concerned both with, “what is good? and what is right?” (p. 9). Thus, throughout this pilgrimage I interrogate what is good and what is right via my participants’ narratives.
**High-stakes assessment.** High-stakes assessment refers to any state-mandated test students are required to take that is tied to school ratings, teacher performance ratings, and whether or not a student is deemed “proficient” in reading, math and/or science.

In this study, I include the following assessments: Florida Standards Assessment for English Language Arts (FSA ELA), Florida Standards Assessment for Math, the Civics, Biology, Algebra, Geometry, and US History End of Course Assessments (EOCs), PERT, SAT, ACT, and the Florida Standards Alternative Assessment (FSAA). Although not technically a high-stakes assessment since it is not factored into school grade, WIDA, Florida’s English proficiency assessment is a state requirement for all ELL students given in accordance with ESSA. WIDA will determine student placement in classes and factors into the VAM scores of ELL teachers. For the purposes of this study, due to the high impact of WIDA, I have included it as part of the research.

**School administrators.** Throughout this journey, I use the terms educational administrators, educational leaders, educational administrators, and school leaders interchangeably. In this dissertation, these terms refer to K-12 public school administrators, such as principals and assistant principals. Although there may be content, which could be useful for higher-education leaders, or even teacher-leaders, this study focuses on school-level leadership.

**Agency.** Agency for the purposes of this proposal refers to freedom or autonomy for educational leaders to make decisions for their school or district organization.

**Obligation.** In this dissertation, I refer to two types of obligation. One type of obligation is organizational in which the educational leaders feels like she is obligated to take a specific action is dictated by someone or something in the organization. The second kind of obligation is
a personal moral obligation an educational leader may feel toward students or teachers within the part of the organization they oversee.

Choice. Choice in this sense refers to the degree to which educational leaders feel they have choices in the implementation of high-stakes assessment.

Dependence. This proposal refers to dependence as educational leaders relying on others above them in the hierarchy of the organization.

Testing Season. Testing season is a term used in K-12 education to demarcate a period of time, typically during the spring, when most high-stakes standardized tests are administered to students.

Over-testing. A term used in K-12 education to describe a frequent and large amount of assessment to be administered to the same students repeatedly.

The culture of measurement. A culture of measurement is characterized by a focus on the assessment students take versus the instruction they receive daily or, in other words, “the overemphasis on the process…and the test-taking competence” (Padilla, 2005, p. 250). Padilla (2005) describes the high-stakes assessment crusade as a “culture of measurement” (p. 249) focused on student test scores, the ranking of students, teachers, and schools, and the narrowing of curriculum, which focuses on basic reading, writing, math, and test-taking skills. According to Padilla (2005), the culture of measurement only views students in the abstract. Thus, the only way to determine if the abstract student is learning is through “presumed norm” referenced tests (p. 250).

The narrowing of the curriculum. Nichols and Berliner (2007) and Au (2009b) describe the narrowing of curriculum as focusing solely on tested subjects, such as reading and
math, to the exclusion of other non-assessed subjects, such as the arts, but it can also relate to students being denied access to curriculum.

**The loss of instructional time.** Berliner (2007) defined instructional time as, “academic learning time” (p. 3). Thus, the loss of academic learning time is the loss of instructional time. For the purposes of my study, I consider any student loss of instruction due to high-stakes assessments as the loss of instructional time.

**The Path of My Pilgrimage**

Throughout this dissertation, I use the terms journey, path, and quest interchangeably. Unlike Chaucer, I am not on this pilgrimage with others. This journey is my own, but I intentionally look to others to help further my understandings. I collect stories or tales from existing literature and methods that have been told by those who have traveled similar paths before me. Along my metaphorical journey, I also stop and collect stories from other school administrators. I refer to each collection of stories encapsulated in a chapter as a tale. My inquiry as pilgrimage begins here in chapter one by situating the reader with the experiences bringing me to this path. I further delve into my motivation for starting the journey, as well as examine my assumptions and background. I present my research questions and purpose. Then I provide the federal and state policy and historical accounts of accountability legislation to provide the reader with the context of the power and pervasiveness of high-stakes assessment policy. I define terms relevant to my journey and describe the structure and the content of the pilgrimage.

My journey continues with chapter two, which is focused on exploring the knowledge and research related to my study. The literature reviewed partly represents the sacred destination I long to arrive at when my pilgrimage ends: attainment and creation of knowledge. I explore
several paths of research in chapter two, including management versus leadership, ethics and educational leadership, ethical paradigms, high-stakes assessment policy implementations, the intersection of ethics, educational leadership, and high-stakes assessment, emotional labor and ethics, authentic leadership, performativity, social justice and advocacy leadership, as well as gaps within the literature reviewed.

The third stop on my journey brings me to my methods for collecting and analyzing the stories of my participants. In chapter three I provide an overview of the epistemology undergirding narrative inquiry and a rationale for utilizing narrative inquiry as my methodology. I present the methods I used to obtain participants and briefly introduce each of the four participants. Next, I describe how I collected and analyzed their narratives, including the limitations of the study.

My fourth stop brings me to my first participant. Each participant represents a stop on my journey and has an individual chapter devoted to them and their stories. My participants are the people and places on my journey. It is their motivation moving me forward on the path toward my sacred destination. Throughout each chapter, I weave relevant research into the tale as the attainment knowledge represents the sacred destination of my pilgrimage. Nora is my first participant. Nora has been a friend of mine for twelve years. Nora has been a principal for thirteen years, six years at Crimson Middle School and the last seven at Blue Middle School. Nora’s tale begins with how she works within a culture of measurement, followed by how she experiences testing season, and finally Nora’s motivations and struggles with the ethics of implementing high-stakes assessment policy. Chapter four concludes with a vignette about one of my experiences with testing season.
The fifth stop on my journey brings me to chapter five: Cher’s tale. Cher is also a friend of mine. Cher is a senior assistant principal at Brown Middle School. She was a district resource teacher for reading before becoming a school administrator four years ago. Like Nora, Cher’s tale begins with how the culture of measurement plays out around her. Her chapter then explores how Cher lives the testing season, and concludes with how her motivations and tensions impact her implementation of high-stakes assessment policy. I conclude chapter five with a vignette about how I work in a culture of measurement.

The sixth chapter represents the sixth stop on my travels: Michelle’s tale. Michelle is also a senior assistant principal. At the time of story collection, Michelle worked with me at Orange High School. She was the senior assistant principal at Orange High School for five years, and a junior assistant principal for three. Michelle begins her tale by describing her experiences working in a culture of measurement. Next Michelle’s narrative delves into the impact of testing season on the school, teachers, students, and her personal life. Michelle’s tale also includes her motivations and ethical dilemmas surrounding the implementation of high-stakes assessment policy. Chapter six ends with a vignette about one of my experiences implementing high-stakes assessment.

My seventh chapter introduces the reader to Dave’s tale. Dave is a testing coordinator in a neighboring school district. A testing coordinator is a teacher whose sole responsibility is to plan, implement, and carry out high-stakes assessments. Although Dave is technically a teacher, his role as a testing coordinator involves primarily administrative tasks. Dave’s tale starts with him describing the pervasiveness of the culture of measurement. Next, he shares how he experiences testing season, and lastly, Dave explains his motivations and ethical struggles with implementing high-stakes assessment policy.
The next stop on my journey is my tale of reflecting on the collective narrative of my participants. In chapter eight, I reflect on the collective narrative of all four of my participants’ experiences with the culture of measurement, testing season, and the ethics of implementation. I connect relevant knowledge and draw conclusions about my findings.

The final stop of my travels is chapter nine: the tale of narrative possibilities. In this tale, I explore the narrative possibilities for practice, scholarship, and school leader preparation. It is on this last stop I consider the future of my journey to my sacred destination: the attainment and creation of knowledge. I hope to have contributed to the knowledge surrounding the nature and purpose of high-stakes assessment.

Conclusion

Using narrative inquiry, I explore the struggles some educational leaders, in particular K-12 school administrators, have with the ethical dilemmas existing between their own personal belief systems, organizational dilemmas, and the requirements of enacting high-stakes assessment policies during testing season. Thus, the purpose of my journey is to explore the stories of the ethical dilemmas K-12 educational administrators navigate when implementing high-stakes assessment policy during testing season. In the following chapter I present the literature related to this purpose.
CHAPTER 2: THE TALE OF THE LITERATURE REVIEW

“With your feet in the air and your head on the ground

Try this trick and spin it, yeah

Your head will collapse

But there's nothing in it” (Francis, 1987).

**********

“Why is this so hard?” I ask no one in particular. I am sitting in a tiny room in my doctoral department’s office suite. They refer to the room as a conference room, but it is more like a two-person meeting room. I have hunkered down in the room almost every day for the past three weeks trying to attain knowledge related to my journey. It is a small, windowless grey room lined with shelves full of empty three-inch binders. In front of me are stacks of books I have borrowed from the university’s library. I have so many books out now I had to find out the maximum number of books a doctoral candidate is allowed to check out, which is 200. I am fast approaching this number, but the document I am writing does not reflect this. Historically, I have agonized about writing almost everything. I guess it is my process but recognizing it is not very helpful and feeling like I am in a time crunch does not make my process helpful either.

Pilgrimages do not always start smoothly or easily I remind myself. Chaucer’s pilgrims may have started strong, but they never make it to Canterbury, their sacred destination. Finishing this dissertation, attaining and creating knowledge, is reaching my Canterbury, but each stop and story are integral to me reaching my sacred destination.
Suddenly, Dr. Karanxha appears at the door. “How are you?” she asks. I explain I have been gathering a great deal of research and my struggle to move from collecting knowledge to writing about the knowledge. She stares at me inquisitively and says, “you are not a squirrel! You cannot continue to gather nuts. At some point you must write!” She hugs me and walks away. It was as if I was struck by lightning. I know exactly where to start: first with my research questions, then an outline of the chapter. Then I will share the literature I have been gathering. I am not a squirrel.

This chapter explores the literature relevant to my journey. Several paths of literature inform my travels: ethical educational leadership, high-stakes assessment history and policy, the intersection of ethics and high-stakes assessment, emotional labor, advocacy and social justice leadership, authentic leadership, and performativity, and narrative as an access point for ethical leadership. Thus, in this chapter, I refer to literature discussing the differences between management and leadership, and then ethics and educational leadership. Subsequently, I outline the intersection of ethics, educational leadership and high-stakes assessment. Next, I review literature relevant to emotional labor, authentic leadership and performativity, advocacy and social justice leadership, the implicated advocate, resource allocation, and then finally I conclude with a chapter by identifying gaps in the literature and a summary.

**Management Versus Leadership**

Since the time of Bobbitt and Dewey there has been a debate in the United States about whether or not schools should be run like businesses or as democratic communities (Au, 2011). Within those competing ideas about what schooling should be there are arguments about whether or not the school administrator should be a manager or a leader (McGowan & Miller, 2001). According to McGowan and Miller (2001), “management tends to focus on maintaining existing
relationships and order, using proven ways of doing things, working within what people think is desirable and, of course, working harder and longer” (p. 32). However, they state, “leadership is about taking risks, striking out in new directions, creating visions, tapping imaginations, changing the way people think about what is desirable, creating excitement about working with children and communities, building new relationships and structures and changing the existing cultures” (McGowan & Miller, 2001, p. 32). Likewise, Gorton and Alston (2009) refer to a school manager as people who ensure procedures and rules are followed while the school leader is a “visionary” (p. 8). These competing ideas are often played out in real-life situations for K-12 school-based leaders as they must be part manager of the school and simultaneously be the leader of the school. Indeed, a school administrator may struggle to reconcile these competing roles when implementing high-stakes assessment policies.

Ethics and Other Closely Connected Terms

In addition to the research to follow, which presents the ethical dimensions of educational leadership, there are other closely connected terms, which interact with ethics or, in some cases, are considered the same or similar to ethics. These are values, morals, moral leadership and/ or literacy, integrity, and authentic leadership. Throughout the literature the terms values, morals and integrity are sometimes used interchangeably and sometimes defined as completely separate terms. It is necessary to note research conducted on the personal values of educational leaders influence all their decision-making (Ashbaugh & Kasten, 1984; Baig, 2010; Begley, 1999; Law, Walker, & Dimmock, 2003; Warwas, 2015). In some cases, researchers call for school leaders to have the moral fortitude to distribute leadership to effect school change (Fullan, 2003). One researcher even suggested some school leaders follow a selfish “ethic of success” citing Machiavelli’s *The Prince* (1999) as an exemplar for power-hungry administrators (Hodgkinson,
1991). Yet, Hodgkinson (1991) hypothesizes moral leadership and moral school organizations are interdependent. Nevertheless, it is in the scope of my journey to acknowledge ethics are messy and there is no one way to define it.

**Ethics and Educational Leadership**

Recognizing the ethical dimensions of compulsory education within American society, Dewey (1909) wrote:

> The moral responsibility of the school, and those who conduct it, is to society. The school is fundamentally an institution erected by society to do a specific work—to exercise a certain specific function in maintaining life and advancing the welfare of society. The educational system that does not recognize that this fact entails upon it an ethical responsibility is derelict and a defaulter. (p. 7).

Whether or not one agrees with the moral responsibility of schooling is for societies benefit or for the individual child’s benefit does not exclude the educational system from having an ethical responsibility. Indeed, as Dewey indicates, the ethical responsibility falls not only to the school system but also, to “those who conduct it” (p. 7). Educational leaders are part of those who conduct the school and ethical responsibility is intricately connected their day-to-day decision making.

Indeed, Campbell (1999) argues the plethora of literature on ethics for educational leaders proves ethics is vital to field, and she posits no one would suggest an educational leader should not act in ethical ways. Yet, Campbell (1999) also suggests the reason for such an abundance of academic work in the field of educational leadership focused on ethics is because there is a disconnect between the practical and theoretical application of ethics. Perhaps some of
the confusion lies in the fuzzy understanding of ethics. For a clear definition of ethics let us turn to Shapiro and Gross’s (2008) description of the concept:

[Ethics has had] numerous meanings over time. Initially, it came from the Greek word ethos, which meant customs or usages, especially belonging to one group as distinguished from another. Later, ethics came to mean disposition or character, customs, and approved ways of acting. (p. 20)

Strike (2007) suggests ethics are concerned both with, “what is good? and what is right?” (p. 9).

Educational leaders are tasked with doing what is best for students, teachers, and the greater school community, thus their role requires ethical decision-making.

According to Beckner, (2004), “educational leaders impact all parts of the education system…and most situations requiring action have moral and ethical undertones” (p. 3). Additionally, educational leaders often “rely more on experience and personal judgment than on theory, particularly theories from philosophy and ethics” (Beckner, 2004, p. 3). The frequency of ethical dilemmas encountered by school leaders is not as important as acknowledging every school leader will have an ethical dilemma they must reckon with at some point in their career. It is how the leader approaches the ethical tension that is important. Accordingly, the literature supports the notion the school leader’s approach to ethical dilemmas is important. In one study, Greer, Searby, and Thoma (2015) surveyed 115 students enrolled in educational leadership graduate studies using an existing measure of moral reasoning: the “defining issues test” (p. 511) and found their participants scored well below the national average. Greer et al. (2015) recommend educational leadership preparation programs include coursework focused on ethical decision-making. This study exemplifies the lack of experience some preservice educational
leaders have in dealing with ethical tensions but due to the limitations of the study it does not take into account the lived experiences of inservice school leaders.

Moreover, Cranston, Ehrich, and Kimber (2006) interviewed seven principals of private schools in Australia about the ethical tensions they face. They found all the principals struggled most with dealing with underperforming personnel and doing what is best for students. They suggested schools, school districts, or leadership preparation programs include professional development on ethics. In the same vein, Dempster & Berry (2003) present data collected from 552 questionnaires completed by Australian public school principals as well as data collected in 25 in-depth interviews. Their findings suggest principals are often in ethical quandaries and receive little professional development in the way of ethical theory. They propose principals create their own informal collaborative groups to learn about ethics. The call for training for educational leaders in ethical decision-making was a common recommendation throughout the literature.

Indeed, within the field of educational leadership it has been theorized the use of ethical paradigms could provide educational leaders guidance in decision-making (Capper, 1993; Greenfield, 1993; Shapiro & Stefkovich, 2001; Shapiro & Gross, 2008; Shapiro, Stefkovich, & Gutierrez, 2014; Starratt, 1991, 1996; Wood & Hilton, 2012). Initially, Starratt (1991, 1996) recommended educational leaders understand three ethical paradigms: ethic of care, ethic of justice, and the ethic of critique. Starrat’s (1991, 1996) calls for the use of ethical frameworks in the practice of educational leadership; however, his stance excludes the ethic of the profession. Starratt (1991, 1996) refers to the paradigms as themes and suggests educational leaders apply the ethic of care, ethic of justice, and ethic of critique in their decision-making process, thus using a multi-paradigmatic approach. He argues these three ethics compliment one another and
educational leaders make more sound decisions when they frame ethical dilemmas using the ethics of care, critique, and justice in concert.

Using Starratt’s (1991, 1996) three ethical paradigms, Shapiro and Stefkovich (2001) recommend educational leaders understand four ethical paradigms: ethic of care, ethic of justice, ethic of the profession, and the ethic of critique. Shapiro and Stefkovich (2001) frame the four paradigms around the “complexity and diversity” of schooling in today’s society (p. 3). They highlight the necessity of approaching ethics from multiple paradigms. Although Shapiro and Stefkovich (2001) recognize individual educational leaders will lean toward one paradigm over the other three, they encourage educational leaders to approach problems using at least two of the paradigms simultaneously. The point of using a multi-paradigmatic approach is for the educational leader to approach complex issues through more than just their individual perspective and embrace the problem with open minds. The idea of using multiple approaches in decision-making is not an isolated recommendation.

Later, Furman (2004) adds of the ethic of the community to Shapiro and Stefkovich’s (2001) four ethical paradigms by grounding the ethics of critique, justice, care, and the profession to the ethic of the community. Wood and Hilton (2012) also add a fifth, the ethic of the local community, as perspectives community college leaders should use as a model for ethical decision-making. Much like Shapiro and Stefkovich (2001), Wood and Hilton (2012), argue for educational leaders’ use of a multi-paradigmatic approach to decision making.

Following the five ethical paradigms introduced by Starrat (1991, 1996), Shapiro and Stefkovich (2001), and Furman (2004), Eyal, Berkovich, and Schwartz (2011) developed the Ethics Perspective Instrument to determine if educational leaders utilize more than one ethical paradigm and which ethical paradigm they utilized most often. Eyal et al. (2011) then asked 52
aspiring Israeli principals to complete the survey. They found more of their participants identified with an ethic of critique then anticipated. Further, they found within the ethic of justice there was a negative correlation between fairness and utilitarianism, a negative correlation between the ethic of care and fairness, and a negative correlation between an ethic of critique and an ethic of community. Finally, they found most of the participants’ ethical preferences were for the ethic of the profession and the ethic of care. Eyal et al. (2011) recommend replicating their work on a larger scale. The use of the ethical paradigms included in their study is a tool for approaching problems in educational administration from different perspectives.

Embracing Starratt’s (1991, 1996) ethical frameworks, Ehrich, Harris, Klenowski, Smeed, and Spina (2015) researched the ways in which six principals balance the tension of high-stakes accountability with other competing aspects of school leadership. They found that all the principals worked with Starratt’s (1991, 1996) frameworks of the ethic of care, the ethic of critique, or the ethic of justice to make sense of the competing values. Further, the principals Ehrich et al. (2015) interviewed described responding to accountability within the framework of Mintrop’s (2012) conceptual responses to outside accountability: “resistance, alignment, and coherence” (p. 201). The authors suggested school leaders could be more direct with their staff about how they use ethical decision making in their day-to-day practice (Ehrich et al., 2015).

Another school of thought concerning ethical educational leadership suggests veteran and aspiring educational leaders benefit from practicing ethical reasoning using case studies (Starratt, 2004; Strike, 2007; Strike, Haller, and Soltis, 2005). Strike et al. (2005) posit educational leaders must struggle with real-life examples of ethical dilemmas ripe for discussion and have no apparent right answer. They believe ethical growth and preparedness arise from grappling with ethical tensions. Similarly, Starratt (2004) walks readers through ethical dilemmas using case
study examples, and Campbell (1999) also recommends case studies as a means for wrestling with ethical issues. However, Campbell (1999) includes two concerns in her recommendations: first, case studies can be too simplistic, thus it is crucial to explore complex cases, and second, case studies can result in moral relativism, thus it is important to ensure case studies are not studied in a vacuum.

There are researchers within the field of ethical educational leadership, which have focused their research on specific ethical paradigms rather than ethics in a broad sense. Although there is not a plethora of research concerned with educational leadership and every ethical paradigm, I have included a brief description of five ethical paradigms. What follows are sections devoted to each of the ethical paradigms as well as social justice: the ethic of justice, the ethic of the profession, the ethic of care, the ethic of critique, including social justice, and the ethic of the community.

**Ethic of Justice**

The ethic of justice is mainly concerned with law, individual rights, good versus bad, rules, fairness, and consequences (Shapiro & Stefkovich, 2001). According to Starrat (2014), “the ethic of justice provides guidance to many school policies and procedures, insisting on some uniform attention to equity and fairness in the way school governs itself in all its daily activities” (p. 54). Beckner (2004) refers to five primary types of justice: procedural, substantive, retributive, remedial, and distributive (p. 37). Procedural justice focuses on the “application of rules” or laws; substantive justice refers to the correctness or fairness of the rules/ laws; retributive justice examines how one who breaks the rules/ laws can make retribution; remedial justice focuses on how the rule or law breaker can be remediated; and distributive justice refers to the equitable distribution of “benefits and burdens” amongst students (Beckner, 2004, p. 37).
Therefore, one should consider the ethic of justice paradigm as the application of rules/laws to ethical tensions.

**Ethic of the Profession**

Similar to the ethic of justice, the ethic of the profession is focused on “predetermined” criteria (Wood & Hilton, 2012, p. 205). Codes of ethics as guidelines for professions are “markers for ethical conduct” (Wood & Hilton, 2012, p. 205). Shapiro and Stefkovich (2001) suggest the ethic of the profession takes “moral aspects unique to the profession” into consideration (p. 18). Shapiro and Stefkovich (2001) further envisioned the ethic of the profession to include: “ethical principles and codes of ethics embodied in the justice paradigm, but it is much broader taking into account other paradigms, as well as professional judgment and professional decision making” (p. 21). The all-encompassing nature of this ethic makes it essential to orient it to individual professions or fields.

Within the field of educational leadership, the National Policy Board for Education Administration’s (NPBEA) Professional Standards for Educational Leaders (2015) outlines ethics for educational leaders, as shown in Figure 1. Thus, if one examines each of the actions NPBEA suggests the educational leader follow, the ethic of the profession has a responsibility to take into account other ethical paradigms. For example, d) asks that the educational leader promote social justice, equity, diversity, and democracy, which all easily fall into the ethic of critique.

In one study, Frick and Faircloth (2007) focused on the ethic of the profession. In particular, they sought to understand if principals viewed the needs of the individual differently than they viewed the needs of the collective or the best interest(s) of the student or students. Frick and Faircloth (2007) interviewed 11 secondary school principals and determined the
principals considered most of their everyday work focused on the best interests of the collective
group of students, but they viewed the best interests of the individual student on a case-by-case
basis. Thus, the perspective was completely different. The principals in the study were quick to
note they did not put the needs of the individual over the needs of the collective but approached
each situation differently depending on the circumstances. In a later study, Frick, Faircloth, and
Little (2012) revisit the ethic of the profession. Again they focused on principals’ reasoning
when working in the best interest of the collective and the best interest of the individual student.
This time they interviewed 13 elementary school principals and focused more intently on ethics
of the profession in working with students with disabilities (SWD). Frick et al. (2012) found the
principals interviewed worked to balance the needs of the collective with the needs of the
individual. One critique of the research is the overlap between their definition of the ethic of the
profession and the ethic of justice, in particular utilitarianism. However, the ethic of the
profession aims to be all-encompassing.

*Figure 1:* Professional Standards for Educational Leaders

<table>
<thead>
<tr>
<th>Standard 2. Ethics and Professional Norms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effective educational leaders act ethically and according to professional norms to promote <em>each</em> student’s academic success and well-being. Effective leaders:</td>
</tr>
<tr>
<td>a) Act ethically and professionally in personal conduct, relationships with others, decision-making, stewardship of the school’s resources, and all aspects of school leadership.</td>
</tr>
<tr>
<td>b) Act according to and promote the professional norms of integrity, fairness, transparency, trust, collaboration, perseverance, learning, and continuous improvement.</td>
</tr>
<tr>
<td>c) Place children at the center of education and accept responsibility for each student’s academic success and well-being.</td>
</tr>
<tr>
<td>d) Safeguard and promote the values of democracy, individual freedom and responsibility, equity, social justice, community, and diversity.</td>
</tr>
<tr>
<td>e) Lead with interpersonal and communication skill, social-emotional insight, and understanding of all students’ and staff members’ backgrounds and cultures.</td>
</tr>
<tr>
<td>f) Provide moral direction for the school and promote ethical and professional behavior among faculty and staff.</td>
</tr>
</tbody>
</table>
One of the main obstacles in the practice of educational leadership from embracing the ethic of the profession is the current political focus on school accountability. Cranston (2013) argues for educational leaders to push back against the demands of accountability and instead focus on professional responsibility. His argument centers on the locus of control for educational leaders. In a climate of accountability, the locus of control is outside of the school and outside of the leader, but professional responsibility brings the locus of control back to the leaders and within their reach and power. The ethic of the profession is integral to professional responsibility.

On the other hand, Black (2008a) provides a poignant description of the principal in his case study justifying her reasons for maintaining a quiet and disciplined school as actions in the best interest of her students. The principal wanted her predominately Hispanic population of students to be disciplined so they could achieve academic success on high-stakes assessments. Black (2008a) suggests this principal’s desire to provide her students with opportunities for “skills and attitudes” (p. 13) that would get them ahead in life was intended to be in the best interest of her students. Black (2008a) characterizes this justification as a “professional ethic” guiding the principal’s decisions (p. 13). This leader believes her professional responsibility to her students is to prepare them for high-stakes assessments. Thus, the application of this particular ethical paradigm can be complex and in some cases paradoxical.

Ethic of Care

According to Noddings (2005), moral life grows out of “love and attachment” (p. 81). The ethic of care puts people before laws or rules (Wood & Hilton, 2012). Noddings (2005) states:
An ethic of care—a needs-and response-based ethic—challenges many premises of traditional ethics and moral education. There is a difference of focus…there is also a rejection of universalizability, the notion that anything that is morally justifiable is necessarily something that anyone else in a similar situation is obligated to do…although an ethic of care puts a great emphasis on consequence…it is not a form of utilitarianism…it is not properly labeled a ethic of virtue…it is an ethic of relation. (p. 21)

Much like the ethic of critique, the ethic of care often questions the patriarchal perspective of the ethic of justice (Shapiro & Gross, 2008). Gilligan (1982) describes the ethic of care as revolving around the idea each of us is interdependent on one another. An education leader working from an ethic of care perspective focuses on student and teacher development and building the school community (Wood & Hilton, 2012). Noddings (1984, 2003) argues the ethic of care is our first and initial moral response as it is learned from being cared for as children. Essentially, the ethic of care is a relational ethic bound to qualities of nurturing, compassion, and loving others (Bergum & Dossetor, 2005). The ethic of care addresses how an issue impacts the community and/ or causes harm. In spite of this, one critique of the ethic of care is school administration, and the field of education as a whole, does not lack compassion or care, thus there is less need for the ethical theory (Willower, 1999). Nevertheless, this ethic calls for the educational leader to consider what is best for students and teachers and how the leader can help meet the needs of students and teachers.

Within the ethic is an approach specific to the African-American community. Bass’s (2012) research focuses on Black feminist care in the practice of five African-American female school leaders. She found that all five leaders actively demonstrated Black feminist care in their
work through mothering and prioritizing care over justice in some cases putting themselves at risk. Bass (2012) concluded Black feminist care was a strength of these educational leaders. In a critical approach to care, Blackmore (2010) argues for an ethic of care with a critical lens. She suggests educational leaders, particularly White female leaders, must examine themselves for “troupes of benevolence,” which addresses the need to make a difference and to care for the “other” but does not center true systemic change allowing for students, teachers, and leaders of color to learn or work in an inclusive system of schooling (p. 51). This feminist approach to caring is also focused on critiquing the social injustices educational leaders face in their daily work in schools.

**Ethic of Critique**

The ethic of critique deals with the “inconsistencies” of the ethic of justice, in which humans “formulate hard questions, and debate and challenge the issues” (Shapiro & Stefkovich, 2001, p. 13). The ethic of critique asks the educational leader to question justice, power, privilege, race, class, gender, and language norms and privilege (Hodgson, 1994; Shapiro & Stefkovich, 2001; Wood & Hilton, 2012). The ethic of critique considers questions such as “who makes the laws, rules, and policies? Who benefits from them? Who has the power? Who are silenced?” (Shapiro & Gross, 2008, p. 25). This ethic is based on critical theory and linked to critical pedagogy (Shapiro, Stefkovich, & Gutierrez, 2014). However, it is important to note to critique injustices or unjust values the educational leader needs to be able to identify them (externally and internally) in the first place (Ryan, 1999, p. 91). Further, Bates (2006) has noted the importance of educational leaders having an ethical understanding of critique by way of social justice even as he laments the lack of research on educational leadership for social justice. Indeed, the ethic of critique is centrally concerned with exploring power dynamics in education.
both in broad and narrow senses. Foster (1989) argues for leadership, which is critical, ethical, transformative, and educative (p. 50). Leadership critical in theory must also be critical in practice (Foster, 1989). Social justice is an important perspective within the ethic of critique paradigm, and social justice leadership is central to this study. Thus, I will explore social justice leadership more in-depth later in this chapter.

Accordingly, Cranston and Kunsanovich’s (2014) research combined performance methodology and reflexive journaling to analyze how 28 school leaders responded to a drama-based exploration of critical ethical dilemmas. Prior to and after their performances using a “fixed-script,” the participants completed a survey concerning ethics (Cranston & Kunsanovich, 2014, p. 52). In addition to the surveys, the participants also maintained journals to record their responses to acting out the ethical dilemmas. Cranston and Kunsanovich (2014) found the participants were more empathetic; more understanding of other’s perspectives and ethnodrama was an effective way of practicing ethical decision-making. Nevertheless, leadership from a critical perspective is not without its perils. For example, Daly, Der-Martirosian, Ong-Dean, Park, and Wishard-Guerra (2011) compared the results of surveys and interviews of transformational principals and transactional principals. Both sets of principals worked at schools that did not perform well within state accountability systems in a high-stakes assessment culture. Daly et al. (2011) found that transformational principals more willing to stand up and speak up to district leadership felt the threat of “sanctions” (p. 194) more acutely than transactional principals. Thus, it can be risky to practice an ethic of critique when it comes to resisting high-stakes assessment policies.
Ethic of the Community

The ethic of the community reflects a communal approach to educational leadership ethics. Furman’s (2004) addition of the ethic of the community to Shapiro and Stefkovich’s (2001) four ethical paradigms connects to the ethics of critique, justice, care, and the profession. Although Furman (2004) proposes the ethic of community, previous scholars have called for communal morals and values within schools (Beck, 1994; Sergiovanni, 1992; Starratt, 1994). Taking the work of Furman (2004) a step further, Wood and Hilton’s (2012) ethic of the local community, initially intended for community colleges, could be useful in the K-12 decision-making process. The ethic of the local community “is the notion that community colleges must serve the needs, interests, and public good of the local community” (Woods & Hilton, 2012, p. 206). In the much the same way, K-12 public schools must also serve the needs, interest, and public good of the community. According to Woods and Hilton (2012), “an ethic of the local community situates the best interests of the local community as a cardinal principle in decision making” (p. 206). A concern for school-based leaders of K-12 public schools should be to serve the needs of the community in which the school resides. Furman (2004) suggests, the ethic of community is a paradigm with which social justice, equity, and student learning can be most easily be put into practice by school administrators. Thus, it is imperative the school leader is aware of the needs and interests of the local community in the application of this ethical paradigm.

Each of the five ethical paradigms outlined are approaches to the practice of school administration. Arguably one of the most time consuming and stressful practices of school and district administrators is implementing high-stakes assessment policy. The literature surrounding high-stakes assessment implementation follows.
High-Stakes Assessment Policy Implementation

The Culture of Measurement

Performances on high-stakes tests have significant impacts on the schooling of students. A great deal is at a stake for both educators and students to perform well on these assessments, thus the term high-stakes assessment. Padilla (2005) describes the high-stakes assessment crusade as a “culture of measurement” (p. 249) focused on student test scores, the ranking of students and schools, and the narrowing of curriculum, which focuses on basic reading, writing, math, and test-taking skills. According to Padilla (2005), the culture of measurement only views students in the abstract, thus the only way to determine if the abstract student is learning is through “presumed norm” referenced tests (p. 250). By extension, the district and school administrators are also abstract beings at the implementation level responsible for increasingly improved student test scores and living in fear of the consequences when the students’ scores do not increase.

The fears of consequences are realized in McGhee and Nelson’s (2008) accounts of four principals whose careers ended because of poor student high-stakes assessment scores. McGhee and Nelson (2008) highlight the experiences of five principals with track records of successfully running schools were suddenly removed from their positions due to falling test scores. In all five cases, once the principals were let go they were not granted access to their superintendents and were shunned by colleagues. McGhee and Nelson (2008) describe this scapegoating of these principals the result of accountability policies, which measures “a person’s worth” (p. 371) on test scores. Taking an ethical stance against high-stakes assessments can be detrimental to a school leader’s career while assimilating politically, or at least pretending to assimilate, to the testing culture can be advantageous.
The Narrowing of Curriculum

McNeil (2000) suggests high-stakes tests may have been intended to close the gap between poor and/or minority and wealthy and/or White students, but, in actuality, the tests have proven to widen the gap even further. The narrowing of curriculum and the incessant skilling and drilling of practice tests has created, “even greater inequalities than had existed before” (McNeil, 2005, p. 91). Additionally, school leaders and teachers have chosen to focus on the students close to passing the high-stakes test to the detriment of students above and below the “bubble” (Booher-Jennings, 2005; Lipman, 2004, p. 43). The pressure to perform on high-stakes tests opens the doors for unethical behavior, as serious as cheating, for all involved (Cannell, 1989; Khan, 2014; Stobart, 2008). The narrowing of curriculum works to excludes students from a well-rounded education.

Although high-stakes assessment policies have the most impact on students, educational leaders also receive rewards and consequences for their schools’ overall performance on high-stakes tests. Hawkes (2011) provides a dramatic personal account of being a principal at school struggling to meet high-stakes assessment goals. He describes in detail the effects high-stakes testing has had on his role as principal in small ways, like being referred to as the CEO of his school rather than a principal, and in substantial ways, like being forced to narrow the curriculum. Hawkes (2011) explains how it feels to send a student on the verge of suicide back to class to finish the assessment already underway and therefore must finish before the days end to count. Hawkes (2011) argues against the complex, controlling and often contradictory policies of fear that come with leading a school in peril of being closed down. Undoubtedly, students also feel pressure to perform well on high-stakes assessments.
Testing Season

Surprisingly, testing season itself has little in the way literature addressing what it is and how it impacts schools, educators, and students. Instead, testing season is merely referred to in a few articles (Crowder & Konles, 2015; Demoss, 2002; Isis, Bush, Siegel, & Ventura, 2010; Maranto, 2016; Neill, 2016; Wright, 2009). Crowder & Konles (2015) call testing season the end of a course or the end of a grade high-stakes assessment timeframe. Isis, Bush, Siegel, and Ventura (2010) discuss mitigating the adverse effects of testing season on students by using art therapy. Neill (2016) states testing season occurs between March and May, and Demoss (2002) explains during this period of time schools focus on test preparation and test-taking skills. Maranto (2016) and Wright (2009) refer to testing season as a timeframe, the spring, when most high-stakes assessments are administered in public schools. Testing season is a time frame, usually during spring when the predominant amounts of high-stakes assessments are administered to public school students.

Student-Level performance

Although not implicitly related to educational leadership, the impact high-stakes assessment has on students is the single most important factor directly related to how school leadership enacts high-stakes accountability policies. Indeed, one scholar even referred to high-stakes assessments as a form of child abuse (Janesick, 2010). The culture of measurement and the narrowing of curriculum are aspects of high-stakes assessment implementation directly felt by children.

The culture of measurement. Black (2008a) describes a principal and an assistant principal intent on their largely Hispanic student population maintaining high levels of performance on their high-stakes assessment. Indeed, Black (2008a) explains the school is
considered an exemplar of Hispanic student success based on high-stakes assessment data and in spite of the high population of English Language Learners (ELLs). The students are micro-managed and attend a highly structured and disciplined school where halls are quiet, and students often take practice high-stakes assessments. Black (2008a) suggests the students “showed a sense of performance agency” (p. 9) in that the students accept responsibility to perform well on the high-stakes assessment. Students were placed in competition with one another as exemplified by bulletin boards showing each students’ performance scores on practice high-stakes assessment written on turtles and placed on the board from high to low depending on the students’ score.

In another research article, Black (2008b) highlights a high-stakes testing-themed pep rally held in a predominately Hispanic elementary school in which the whole school community placed the pressure of “collective performance” (p. 47) on the assessments squarely on the shoulders of fourth graders. At the same time, the pep rally served to remind the fourth grade students of the collective performance the current fifth graders attained the previous year, thus further creating a competitive atmosphere of maximized performance. Indeed, Black (2008b) describes witnessing the principal imploring students be successful on the high-stakes assessment for both their individual good and the good of the school. Black (2008b) notes the high-stakes assessment the students were preparing for, “labeled a student and shaped his/ her educational experience” for the foreseeable future.

At the elementary school level, Dutro and Selland (2012) analyzed the language students at one urban high poverty school used to describe high-stakes reading assessment and related adult practices. The student responses indicated the personal and institutional responsibility they felt to perform well on their tests. Students also recognized the “the man behind the curtain”
providing the rewards and consequences for student performance outcomes by referring to these unknown adults as “they” or “them” during their interviews (p. 351). Dutro and Selland (2002) note students who were “proficient” (p. 359) on their high-stakes reading assessment were more likely to demonstrate positive feelings about the test than the students who did not score as well. Likewise, Watson, Johanson, and Dankiw (2014) collected information from elementary school students about their feelings concerning high-stakes assessment. In their study, Watson et al. (2014) provided journals for students to voluntarily write their feelings about testing each day of a multi-day multi-subject high-stakes state assessment. After analyzing the student journals, the researchers found high incidents of students recording feelings of stress and nervousness sometimes resulting in physical illness. Considering these studies in contrast to the McNeil, Coppola, Radigan, and Helig’s (2008) study of students likely to drop out and the pattern of long-term negative feelings toward high-stakes assessments for students who do not perform well is demonstrated.

The narrowing of curriculum. In her research concerning the relationship between teachers pay-for-performance and student achievement on high versus low stakes assessments, Lauen (2013) found students at high-poverty schools did not fare as well on low stakes science assessments as they did on high-stakes assessments. Yet, students at high poverty schools performed better on high-stakes reading assessments then they did on low-stakes science assessments when comparing their performance with peers in low poverty schools. Lauen (2013) attributes these results to the narrowing of curriculum at high poverty schools to focus on subjects with high-stakes assessment at the expense of subjects like science. Lauen (2013) also found, in spite of teacher pay-for-performance incentives, schools where teachers “failed to receive bonus thresholds” (p. 110) demonstrated greater student achievement gains on high-

42
stake assessments than the schools where teachers met bonus thresholds. However, Lauen (2013) recommends further research to test these results. In their research, Maltese and Hochbein (2012) looked at the long-term effects of the narrowing of elementary school science curriculum on 12th grade students for three consecutive years. They found students who attended elementary schools demonstrating improvement on high-stakes assessments did not score as well on the ACT Science as students who attended schools with “declining” (p. 824) improvement on high-stakes assessments. Maltese and Hochbein (2012) posit this difference is due to hyper focus and amount of time the improving schools spent on teaching reading and math to the detriment of science. They further suggest the accountability measures put in place by NCLB to help prepare students for college majors in STEM areas have actually done the opposite. Students bear the weight of accountability to the detriment of a well-rounded education.

Further, McNeil at al. (2008) researched the impact of high-stakes assessment on dropout rates of high school students. Using a combination of existing student dropout data in a large urban school district, student survey data and ethnography, McNeil et al. (2008) found student dropout rates dramatically increased after the implementation of high-stakes accountability policies. Additionally, McNeil et al. (2008) report students attending high poverty high schools and African American and Hispanic students had much higher dropout rates than White or Asian students. The authors argue high-stakes accountability leads to greater dropout rates for several reasons. First, students who have been retained due to poor test scores are more likely to dropout of school. Second, students negatively impacting school assessment data are encouraged to dropout by school personnel, or worse, punitive measures are used to expel students who did not perform well on high-stakes assessments with the support of zero tolerance policies. Third, students surveyed reported the narrowing of the curriculum at high poverty schools took the joy
out of learning and providing reasons to attend school. Performance, or lack of, on high-stakes assessments, has long-term consequences for the most vulnerable students.

Many of the articles included in this section referenced the narrowing of curriculum in relationship to high-stakes assessment; however, Berliner (2011) regards the narrowing of the curriculum as the single most detrimental aspect of high-stakes assessment. Using existing data, Berliner (2011) examines the narrowing of curriculum in both the United States and Britain. He suggests the narrowing of curriculum removes the fun and enjoyment from school, ignores student talents, and “restricts thinking skills” (p. 287). Berliner (2011) admits this may be economically (i.e., testing companies) and politically beneficial in the short term but warns the long term economic consequences will be harsh as children will be unprepared for a 21st century world. Additionally, Amrein and Berliner (2002) found the amount of time devoted to skill and drill practice testing as well as the time spent taking the tests further narrowed the curriculum as there was less time to teach subjects like social studies or art.

The literature presented suggests students, especially low-performing students, carry the weight of their whole school’s success or failure. Students as young as eight are set up for success or failure and whether or not they are placed in academically challenging curriculum through the end of high school and beyond based on their performance on one reading and one math test, as Berlak (2000) suggests, “curriculum linked to so-called high-stakes testing is among the most powerful and pervasive forms of social control” (p. 195). Likewise, Hanson (2000) concurs, “of equal importance with the evaluation of past accomplishments is the use of testing to predict future behavior. In its future-oriented applications, testing is an instrument of social efficiency” (p. 69). The politics of high-stakes assessment and accountability rest on the shoulders of children. The educators, too, are socially controlled, and, in some cases, play into
the role created for them by policies of testing. Au (2009b) argues, “in addition to content and form controls, we see high-stakes testing operating as a pedagogic control over teachers’ practice” (p. 90) and, by extension, school leaders’ practice. Educators are certainly held responsible, and, in some cases, they struggle with how to do what is best for students while enacting high-stakes accountability policy. In the next section, I will present literature focused on educational leaders dealing with the ethical tensions of high-stakes assessment.

The loss of instructional time. Much like the narrowing of the curriculum, students who live in poverty or with disabilities disproportionately experience a greater loss of instructional time due to high-stakes assessments. Although not heavily studied, the loss of instructional time due to high-stakes assessments is noted in the literature (Invernizzi, Landrum, Howell, & Warley, 2005; Rodgers & Mirra, 2014; Smith, 2000, Zellmer, Frontier, & Pheifer, 2006). Berliner (2007) defined instructional time as, “academic learning time” (p. 3). As early as 2000, Smith argued that one of the common causes of lost instructional time were high-stakes assessments in her research on lost instructional time in Chicago elementary school students.

Further, Smith (2000) found even when 1/3 of the school year remained after testing season was over, instruction slowed or even stopped after high-stakes assessments were completed for the school year. The “post test period” was characterized by activities unrelated to instruction (Smith, 2000, p. 669). Rodger and Mirra (2014) surveyed teachers across California and found teachers in high-poverty schools reported twice as much instructional time was lost due to high-stakes assessments than in more affluent schools. Additionally, Invernizzi at al. (2005) found redundant high-stakes reading assessments led to the loss of instructional time in the elementary schools they researched. In their survey of administrators throughout the state of Wisconsin, Zellmer at al. (2006) found students with disabilities lost more instructional time due
to high-stakes assessments than students in general education. They explained students with disabilities lost as much as “15 days of disrupted instruction” because their teachers were administering high-stakes assessments (Zellmer et al., 2006, p. 44). Based on this research it is evident vulnerable populations of students experience a more significant loss of instructional time than less vulnerable students.

The Intersection of Ethics, Educational Leadership, and High-Stakes Assessment

Researchers have found when ethics and school leadership are forced to intersect with the implementation and administration of high-stakes assessment, school leaders find themselves having to balance managing and leading (Mintrop 2012). Applying the definitions provided by McGowan and Miller (2001) and Gorton and Alston (2009), the manager maintains the status quo, and the leader is a risk-taking visionary, school administrators struggle to do both simultaneously. This is described by Carpenter and Brewer (2014) as the, “implicated advocate,’ whereby the school administrator is tasked with advocating for what is best for his or her students while simultaneously implicated as an agent of the state through the implementation of high-stakes accountability policies. The literature demonstrates the destructive, neutral, and beneficial aspects of high-stakes assessments for school administrators.

Destructive Impacts

Mintrop (2012) found the managing the forces of accountability policies and keeping the best interest of students and personal integrity at the forefront of practice is a problematic balance for principals. Mintrop (2012) examined the overall integrity of nine schools based on their performances on high-stakes assessments, five high performing, and four low performing, in a mixed methods study. His findings indicated schools must, “strive…to develop or maintain their integrity in the face of incontrovertible accountability demands” (p. 716). Mintrop (2012)
suggests the implications of his research for school leadership is the need for principals to place student best interest before implementing accountability policies.

School leaders often have to adjust their leadership approaches in order to comply with high-stakes accountability policies. Shipps and White (2009) collected leadership practice data prior to and then immediately after the implementation of accountability policies. This research provides a unique perspective demonstrating the changes to leadership practices directly influenced by high-stakes assessment. In their first data collection, Shipps and White (2009) found principals “focused on the school’s internal instructional environment and school relationships defined by value commitments” (p. 370). In contrast to the findings to the first data collection, the second data collection yielded results suggesting principals focused on professional relationships outside of their schools and “they attended less to ethical precepts or experienced moral accountability” (p. 370). Shipps and White (2009) suggested these changes in leadership practices were often in the best interest of the school leader rather than the best interest of students or teachers. In some cases, leaders make do with the difficulties of high-stakes assessment policies, and in other cases, leaders make what appears to be selfish decisions.

Moving beyond the adjustment of leadership approaches, the notion of having to give something like personal integrity up to ensure student success on high-stakes assessment was bore out by Lipman (2004) in cases studies she conducted in Chicago public schools. Lipman (2004) describes teachers and school leaders slowly letting go of ideals about education and “assimilating” to the testing culture (p. 131). This exemplifies how educational leaders living within the ethical tensions may be worn down over time and stop resisting the realities of enacting high-stakes assessment policies. Yet, there are examples of school leaders who have gone too far to assimilate.
Using Callahan’s (1962) Vulnerability Thesis in relationship to educational leadership and accountability policies, Khan (2014) explains the problems facing educational leaders are a direct result of their job description. Khan (2014) uses the cheating scandal in Atlanta as an example of how leadership positions in education are by nature vulnerable due to public scrutiny and the value of reputation. Citing the adoption of business practices in schools as the source of the accountability movement, Khan (2014) describes the climate in Atlanta Public Schools as placing an, “over-emphasis on test results and public perception to the exclusion and detriment of integrity, ethics and individual student progress” (p. 5). Although he does not completely disagree with applying business practices to public schools, Khan (2014) suggests the “indiscriminate” adoption of business is practices are the root of the problem. Accountability, particularly high-stakes assessments, created the conditions for the ethics of educational leaders to be vulnerable.

Arguing for the need to move leadership away from its current focus on accountability policies and onto moral leadership responsibility, Cranston (2013) implores school leaders to lead the change in policy rather than remain “the ‘doers’ of the bidding of” policymakers (p. 131). He explains the focus on “managerial efficiency” needs to be replaced by a focus on students. Cranston (2013) defines leadership responsibility as leadership with a vision, leaders learning alongside teachers and students, and leading the national debate about a new school reform “grounded in notions of professional responsibility” (p. 129). By collectively reshaping educational leadership, the restraints high-stakes accountability policies have on public schools can be broken.

Furthermore, Vasquez Helig, Young, and Williams (2012) applied the business concept of risk-management to Texas schools working within high-stakes testing policy. After
interviewing 89 Texas school principals and teachers, they found three unintended consequences of the policy: concerns about possible job loss, finding loopholes in the policy implementation, and the viewing of at-risk students as “liabilities” (Vasquez Helig et al., 2012). Teachers and principals were likewise concerned about being at-will employees and felt they had to teach to the test because they could lose their jobs if they did not. They also reported looking for loopholes in the testing policy so they could enact the policy in a manner most advantageous to the school. The most striking finding, however, was the ethical quandaries associated with at-risk students as participants described being caught between doing what was best for the student or risk having the school closed down. Vasquez Helig et al. (2012) argued this behavior was in fact risk management and application resulted in schools being risk-adverse to at-risk students. Vasquez Helig et al. (2012) suggested participants admitted to encouraging at-risk students to dropout or discouraging at-risk students from enrolling in their schools, and in one instance the school administrator gathered all at-risk Latino students and told them he would have them deported.

Neutral Impacts

Spillane and Kenney (2012) explore how principals deal with the competing demands of accountability policies and classroom teacher agency. They suggest that these demands can be seen as “organizational legitimization” and “organizational integrity” (p. 547). By adhering to accountability policies principals legitimize their organizations to the public and by providing teachers agency they create organizational integrity within the school community. Coming from a completely different point of view, Sanzo, Shermand, and Clayton (2011) looked at the practices of principals with schools and students that performed well on high-stakes assessments. Sanzo et al. (2011) determined these practices included distributive leadership, a heavy focus on
professional development and behaviors demonstrating integrity, such as honesty and a willingness to listen. These principals managed the demands of high-stakes assessments while maintaining their integrity. Within the same vein, Feuerstein (2013) discussed the ethics displayed by superintendents enacting accountability policies in their school districts. Feuerstein (2013) suggests the nature of the job of the superintendent requires a transactional style of leadership in the face of accountability policies. Nevertheless, Feuerstein (2013) also suggests most superintendents act in the best interest of schools, teachers, and students. The complex and contradictory nature of acting in the best interests of students and implementing high-stakes assessment policies requires leaders to manage from the middle or to work within the tensions.

Working within the tensions relates not only to principals and superintendents, but also district-level administration and school board members. Black and Shircliffe (2013) examined how school and district-level leaders managed accountability policies while also working toward equitable practices focused on social justice. Black and Shircliffe (2013) analyzed interview data from school board members and district-level administration in a large school district. They found these district leaders were able to “make do” with the state and federal accountability policies while simultaneously finding ways to create more equitable district policies, like dedicating Title 1 dollars to pay for faculty, and also avoiding disciplinary consequences from the state (Black & Shircliffe, 2013, p. 109).

**Beneficial Impacts**

In other research focused on moral leadership, Ylimaki and McClain (2009) call for wisdom-centered leadership practices. According to Ylimaki and McClain (2009), “today’s principals must navigate the maze of the current high-stakes assessment and accountability from a deep understanding of wise leadership practice as evidenced through the attitude of equanimity,
or balance, compassion and joy in themselves and in the communities they lead” (p. 14). They examined the characteristics of principals who found ways to “navigate” accountability policies in “wise and humane ways” (p. 14). They suggest these characteristics are congruent to Buddhist wisdom. Drawing on tenets of Buddhism, the article provides a framework for wisdom-centered leadership to help principals deal with accountability policies. The notion of wisdom allows for ways to work within the tensions of high-stakes accountability and personal ethics.

Additionally, Hughes and Jones (2011) sought to determine the relationship between student achievement scores and principals who attended professional development trainings on ethics. After surveying 600 principals, they found the principals who reported having attended a pre-service or in-service professional development training also reported student achievement on high-stakes assessments were higher than the principals who did not attend an ethics training. Hughes and Jones (2011) imply this correlation between principals trained in ethics and better student performance is cause for all principals to receive ethics training. In spite of the correlation, it is difficult to determine what other factors may have also impacted this outcome; nonetheless, ethical training for educational leaders is vital.

Accountability policy, in particular, high-stakes testing, has shaped the relationships between educators and students in that every educator has a personal interest in students to perform well on high-stakes assessments instead of having a personal interest in students reaching their potentials. Thus, teachers need students to perform to please their principals and to help them receive pay-for-performance bonuses. School administrators need students to perform to stave off consequences from central offices and maintain professional stances of success. Districts want to avoid state-level sanctions, thus they need students to perform well.
In a strange way the relationship between educators and students is inverse to what children need and what education should be doing. Rather than the students depending on educators to help them, the educators now depend on the students. Admittedly, the ethical aspects of accountability policies which judge the quality of students, teachers, leaders, districts, and state systems of education based solely on test scores are questionable at best.

**Emotional Labor and Ethics**

**Potential Spaces of Negotiation for Educational Leaders and Emotional Labor**

Hochschild (1983) introduced the notion of emotional labor into the literature in her book *The Managed Heart*. According to Williams (2013), Hochschild described emotional labor as centering on:

…Creating a desired emotion in another and requires individuals actively to work on their own emotions when they do not spontaneously feel what they ought to in a given situation. Feelings have to be induced or suppressed in order to present an appropriate outward appearance…(p. 5-6)

Hochschild (1983) recognized because emotions are pliable and often momentary they have been commoditized by the labor market (Williams, 2013). The commoditizing of employee emotions by employers is mistreatment and mismanagement, but most of all it is “exploitation” (Williams, 2013, p. 5). Indeed, the employee’s real emotions are concealed from the customer creating a situation in which the employee is distanced from their true feelings (Williams, 2013). Williams explains, “Hochschild argued essentially that if a worker can become estranged from the product of his labor in a factory, this could equally apply to emotions in a service industry” (p. 6).

Hochschild’s work has since inspired a plethora of research focused on emotional labor in various contexts and fields, including business management (Ashforth & Humphrey, 1993;
Brotheridge & Grandey, 2002; Grandey, 2000; Liu, Liu, & Zeng, 2011), medicine (Henderson, 2001; Larson & Yao 2005; Lopez, 2006; Mann & Cowburn, 2005; McQueen, 2004), education (Bellas, 1999; Blackmore, 1996; Constanti & Gibbs, 2004; Zembylas, 2004), and tourism (Kim, 2008; Van Dijk, Smith, & Cooper, 2011), and too many others to be adequately addressed here as according to Google Scholar *The Managed Heart* has been cited almost 17,000 times.

Morris and Feldman (1996) operationalized emotional labor in four parts: “(a) frequency of appropriate emotional display, (b) attentiveness to required display rules, (c) variety of emotions required to be displayed, and (d) emotional dissonance generated as the result of having to express organizationally desired emotions not genuinely felt” (p. 987). Thus, emotional labor operationalized in these terms can be viewed as a constant performance the employee must play a role in until the employee begins to doubt or question his or her true emotions.

Tracy’s (2000) article, “Becoming a Character for Commerce,” describes Tracy’s experiences and research she conducted during her 10-month stint as a cruise ship’s activities director. The research “analyzes emotional labor, self-subordination, and discursive construction of identity in the context of a total institution” (p. 91). Further, Tracy (2000) suggests discursive construction of identity combined with self-subordination and emotional labor lead to “burnout” (p. 91) of cruise ship employees. Tracy (2000) connects emotional labor with ideas of organizational “excellence” (p. 93) in which employees who interact with the public must always pretend to be happy and helpful. Negative thoughts and private feelings must be suppressed at all times. Thus, those engaged in emotional labor are forced to fake happiness when they are the face of the organization and the “real self” (p. 96) is repressed. Therefore, the organization exerts power over the individual’s identity.
**Emotional labor and educational leadership.** Although the literature examining the emotional labor of educational leaders is limited, there are a handful of studies looking more broadly at emotions and the work of educational leaders and a few that specifically look at the emotional labor of educational leaders. For example, Blackmore’s (2011) research uses a feminist lens to study the emotional aspects of female school leaders. Blackmore states within “educational leadership literature, there is evidence of how women have often felt their emotions pathologized as private and women’s emotionality being used as a rationale for their exclusion from decision making” (p. 218). If emotions have been pathologized for women then for men demonstrating emotions is entirely unacceptable behavior in educational leadership.

**Authentic Leadership Versus Performativity Versus Emotional Labor**

In recent years there has been increasing interest in the concept of authentic leadership (Begley, 2001; Duignan & Bhindi, 1997, 2009; Duignan, 2014; Starratt, 2004). According to Duignan and Bhindi (1997), authentic leadership should be thought of in terms of: “leadership praxis linking theory, practice, and ethics of leadership” (p. 208). Further, Duignan & Bhindi (1997) describe an authentic leader as being his or herself in all matters but, at the same time, putting the needs of others first. Additionally, the authentic leader considers what is best for both the individuals and groups following him or her and what is ethical in all of their decisions and actions (Duignan & Bhindi, 1997). Duignan & Bhindi (1997) describe leadership as the opposite of authentic: “artifice” style leadership. They suggest fake or phony leaders frequently pretend to be authentic but in actuality are putting on a performance. The educational leader may need to be inauthentic at times when implementing policies they might not agree with, for example, Hodgkinson (1991) describes the school leader has an “agent of acts” rather than the “author of acts” (p. 123). Thus, the challenge for educational leaders is to be simultaneously
another’s agent yet also be authentic as there will always be outside forces attempting to control schools.

Being an inauthentic leader requires a certain level of performativity. Ball (2003b) suggests systems of accountability in public schools force educators to, “set aside personal beliefs and commitments and live in an existence of calculation” (p. 215). Likewise, being an agent of acts for accountability policy implementation may require educational leaders to put aside their values, morals, and ethics to live an existence of inauthenticity. Indeed, Ball (2003b) refers to this as a “struggle for the soul” of the educator. Educational leaders struggle about the sole purpose of accountability versus the soul purpose of education and their role in those contradictions.

As Smith (2004) suggests, “in the political spectacle, persons become actors, take on roles, and polish their image” (p. 18). Policies of accountability attempt to create transparency; however, this can result in attempts at appearing accountable or doing what is expected when in reality it is merely pretense.

However, Hanley (2015) suggests neoliberal policies within education and the pressure educational leaders feel when these reform policies are put into practice “create a demand for emotional performance” (p. 105). Hanley (2015) argues neoliberal policies such as high-stakes assessment force educational leaders to perform in a way they might not typically perform emotionally. He posits, “the emotionally skilled educational leader presents emotional responses in ‘positive’ language. This can create difficulty for educational leaders whose emotional responses may be ambivalent and troubling” (Hanley, 2015, p. 110). Thus, for educational leaders the need to perform emotionally is equal to the need for their schools to perform in the age of accountability measures in public education.
In his ethnographic study, Iszatt-White (2009) suggests there are performance norms embedded into the job of a school leader, norms governing how a leader ought to act in given situations. Moreover, Iszatt-White (2009) found the “emotional laboring of educational leaders suggests a reversed causality in which the beliefs may be said to prompt or cause the acting, rather than deriving from them, without doing away with the requirement for acts of labor in specific instances” (p. 462). Thus, the beliefs steeped in neoliberal policies may cause educational leaders to perform contrary to personal beliefs.

This then suggests a more profound tension for educational leaders attempting authentic leadership when policies of high-stakes testing require emotional labor. The contradictions between being an authentic leader but needing to perform the role of the school leader, which includes setting aside or hiding feelings, acting in a manner expected by colleagues, and enacting policies of high-stakes assessments, are staggering.

**Social Justice and Advocacy Leadership**

**Social Justice Leadership**

Social justice leadership is defined by Theorharis’s (2007) as school administrators focusing on, “issues of race, class, gender, disability, sexual orientation, and other historically and currently marginalizing conditions in the United States central to their advocacy, leadership practice and vision” (p. 223). Within the field of educational leadership there is a plethora of research with several nuanced approaches to social justice leadership (Blackmore, 2009; Bogotch, 2000; Dantley & Tillman, 2006; Diem & Carpenter, 2012; Frattura & Capper, 2007; Jean-Marie, Normore, & Brooks, 2009; Marshall & Oliva, 2006; McKenzie et al., 2008; Rusch & Horsford, 2008; Theoharis, 2007, 2009, 2010). Jean-Marie et al. (2009) suggest school administrators have a responsibility to students and their communities to advocate for
educational and social changes, which extend educational and social openings to “marginalized and oppressed” students (p. 4). Blackmore (2009) argues Jean-Marie et al.’s (2009) call is too generalized and notes there are many different ways humans are oppressed and marginalized; instead, both issues of inequity and difference should both be considered instead of just equity. She states, “social justice is central to leadership preparation because without addressing issues of difference and inequality then there will be no substantial improvement of student learning for those in challenging circumstances” (Blackmore, 2009, p. 8, emphasis in original text). Bogotch (2000) posits social justice educational leadership means different things to different people, what it looks like in theory may not be mirrored in practice or vice versa, but individual educational leaders working toward socially just systems of education is what is important. Perhaps, most important to this study of educational leadership is “deliberate intervention that requires the moral use of power” (Bogotch, 2000, p. 2). Dantley and Tillman (2006) further suggest social justice educational leadership is moral transformative leadership, which is focused on the, “use and abuse of power,” and the desire to create educational and social equity for all students in educational institutions (p. 24). To truly be an ethical leader, the school administrator must apply power with social justice at its center.

**Advocacy Leadership**

School administrators should be the advocates for students with disabilities, students who are English Language Learners, students living in poverty, students of color, and students who are LGBTQ. The inclusion of all students in all curriculums is the true definition of equity. Thus, social justice leadership is similar to advocacy leadership in that the leader advocates for social justice within schools, but advocacy leadership includes a “more politicized notion of leadership” which “acknowledges schools are sites of material and cultural resources and
ideological commitments” (Anderson, 2009, p. 13). According to Anderson (2009), advocacy leaders recognize issues of social justice at school, district, community, and political levels simultaneously and work to eliminate barriers. Advocacy leaders also realize some systemic problems may be beyond their locus of control but work to ensure a democratic education to enable vulnerable students to take part in the politics of the future (Anderson, 2009). It is within the push and pull of systemic social justice problems the implicated advocate resides.

The implicated advocate. As discussed earlier in this chapter, the implicated advocate is the school leader who works to advocate for social justice but at the same time is implicated as the implementer of neo-liberal systems of accountability (Carpenter & Brewer, 2014). Carpenter & Brewer (2014) do not define advocate the same way Anderson (2009) does, instead their advocate is less political and more in line with social justice leadership. For example, Thompson and Mockler (2016) studied how principals in Australia managed to become the “heroes” of school reforms and implementers of high-stakes accountability (p. 2). They found the principals in their study voiced this contradiction by working to find the positive aspects of high-stakes assessments in the negative aspects by working to find the positive aspects of high-stakes assessments in the negative aspects (Thompson & Mockler, 2016). Werts, Della Sala, Lindle, Horace, Brewer, and Knoeppel (2013) connect Carpenter and Brewer’s (2014) research to their findings in their study about educational stakeholder’s sense-making of South Carolina’s accountability policies and suggest, “educational leaders have to not only shape the messages received by education stakeholders, but also to encourage a more robust political discourse about education policy” (p. 416), thus bringing more advocacy to their roles.

In a study of assistant principal social justice identity in schools that are low-achieving on state grading systems, Carpenter, Bukoski, Berry, and Mitchell, (2017) suggest assistant
principals use their power for advocacy in spite of their implicated roles in federal and state systems of accountability. Lac and Cumings Mansfield (2018) argue as implicated advocates, school administrators have a responsibility to help their students become advocates for themselves. Frankly, using Carpenter and Brewer’s (2014) framework, many public school administrators can be defined as an implicated advocate because all school administrators should be advocating for their students and all are implicated in larger systems of accountability.

**Resource Allocation**

Resource allocation is part of project management. When managing a project, the project manager must determine the best ways to distribute resources. In K-12 public schools, school administrators determine where human and capital resources, such as teachers or budgets, should be used. The discussion of resources and the planning for the use of resources in public K-12 schools is limited in the literature, save human resources and school finance. Research on human resource management in public K-12 schools is focused primarily on the hiring and retention of teachers (Briggs & Desmond, 2014; Bryk & Schneider, 2003; Cunningham & Cordeiro, 2003; Seyfarth, 2002, 2008). Cunningham and Cordeiro (2003) devote a chapter to resource allocation in their book about school leadership, but the focus is on school finance rather than the allocation or reallocation of other capital resources. Another example, Halverson and Plecki (2015) studied how a school district made changes to how they allocate financial resources to be more equitable. Yet, the literature concerning resource allocation in public schools in relationship to high-stakes assessments is concerned with the diversion of school instructional resources toward students close to passing high-stakes assessments (Booher-Jennings, 2005; Neal & Schanzenbach, 2010; Reback, 2008; Reback, Rockoff, & Schwartz, 2011). There is no research focused on the reallocation of resources such as classrooms,
computers, access to resource teachers, and the reallocation of time due to high-stakes assessments in the literature except for Crowder and Konle’s (2015) mention of it in passing in their article concerning the testing company Pearson.

Gaps and Summary

The literature reviewed in this chapter demonstrates the lack of research focused on the ethical dimensions of implementing high-stakes assessment for K-12 educational administrators. Further, there is almost nothing in the research reviewed in regards to principals’ perceptions of their ethical obligations to students when it comes to the implementation of high-stakes assessment or principals’ feelings about the amount of time students spend taking assessments throughout the school year, in particular, high school students who have not historically been successful on high-stakes assessments. Further, there is a lack of research concerning the allocation or reallocation of resources due to high-stakes assessments.

Additionally, recent literature demonstrates a need for voices in educational leadership to share their narratives surrounding personal belief systems and high stakes-assessment. For example, McGhee and Nelson (2005), both former school principals, expressed living with the dilemmas associated with high-stakes assessments and meeting the needs of their students. Indeed, Vernaza (2012) calls for educators to have a voice concerning high-stakes assessment policies. Larson (2014) argues for “in-depth ethnographic and case study research” to begin to understand the nature of unjust institutional systems of achievement and argues for a “capability” approach to transform education (p. 163). Additionally, Muzaliwa and Gardiner (2014) posit for narrative inquiry as a method that can “advance” educational leadership (p. 184).

Another area I found the literature lacking was in respect to educational leaders finding small ways to resist accountability policies. Although Theoharris (2010) presents narratives of
principals working to resist social injustice in their respective schools, and one of the articles reviewed did robustly address leadership resistance to high-stakes accountability policies (Black & Shircliffe, 2013), more research in this arena has the potential to help educational leaders in their day-to-day practice. Research delving deeper into how educational leaders find loopholes in the high-stakes assessments policies and use them to the benefit of schools, teachers, and ultimately students would help public schooling and perhaps other educational leaders. Furthermore, there is a gap in the literature surrounding the allocation of resources beyond human resources and budgeting for public school administrators. Further, no literature wholly treats the phenomena of testing season beyond demarcating it as a period of time when high-stakes assessment takes place in public schools.

Another gap in educational leadership literature is related to emotional labor. Apart from one study (Blackmore, 2011), little attention has been paid to emotional labor in the field of educational leadership. Additionally, no research within the field has examined the connections between performativity, emotional labor, and authentic leadership. Similarly, there is a gap in the literature analyzing gendered notions of high-stakes assessment policy. In particular, there is no literature exploring how these concepts further connect to ethics.

This chapter began examining the differences between management/administration and leadership and exploring ethics pertaining to educational leadership as well as the ethical paradigms: ethics of justice, community, care, profession, and critique. Next, the chapter moved into an examination of the implementation of high-stakes assessment policies and how it pertains to school and student performance. Then, the chapter reviewed the literature at the intersection of ethics, high-stakes assessment, and educational leadership. This was proceeded by a brief explanation of emotional labor, performance, authentic leadership followed by a discussion of
the notions of authentic leadership in relationship to both performativity and emotional labor.

Next, the chapter continues with a review of advocacy and social justice leadership, the implicated advocate, and resource allocation. Finally, I conclude the chapter by identifying gaps in the literature and a summary.

**Conclusion**

Throughout this chapter and on this stop of my journey I have reviewed literature related to ethics, educational leadership, high-stakes assessment, the culture of measurement, testing season, the narrowing of the curriculum, emotional labor, social justice leadership and advocacy leadership. I also noted the call for more narratives from researchers in the field of educational leadership. The next stop on my travels I will explain for my methods for gathering stories to help me move forward toward my sacred destination.
CHAPTER 3: THE TALE OF METHODS

“See I’m not walkin’ on it
Or tryin’ to run around it
This ain’t no acrobatics
You either follow or you lead, yeah” (Irvin III, Joseph II, Patton, & Monáe Robinson, 2010).

*******

“Start with your methods chapter,” Dr. Black tells me as he gingerly puts the stack of books I just returned to him back into his vast collection one at a time. I am standing in the doorway to his office late on a Saturday afternoon. Each book on the shelf represents aspects of the sacred destination I am seeking on my journey. “That’s a good idea,” I agree. I know I am planning to write a narrative including the narratives of other school leaders as well as my own as I continue my pilgrimage to better understanding ethical dilemmas in relationship to high-stakes assessment. For this stage of my journey, I focus on how I can best understand my participants through narrative as a method.

The purpose of this journey is to explore the stories of the ethical dilemmas K-12 school administrators navigate when implementing high-stakes assessment policy in a culture of measurement during testing season. This stop on my travels focuses on the methods I used for research. I begin with the epistemology undergirding the research methods, followed by the methodology, description of my participants, and then the methods of story collection, selection, and analysis I utilized. I then describe ethical considerations, and I conclude with a brief summary of the chapter.
Epistemology

The epistemological stance I embrace in this study is interpretivism. Bochner (2005) refers to interpretivism as “interpretive social science,” which “embraces the power of language to create and change the world, to make new and different activities and meanings possible” (p. 66). The interpretivist does not believe in separating the heart and the mind. Rather, he or she attempts to join the heart and mind as well as to join social science and literature. Bochner (2005) suggests interpretivists are instruments and participants within their own research as it is impossible for researchers to be value-free. Thus, the focus of inquiry for the interpretivist is on meaning instead of objects.

According to Schwandt (1994), the interpretivist, “believes that to understand this world of meaning one must interpret it” (p. 118). Therefore, each human interprets everything uniquely. Schwandt (1994) explains this perspective focuses on the creation, negotiation, modification, and the sustaining of meaning for each individual human. The act of meaning-making is the primary focus of the interpretivist. The interpretivist’s inquiry accepts the subjectivity of the inquirer. At the heart is the concept of Verstehen (emphasis mine) (Schwandt, 1994). Schwandt (1994) describes Verstehen as the process of “grasping or understanding the meaning of social phenomena” (p. 119). Crotty (1998) describes Verstehen as an “inquiry into human affairs” as the researcher is seeking to understand (p. 94). The interpretivist’s methods, “interacts with participants and the world by using methods that are closer to literature than to physics; interpretivists probe how meaning is performed and negotiated in the everyday world” (Paul, Gaffam, & Fowler, 2005, p. 47). The interpretivist uses a narrative structure to find meaning within the, “relationships between author/ researcher and subject” (Bochner, 2005, p. 66). Interpretive research methods in the social sciences explore the role of “stories and
storytelling” in the production of individual and cultural identity (Bochner, 2005, pp. 66-67). The stories often take the form of narrative ethnographies, autoethnographies, and case studies.

**Historical Roots**

The beginning of interpretivism as a theoretical perspective has deep roots stemming out of the study of literature and interpretation of biblical scripture. This study and interpretation have been termed *Hermeneutics* (emphasis mine; Crotty, 1998, p. 88). Although the term hermeneutics is less than three hundred years old, the practice of interpreting texts traces back to the ancient Greeks and Jews (Crotty, 1998). Crotty (1998) explains the modern study of hermeneutics began in the 1800s with German philosophers Fredriech Ast and Fredriech Schleiermacher. According to Crotty (1998), Schleiermacher’s philosophy of hermeneutics centered on both grammar and interpretation. A student of Schleiermacher, Wilhelm Dilthey’s study of hermeneutics advanced his philosophical stance by including both history and social phenomena as part of the interpretation of text (Crotty, 1998). Thus, “we are our history…the fact that language and history are both the condition and the limit of understanding is what makes the process of meaning constructing hermeneutical” (Schwandt, 1994, p. 120). It is important to note, these philosophers considered the process of hermeneutical Verstehen as objective rather than subjective.

Additionally, hermeneutics was also the focus of Martin Hiedegger’s work in the 1960s. However, unlike his predecessors, Hiedegger considered hermeneutical Verstehen as subjective. He believed hermeneutics is the phenomenology of ontology or how the experience of being is interpreted (Crotty, 1998). Also in the mid-twentieth century, Alfred Schutz’s work in the sociological tradition of phenomenology heavily influences interpretivism (Schwandt, 1994). Schutz explored the process of phenomenological Verstehen. This process “refers to the
intersubjective character of the world and the complex process by which we come to recognize our own actions and those of our fellow actors as meaningful” (Schwandt, 1994, p. 121). The late twentieth-century interpretivists draw on both phenomenology and hermeneutics.

The ontological assumptions of interpretivism are: “reality is mediated by language; because the world does not exist in the shape of our sentences, the mind plays an active role in the construction of reality” (Paul et al., 2005, p. 47). The epistemological assumptions of interpretivism are: “knowledge and the knower are inextricably linked; events are placed in intelligible frames by a mind that actively engages the world, attaching significance to those events” (Paul et al., 2005, p. 47). The goal of the interpretivist perspective to examine, “how people breach circumstances…how they make the absurd sensible and the disastrous manageable; and how they turned calamities into gifts or lessons for living” (Bochner, 2005, p. 67). Richardson (1990) writes, “we craft narratives; we write lives” (p. 10). The interpretivist not only acknowledges their writing is entwined with their research, the interpretivist embraces the enmeshment. “Ful wys is he that kan himselve knowe” (Chaucer, 1987). Chaucer’s words sum up the interpretivist perspective because interpretivists know they—the researcher—are as much a part of their research as those they are researching. Interpretivists know and acknowledge their own participation.

**Central Arguments of Interpretivism**

The first interpretivist persuasion is interpretive anthropology, which is a “theory of culture” (Schwandt, 1994, p. 123). Interpretive anthropologist Clifford Geertz (2001) describes this as an “explanation” focusing “attention on what institutions, actions, images, utterances, events, customs, all the usual objects of social-scientific interest, mean to those whose institutions, actions, and customs, and so on they are” (p. 167). Hence, both the researcher and
the participant interpret culture in context. The second persuasion is symbolic interactionism. According to Schwandt (1994), symbolic interactionism is hard to describe due to the plethora of theories about it. However, in its simplest terms, symbolic interactionism posits, “human beings act toward the physical environment” and other persons (Schwandt, 1994, p. 124). The meanings made from the interaction between human and the environment results in communication. Thus, communication is symbolic (Schwandt, 1994, p. 124).

The third persuasion is interpretive interactionism. Norman Denzin (1994) suggests “interpretivism is an art; it is not formulaic or mechanical. It can be learned, like any form of storytelling, only through doing…writing is interpretation or storytelling” (p. 502). This art is shaped by style, society, genre, archetype, creates deep interpretative understandings. In essence, the interpretivist believes objectivity is impossible because the inquirer must interpret what is seen essentially rendering the inquiry subjective. This form of interpretivism takes on poststructuralist leanings (Schwandt, 1994, p. 123). For example, Holstiend and Gubrium (1994) suggest one of the tenets of interpretivism is the centrality of language for communicating meaning. These three main “interpretivist persuasions abound” today (Schwandt, 1994, p. 122). For this dissertation, I will mainly be utilizing two of the persuasions, symbolic interactionism and interpretive interactionism. Indeed, narrative inquiry uses aspects of interpretive interactionism, storytelling, to explore symbolic interactionism, how humans interpret the word.

Firmly rooted in interpretivism, narrative inquiry assumes knowledge interconnects to the life events (Paul et al., 2005). According to Denzin (2014), the written and spoken word are not sufficient to detail specific human experience, rather language represents our experiences. It is through language the knower must create a representation of experience. The knowledge cannot exist without the knower, and meaning is presented through descriptions of events (Paul et al.,
Indeed, interpretivism privileges stories and the performance of stories based on human experiences. “Theory merges with the story when we invite others to think with the rather than about it” (Bochner & Riggs, 2014, p. 207). The focus on the stories, desire to make meaning from experience, the relationship between the researcher and the participants, and the meaning they create together are the tenants of interpretivism and, in turn, narrative inquiry.

**Methodology**

I have chosen narrative inquiry as the method for my journey because narrative inquiry “focuses on stories told by participants” (Grbich, 2013, p. 216). According to Pinnegar and Daynes (2007), “narrative inquiry embraces narrative as both the method and phenomena of study” (p. 4). Clandinin and Connelly (2000) support the notion of narrative as both method and phenomena. Clandinin (2016) defines narrative inquiry as, “an approach to studying human lives conceived as a way of honoring lived experience as a source of important knowledge and understanding” (p. 17). According to Riessman (1993), “individuals construct past events and actions in personal narrative to claim identities and construct lives” (p. 2). Thus, I wish to honor the lived experience of my participants as a source of knowledge and understanding.

**Rationale for Narrative Inquiry**

The purpose of using narrative inquiry is to explore the stories of my participants, in particular, their ethical and emotional experiences of being K-12 educational administrators during a time the American public education system is hyper-focused on high-stakes assessments as a means of accountability (Amrein-Beardsly, 2009; Au, 2007, 2008, 2009a, 2009b, 2011, 2016; Berliner, 2011; Booher-Jennings, 2005; Horn, 2003; Kempf, 2016; Lipman, 2004; McNeil, 2005; Neal & Schanzenbach, 2010; Nelso, 2013; Nichols & Berliner, 2008; Nichols, Glass, & Berliner, 2012; Padilla, 2005; Pandina Scott, Callahan, & Urquhart, 2008; Torres, 2004;
Valenzulela, 2005, Vasquez Helig & Darling-Hammond, 2008). As Clandinin (2016) suggests through narrative the researcher deliberately has a relationship with our participants, and:

We, as inquirers, think narratively about our experiences, about our participants’ experiences, and about those experiences that become visible as we live alongside, telling our own stories, hear an other's stories, moving in and acting in the places - the contexts - in which our lives meet. (p. 23)

As the researcher I am part of the research in narrative inquiry, and by working with participants I am already acquainted with I was interested where their stories of implementing high-stakes assessments meet with my life and my story. The method of narrative inquiry I took is referred to as “narrative-under-analysis” (Bochner & Riggs, 2014). Narrative-under-analysis is focused on the content or the stories within the narrative rather than on structure or form or, in other words, what participants say rather than how they said it. I was interested in knowing the stories of the tensions they felt when implementing high-stakes assessments. More importantly, I present research where the stories of the participants meet with the stories other aspiring or veteran school leaders.

This research was conducted in a high-stakes assessment educational context. High-stakes assessment refers to any state-mandated test students are required to take that is tied to school ratings, teacher performance ratings, and whether or not a student is deemed “proficient” in reading, math, history, and/ or science. Beginning in states like Texas, Florida, and North Carolina in the 1990s and then adopted nationally with the No Child Left Behind (NCLB) Act of 2001 and until December of 2015, the state and federal mandates demanded students be proficient in reading and math. Although students did not have to demonstrate proficiency in each and every grade, students had to be proficient in benchmark years. However, NCLB did
mandate students take a state assessment annually. In December 2015, President Obama signed the Every Student Succeeds Act (ESSA, 2015) into law. Taking the place of NCLB, ESSA made some significant changes, but the requirement of an annual assessment remained. In Florida, students who are not proficient on state reading assessments may be retained in third grade, may be required to take additional remedial reading courses, and must demonstrate proficiency by graduation on the state assessment or another assessment the state has determined an acceptable concordance assessment, such as ACT or SAT, or the student receives a certificate of completion rather than a standard diploma (The Florida K-20 Education Code, 1995).

Enactment and implementation in the context of my journey refer to how educational leaders practice educational policy implementation within their organization. Policies guiding high-stakes assessments are implemented by each organization differently. Thus, part of the narrative inquiry focuses on the ways each participant carried out high-stakes assessment policy at their level during testing season.

My intent was to hear the stories of K-12 school administrators’ experiences of implementing high-stakes assessments during testing season. Through learning their stories, I raise questions about the purpose of high-stakes assessment and the impact of policies at the school level and illustrate how my participants, school administrators, negotiate ethical tensions arising during implementation. Further, this research seeks to illuminate how educational leaders work within the organizational boundaries guiding the implementation of high-stakes assessment. I hope this work shines a light on how they rationalize or reconcile the differences between their own personal ethics and the requirements of their jobs. By exploring their experiences, I highlight how educational leaders navigate these tensions. The purpose of my journey is to explore the stories of the ethical tensions K-12 educational administrators navigate.
when implementing high-stakes assessment policy in a culture of measurement during testing season. Some educational leaders, in particular, K-12 school and district administrators, struggle with the tensions existing between their own personal belief systems, organizational dilemmas, and the requirements of enacting high-stakes assessment policies.

My research questions are: 1) In what ways do K-12 educational administrators navigate potential ethical dilemmas between personal ethical obligations to students versus organizational directives associated with the implementation of high-stakes assessment policy? 2) How do K-12 educational administrators experience and live the testing season? 3) How does the culture of measurement play out in the stories of K-12 educational administrators?

Methods

Participant Recruitment and Selection

The goal of my travels was to gather the narratives from four K-12 school-level educational administrators. The eligibility criteria for the people participating in my study were they must have been working as a K-12 public school administrator in Florida because they were directly tied to high-stakes assessment implementation and this led to a deeper understanding of the impact of high-stakes assessment policies at a school organizational level. Preference was given to people I already knew who worked in the same region as I because I wanted my research to be relational and situated in common understandings. Wolf (1996) refers to this as “situated knowledge” reflecting shared knowledge of location and a relational positionality (p. 14). Working in the same state and in the same regional area means we have all experienced the similar directives for high-stakes assessment policy implementation and are in similar relational positions of school-level administrators responsible at some level for carrying out the
implementation. My research has benefited from a shared understanding of the state and institutional organizations and relational positionality.

I asked four school administrators to participate in my narrative inquiry because in separate incidents they have expressed frustration with high-stakes assessments. I approached each participant in person and explained what the purpose of my journey was and asked him or her to be a participant in my study. Each one agreed to participate. I collected stories from Nora, a middle school principal, Cher, a middle school senior assistant principal, Michelle, a high school senior assistant principal, and Dave, a testing administrator. A senior assistant principal is tasked with overseeing the implementation of high-stakes assessment at the school level. A testing administrator is a teacher whose sole job is to oversee the implementation of high-stakes assessments at the school level. Each participant chose a pseudonym.

**Nora.** Nora was the fourth stop on my travels. She is a middle school principal. She has been a principal at her current school, Blue Middle School for six years; prior to that she was a principal at Crimson middle school for five years, where she hired me as a teacher. She has been an educator in the school district for 30 years, twenty of which she has been a school administrator. We have known each other for twelve years.

**Cher.** Cher is the fifth stop on my journey. Cher has been a school administrator for three years. She is in her second year as a senior assistant principal at Brown middle school in the same school district as I. Cher was a junior assistant principal at the same school for one year prior to becoming the senior assistant principal. She was a teacher peer evaluator and district-level reading resource teacher prior to becoming an administrator. She has worked for the school district for a total of 22 years. We met when she was training a new teacher reading
course I took part in when I was in my first three years if teaching. We have known each other for eleven years.

Michelle. Michelle is the sixth stop on my journey. Michelle is a senior assistant principal at a high needs high school, Orange High School, the same school where I work. She has been in this role for five years. Prior to that she was a junior assistant principal at the same school for three years. She has approximately 15 years of experience in the school district. We met when I was hired to be a junior assistant principal at Orange High School. We have known each other for three years.

Dave. Dave is the seventh stop on my travels. Dave is a testing administrator at Bronze Middle School in a large school district bordering the one Nora, Michelle, Cher, and I work in. Dave worked as a high school math teacher for six years before becoming an administrative resource teacher at a high school for one year, which is a teaching position with administrative responsibilities, such as handling students’ affairs. Dave was testing administrator at Pink Middle School before moving into the same role at Bronze Middle School. Dave is also a doctoral student in the same program as I. He began his program a few years after I did, thus we did not know each other except by face prior to this study.

People, places, and motivations. Each of my participants represents the places I am stopping at and traveling through on my quest to reach the sacred destination: attainment and creation of knowledge. Although it is my journey, these people tell tales along my way. It is their stories, their motivations and struggles with high-stakes assessment implementation, which propel me down the path of my pilgrimage toward my destination.
**Story Collection**

**Site.** I met with my participants at a location of their choosing. I met with three administrators at my home on three separate occasions for focus group interviews. I met with a fourth participant twice at the University we both attend. I conducted a follow-up interview with two of my participants at one of their homes. Although four of the five of us work within the same large school district in the southeastern United States, I did not use any of the school district’s buildings as sites for interviewing my participants because I wanted my participants to feel comfortable sharing their stories without worrying about who might overhear.

I used three primary forms of story collection: electronic diary entries, focus group interviews, and one-on-one interviews to gather stories from my participants:

**Focus group interviews.** The focus group interviews were designed to be more like conversations between the participants and myself than like a formal interview (Berg & Lune, 2012). The focus group interviews took place at the beginning, the middle, and the end of testing season in Florida. My initial focus group interview took place in early February. I took part of the time to explain my plans for story collection, selection, and analysis and provide each participant with consent forms. Three of the four participants attended the first focus group interview: Nora, Cher, and Michelle. One participant, Dave, was unable to attend the initial focus group meeting. The focus group meeting was recorded and any stories shared were transcribed; however, the collection of stories was not the primary goal for the initial meeting. Prior to the first focus group meeting, I asked each participant to bring me a list of questions they might have about the research, the process, and questions they have regarding high-stakes assessment and ethical dilemmas we all experience. However, none of the participants prepared any questions.
During April, I conducted my second focus group. Three of the four participants attended the second focus group interview: Nora, Cher, and Michelle. Dave was unable to attend this focus group interview. This meeting was recorded and guided by interview questions, which were open-ended and were crafted to elicit personal experiences from my participants (see Appendix A). A great deal of my questions begin with the following stem: “tell me about a time when...” The questions are not necessarily sequenced; rather I adjusted my questioning based on how the conversation went (Gibson & Brown, 2009). I was prepared to include school grade or district grade data as entry points for eliciting stories from my participants; however, my participants were forthcoming with stories and made it evident during our conversations they were well aware of the school and district data, thus it was unnecessary.

In late May I conducted my final focus group meeting, again using open-ended questions crafted to elicit personal experiences from my participants (see Appendix A). Three of the four participants attended the third focus group interview: Nora, Cher, and Michelle. Dave was unable to attend the final focus group meeting. At this final meeting I collected the electronic diaries via digital voice recorders from each participant. In mid-June, I met with Nora and Cher to conduct a follow-up interview, during which I used questions I had about stories they had shared during the previous three focus group interviews. I recorded this meeting using a digital recorder. The total time of the focus group recordings was approximately five hours long, and the follow-up interview with Nora and Cher was approximately one hour.

**Electronic diaries.** During the first meeting, I provided each participant with the digital voice recorder they used to record their electronic diary entries (Bruner, 1986). The participants were asked to record stories of implementing high-stakes assessment during testing season, which roughly begins in late February and ends in early May. Participants were encouraged to
record weekly, but they could record more often if they so choose. The participants were provided several open-ended questions to guide them in their electronic diary entries (see Appendix A). Unfortunately, my participants did not use the electronic diaries much in spite of weekly text-message reminders from me. In fact, the electronic diary recordings amounted to less than 30 minutes from Nora and Cher. Michelle and Dave did not make any electronic diary entries.

**One-on-one interviews.** I met with Dave at the University for a few minutes in February to give him the voice-recording device and provide him with his consent forms and questions to guide his electronic diary entries. Then, in late June, I met with Dave to conduct an interview and collect his electronic diary. I used the same questions I used during the focus group interviews to guide the conversation and elicit stories from him. I recorded this meeting using a digital recorder. In mid-July, I met with Dave again to conduct a follow-up interview. I also conducted a followed up with Michelle one-on-one after school at a local café in June. The follow-up interviews were each approximately one hour long. The one-on-one interviews were approximately 2 hours long.

**Transcription.** I only transcribed the stories recorded during the focus group meetings, the one-on-one interviews, the follow-up interviews, and the stories shared using digital diary entries rather than transcribing each meeting and diary entry word for word since the focus of the story analysis is on narratives. I used narrative elements described in the story selection section below and the concepts of the culture of measurement, testing season, and the ethics of implanting high-stakes assessments drawn from the literature to guide my decisions about what I ultimately transcribed.
Prior to, throughout, and after collecting stories from my participants, I was simultaneously collecting my own stories of ethical dilemmas I feel when enacting high-stakes assessment policies as well as noting my responses to their interviews and the narratives. For several years now, I have been writing about the concerns I have with high-stakes assessment in my practice as a teacher and now as an administrator. As well as facilitating the focus group, I shared my stories during the focus groups as well because I am a participant/researcher, and my stories may have sparked a memory for my participants’ and/or created a setting for my participants to feel more comfortable sharing their stories. Additionally, I wrote in response to the stories collected during the interviews after each interview. Ellis (2017) suggests, “we come to understand others through our own self understanding and we come to understand ourselves through others” (p. 434). Thus, narrative inquiry is a reciprocal process. Lastly, the final chapters in this dissertation will contain some of my story and my conclusion to the narrative inquiry.

Once I collected stories from my participants via electronic diaries, focus groups, and interviews, I began the transcription process. To transcribe I used YouTube’s closed captioning feature. To do this, I combined the digital recordings with random photos from my library in the iMovie application to create a video file. From the iMovie application, I was able to upload each file to YouTube being careful to make sure the file was marked “private” to ensure the files remained confidential. I then turned on the closed captioning feature for each video. After YouTube added the closed captioning to each file, I then downloaded the text. To do this, I clicked on each video and followed these steps:

- Paused the video,
- Right-clicked on the web page,
• Clicked “Inspect,”
• Clicked “Network” in the sidebar,
• Turned on the closed captioning by clicking “CC” at the bottom of the video,
• Right clicked on “timed text” in the list of data URLs,
• Clicked on “Open in New Tab,”
• Copied the html text,
• Opened an Excel document,
• Pasted the text in Exel,
• Clicked “Edit,”
• Clicked “Replace,”
• Typed “<*>” in the Find box and pressed the space bar once in the Replace box,
• Clicked “Replace All,”
• Copied the text,
• Opened Text Editor application,
• Pasted the text in TextEdit application,
• Clicked “Format,”
• Clicked “Make Plain Text,”
• Copied the text,
• Opened a Word document,
• Pasted the text in the Word Document,
• Used Replace to remove any excess spaces between words, and
• Formatted the transcript.
Once I completed these steps for each file, I listened to the recordings as I followed along with the transcription. I did not fix every mistake or make sure each word was accurate because I chose to only check for accuracy once I had completed the analysis. At that time I focused on checking for accuracy for only the stories and narratives used in each participants’ respective chapter. The transcription resulted in approximately 50 double-spaced pages for each of my four participants.

**Story Selection**

I explored the transcripts looking for the following narrative elements:

1. **Sense of Temporality** (Birkerts, 2008; Bochner & Riggs, 2009; Crites, 1986; Freeman, 1998; Kerby, 1991),

2. **Context, setting, or place** (Bochner & Riggs, 2014; Coulter & Smith, 2009),

3. **Characters and, in many cases, dialogue** (Bochner & Riggs, 2014; Coulter & Smith, 2009),

4. **Moral or point of the story** (White, 1981; Bochner & Riggs, 2014), and


Temporality refers to the sequence of events unfolding in a story. Context situates the story in a place. Characters are the people the narrative is about; the rising action is the crisis centering and moving the story forward to a resolution (Bochner & Riggs, 2014). The moral of the story is the reason the narrative is told in the first place, it “gives meaning and value to experiences depicted” (Bochner & Riggs, 2014, p. 202). White (1981) suggests narratives inherently include implicit or explicit moralizing. As we narrate stories to one another we work to find the purpose or the moral of the stories we tell. Indeed, “the moral contradictions in our personal lives more than resonate with those in our social order, our nation’s politics, our culture” (Coles, 1989, p. 202).
It is in the moral contradictions of the politics of assessment playing out in the moral contradiction of the lives of my participants as they work to find the greater purpose or moral of their stories I sought. I looked for all of these narrative elements throughout the transcripts, but it is possible for a story to be missing one or more of the elements and missing an element did not necessarily lead me to exclude that particular narrative. However, I excluded narratives with all the elements but without any relationship to my research questions.

**Story Analysis**

The goal of the analysis was to locate stories grounded in the concepts drawn from the literature on testing season, the culture of measurement, and the ethics of implementing high-stakes assessment policy within the electronic diaries, the focus group, and interview transcripts. An important aspect I was looking for during my story analysis were stories related to the purpose of my pilgrimage, and I excluded stories unrelated to ethical dilemmas, testing season, or high-stakes assessment. I chose narratives collected from the electronic, focus group, and interview transcripts, and I present the participants’ narratives in a chapter devoted to each participant in which I included and connected relevant research from the literature.

After I completed the transcription of the recordings, I isolated each participant’s narratives into four individual documents. Again, I listened to the recordings and at this point I made corrections to the transcript to ensure accuracy. As I listened, I took notes and also noted in the document the stories in which my participants shared about their ethical conflicts with the testing season, ethical issues concerning the implementation of high-stakes assessments, and the culture of measurement. I additionally noted where the literature reviewed in chapter two connected to my participants’ stories.
Once the transcription and analysis were complete, I gave participants the opportunity to provide feedback, which I offered to add to their chapter or append depending on the nature of the feedback (Wolgemuth, personal communication). This way of reporting findings attended to the relationships I have with the participants and worked to ensure my story collection, selection, analysis, and reporting remained ethical. The inclusion of participant feedback comes in the form of comments, corrections, or additions to stories they shared with me. However, after reviewing transcripts and final drafts of their chapters none of the participants provided any feedback other than to tell me everything was accurate.

**Presentation.** I took an anti-reductionist stance on the application of narrative inquiry to the stories shared by my participants. I privileged the storyteller instead of the analyst, myself. I sought to talk to, to talk with, and to “inquire as an empathetic [witness] on behalf of [my] research participants” (Bochner & Riggs, 2014, p. 206). As the researcher, I wished to present the stories of my participants in an alternative to traditional narrative analysis. Bochner and Riggs (2014) suggest:

> The stories seek to activate subjectivity and compel emotional responses from readers; they long to be used rather than analyzed, to be told and retold rather than theorized and settled, to offer lessons for further conversation rather than truths without any rivals, and the promise companionship of intimate detail as a substitute for loneliness of abstracted facts. (p. 206)

The narrative analysis I utilized is referred to as “narrative exceptionalism,” which embraces caring relationships with participants rather than treating them as mere subjects to be studied (Bochner & Riggs, 2014). Indeed, Riessman (1993) posits, “precisely because they are essential meaning-making structures, narratives must be preserved, not fractured, by investigators, who
must respect respondents’ ways of constructing meaning.” (p. 4). Thus, I present and analyze the narratives of my participants in four separate chapters with each chapter devoted exclusively to one participant.

Rather than approach the narrative analysis using the socio-linguistic or socio-cultural approaches to narrative outlined by Grbich (2013), I have chosen to use Tierney’s (1994) method of narrative analysis, which is aligned more closely with life history methods. There are several reasons why I have made this choice. First, I agree with the five criticisms Grbich (2013) has of the socio-linguistic approach: (a) the focus is on the text and not on the human interaction, (b) the assumptions about reality and logistic representation, (c) the idea “it possible to break down all narratives into units of meaning and map them in such a way that their common properties will be revealed” (p. 219), (d) the context is left out, and (e) the life of a human described in a narrative is not always presented in a sequential manner. Second, I approached narratives as complete and whole stories not broken down into units of meaning. Third, I approached all narratives in context.

Although the socio-cultural approach to narrative analysis comes closer to the approach I took, Grbich (2013) explains this approach requires the researcher “interpret stories” (p. 222). Throughout the interpretation researchers must be aware of her “own positions and reactions and how these shape the text” (p. 222). I think my participants’ narratives deserve more. Thus, I present their stories in conjunction with my experience of knowing them and learning about their experiences with them. I followed Tierney’s (1994) model and interwove the story of my history of working with these educational leaders and the thoughts and feelings I had in response to each of their narratives. Tierney (1994) describes himself and his research participant as engaging “in an encounter that sought to enable us to understand one another and to allow others to understand
the challenges” of his participant (p. 113). Thus, the narrative analysis I present are the stories of my encounters with my participants, our efforts to understand their experiences, and hopefully, the presentation of this narrative inquiry will allow others to understand each participant’s experiences in the context of being a school leader implementing high-stakes assessments policies during testing season.

Tierney’s (1994) approach to narrative analysis and life history is in line with my belief system and an interpretivist epistemology. According to Richardson (1990), “narrative provides powerful access to this uniquely human experience” (p. 22). Tierney’s (1994) way of providing access to a unique human experience is inspiring and beneficial. Rather than simply presenting and interpreting these leader’s stories, Tierney’s (1994) approach allowed me to “think with [each] story” and write with it (Bochner & Riggs, 2014, p. 207).

Entering into the thoughtful acts of collecting, compiling, and presenting these narratives I was well aware of my position as the researcher, and, as Frank (2000) suggests, my “standpoint as the storyteller” (p. 357). Frank (2000) refers to a standpoint as, “a political and ethical act of self-reflection: to take a standpoint means to privilege certain aspects” (p. 356) of my own story and the stories shared with me. Frank (2000) also points to community membership as common experience, but within the community each person has a distinctive experience. Thus, my membership in the community of educational administrators certainly impacts my standpoint, as does my unique experience as a former reading coach. According to Bochner and Riggs (2014), “the burden of the academic storyteller is to find the story in the experience” (p. 203). As the researcher making the choices puts me in the position of authority or power, but I attempt to mitigate that by providing space for the stories to lead the research rather than the research taking the lead. The ultimate goal of presenting my research in this manner is to honor the narratives of
my participants and open up spaces for dialogue and understanding concerning the ethical
dilemmas of enacting high-stakes assessment policies during testing season.

**Reflections on the limitations.** By choosing to work with participants I know well who
also work within the same school district as I do, my closeness to the stories was a limitation of
this study. As discussed earlier, the relational aspects of this study have advantages. However,
they also have disadvantages. One disadvantage was my closeness to my participants, which, at
times, rendered me unable to identify or unwilling to address ethical dilemmas in their narratives.
The cultural norms of our school district I am embedded within impacted my ability to critically
“see” ethical dilemmas. At other points, I found my friendships with my participants left me
feeling reluctant to interrogate their beliefs, language, or actions. It was through my analysis of
Dave’s narrative and through many critical conversations with my major professor I was able to
recognize these limitations. I analyzed the stories further to reveal any cultural norms or
reluctance I may not have addressed in my initial story analysis. Additionally, I further
addressed the ethical dilemmas in chapter eight along with noting my own ethical obligations as
a friend and employee. As with all qualitative research, the findings of this study are not
generalizable. However, the results may be transferable.

**Attending to the relationship.** I followed Ospal, Woglemuth, Cross, Kaatna,
Dickmann, Colomer, and Erdil-Moody’s (2015), “strategies for attending to the relationship,” to
assist me with ethical considerations (p. 9). As noted in Figure 2, Ospal et al. (2015) have
recommendations for interviewing, methods, and data analysis, as well as material assistance and
emotional resources. Ellis (2017) explains relational ethics are intentional reflections on the
researchers part to consider our “role, motives, and feelings” before during and after an interview
(p. 435). I think it is important to note my role as a researcher in this particular instance does
not wholly place me in a position of power because the administrators I collected stories from work in the same school district as I, and they are either in jobs equal to mine or above mine in the organization’s hierarchy. Although I determined what stories I used and which I did not use in the final dissertation, which is certainly a position of power, I was not interviewing marginalized groups, per se. However, to further attend to the intentionality of the relational ethics involved with eliciting stories from my participants and following Ospal et al.’s (2015) recommendations, I created focus group protocols, which I provided participants prior to the focus group.

*Figure 2. Strategies to Attend to the Relationship.*

<table>
<thead>
<tr>
<th>1. Interviewing strategies</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Validating participants experiences</td>
</tr>
<tr>
<td>b. Listening skills</td>
</tr>
<tr>
<td>c. Personal disclosure</td>
</tr>
<tr>
<td>d. Opportunities for self-reflection</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>2. Methodological strategies</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Multiple interviews</td>
</tr>
<tr>
<td>b. Location chosen by participant</td>
</tr>
<tr>
<td>c. Engaging relevant others (principals, supervisors)</td>
</tr>
<tr>
<td>d. Inquiry of discomfort</td>
</tr>
<tr>
<td>e. Community-based research</td>
</tr>
<tr>
<td>f. “What else should I know about”</td>
</tr>
<tr>
<td>g. Explicit opportunities to reflect on the experience</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>3. Material assistance</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Material assistance</td>
</tr>
<tr>
<td>b. Referral to community resources</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>4. Emotional resources</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Emotional support/processing of experiences/safe discussion</td>
</tr>
<tr>
<td>b. Sharing information about their concerns</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>5. Data analysis/reporting findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Member checking</td>
</tr>
<tr>
<td>b. Sharing of final results with community</td>
</tr>
<tr>
<td>c. Sharing of final results with participants</td>
</tr>
</tbody>
</table>

(Ospal et al., 2015, p. 9)
Before each focus group and one-on-one interview I reviewed the interview strategies Ospal et al. (2015) suggest, and throughout the conversations I can be heard validating the participants’ experiences, disclosing my own ethical struggles with high-stakes assessment, and providing opportunities for self-reflection. Although it is not something immediately evident in the recordings, I made sure was careful to listen attentively, which is demonstrated by the fact that I rarely disrupt my participants. In fact, throughout the recordings, there are times when no one is speaking because I wanted to provide the space for my participants to reflect and provide me with their narratives.

For the methodological strategies following Ospal et al. (2015), I conducted multiple interviews. As previously mentioned, I asked participants to chose the site(s) of the focus group and interviews. Throughout story collection, I remained in close contact with my major professor and kept him informed and engaged in the process. Additionally, I frequently checked to make sure my participants were comfortable with the process, and once completed I checked with each one to provide them with an opportunity to reflect on the experience of participating in the study, as well as asking them if there was anything additional I should know. Due to the nature of the study, none of my participants required any material resources or emotional support.

**Internal review boards and consent.** I sought and received approval from my school district’s Internal Review Board as well as my University’s Internal Review Board (See Appendices B and C). I followed both institutions protocols during this study. I provided my participants with consent forms prior to our first meeting, spent several minutes explaining the contents of the consent forms, allowed them time to review and read the consent forms during the first meeting. My participants received nothing in exchange for participating in my study.
Conclusion

The purpose of this journey was to explore how K-12 educational administrators narrate their lived experiences of working within dilemmas between their personal ethics and the requirements of their jobs. On this stop of my journey, I have outlined the epistemological underpinning for my research. I described my use of narrative inquiry as a method for collecting and analyzing my participants’ stories. In the coming five stops on my travels, I used narrative inquiry to explore the stories of K-12 educational administrators’ experiences of implementing high-stakes assessments during testing season. By eliciting and sharing these stories I raise questions about the purpose of high-stakes assessment and the impact of these policies at the implementation level. Perhaps my research exposes some of the ways educational leaders work within the boundaries of high-stakes assessment. The next stop leads us to my first principal and mentor, Nora.
CHAPTER 4: NORA’S TALE: HERE'S THE CRAPPY THING, WE HAVE TO TEST BUT THE GOOD NEWS IS...

I don’t think it's right but when you're judging myself and my school that's how you're judging me. You have to consider all of that. It's like you're coaching. It's a game and how are you going to strategize to get the most points or the best win out of this situation. I mean if you ignore all that you're saying, “ah I'm sorry but I'm not worrying about how I'm graded?’ This is how you're judged and you're going to just toss that away? I don't think so, Nora explains.

**********

My doorbell rings and I greet Nora. She and I are getting together for dinner at my house, and, as usual, she is warm and engaging. “Hello!” she greets me with a huge smile on her face and draws me in for a hug. Nora is one of the few people who manage to simultaneously be serious about the importance of her work while being the most fun-loving person you have ever met. She is known for making schools a fun place to learn and work.

This stop on my journey has taken me to Nora. Nora hired me into my first teaching job 12 years ago. She has been a principal for 13 years. The first six years Nora was at Crimson Middle School, the middle school with the highest poverty levels in the district, but for the last seven, she has been at Blue Middle School, which is also a Title 1 school. Nora and I have been friends ever since we worked together. Nora’s position as a veteran principal is unique to my journey. In fact, Nora is the only principal in the study, thus I am hoping her perspective will shed light on the answers to the questions that began my pilgrimage. After a light meal with my husband and daughter, Nora and I begin to discuss high-stakes assessment.
Culture of Measurement

I read over the questions you sent. One of the questions you asked about the ethical dilemmas, do you believe in the testing that’s going on? And, quite frankly, the State testing, first of all, I think this district we test way too much, district and/or state. Two, the State testing, which is the FSA at this current time, is basically—we're giving this test to give schools a grade—which basically all you're doing is rating the poverty level of each school, in my opinion. It's not fair for the work that's being done at each school because if your kids come in...and they started reading before they even went to kindergarten and then you have another school where...kids are coming in to kindergarten not knowing how to read...but yet this test rates your school. The test does nothing we don't retain kids [in middle school] whether they are a Level 1, 2, 3, 4, or 5. If they are a one or two they could be placed in remediation classes instead of an elective, but the bottom line is this test is for a grading system that is flawed and is not fair to the work that's being done at the schools. Why are we giving this test that just grades a school? And, what does it really mean? I have an ethical dilemma as an educator, I think we shouldn't be doing this test period or if you're going to do it needs to be for the right reasons and it's not.

Nora points to the persisting achievement gap (Cohen, Garcia, Purdie-Vaughns, Apfel, & Brzustoski, 2009; Haycock, 2001; Ladson-Billings, 2006; Lee, 2002; Reardon, 2011, 2013; Williams, 1996) prevalent between White middle-class students and students who are poor, Black, and Brown. Reardon (2013) suggests the achievement gap has little to do with school because students from poor and racial minorities come to school at a disadvantage, which are due to larger societal problems. Nora argues it is unethical to grade a school using test scores when students come to some schools disadvantaged and other students come to school with advantages, but I wonder if the larger ethical dilemma rests in why the societal problems are not
addressed and some families continue to experience the negative consequences for not being White and middle class. I agree with Nora, high-stakes assessments exist simply to provide schools with a grade, which serves to center the entire educational experience for students and educators on the preparation and administration of high-stakes assessments. This is a systemic problem arising from a state policy, Jeb Bush’s A+ Plan, which requires a grade is given to each school based solely on students’ performance on high-stakes assessments. Further, the school district, in complying with state policy, encourages schools to focus on student test preparation and performance.

Nora states she has an ethical dilemma as an educator because she does not agree with the amount of testing. She wonders what the point of the all testing might be, and seems to struggle with the purpose of high-stakes assessments. I ask her to explain:

*You have been in education for a long time, what do you think some of the adverse effects of this practice, I ask.*

*Well because of the outcome of this test and what the results show, teachers are less apt, who have the heart for struggling kids, are less apt to go to a school and teach at those tough schools because their grade is not going to be as good as it would be if they went to a school that had kids coming in prepared. I find that even the teachers, the new teachers that come into those struggling schools, will leave after a couple of years because of the outcomes of that test. I think that is a disservice to kids, and it's not intentional but the adults are looking at it from, ‘I have to look at me first because this could derail my career based on the outcome.’ I think what I see now, and I hate to say it, teachers teaching to a test, but if that's what you're graded on and assessed on that's the focus. In the past teachers would be more creative and do other lesson enhancements, but they're not apt to do that now because ‘I've got to focus on what this banner*
benchmark is or what the standard is because if the kids don't pass it then it's reflected upon me and the school.’ I don't think they're as apt to adding in supplementary things and being creative as they could be because they feel like ‘I have to do this, I have to do this, I have to do this.’

Nora reiterated what she thought was purposelessness of high-stakes assessment in middle school throughout my conversation with her and is outlined in the literature as geographical opportunity or lack thereof (Au, 2009a, 2009b; Orfield & Lee, 2005; Rodríguez et al.; 2016; Tate IV, 2008). Au (2009) describes the “zip code effect” as the, “deeply entrenched race and class inequality” reflected in the high-stakes assessment scores of students and thusly the grades of the schools located in poorer neighborhoods and neighborhoods with high populations of people of color (p. 1-2). This is true of the school district Nora and I work in, the schools located in the neighborhoods with the highest rates of poverty and large Black and Hispanic populations are predominately schools with the lowest school grades. According to Nora, the low school grade is not reflective of the hard work being done at the schools with the lowest grades.

In my experience, the zip code effect is real, and the hard work is real, but the hard work does not need to be focused on test preparation or administration. This is where school administrators have space to make decisions about curriculum and instruction, for example adopting literacy-rich culturally relevant curriculum could have a profound effect on student learning, which may in turn impact test scores. When I consider how far off the right track the system I work within has gone, it frustrates me and also makes me feel helpless and alone. I feel responsible for subjecting students to test preparation, and I want to stop it, but it looms so large over the landscape of education. I do not know where to begin or what to do.
Further, teachers are driven away from schools with low school grades because the zip code effect equates to poor student performance on high-stakes assessment and is also tied directly to teacher Value Added Measurement (VAM) scores. According to the Florida Department of Education (FLDOE), “value-added models are used to measure the contribution of a teacher or school on student learning” (2013). VAM scores use student growth on high-stakes assessments to determine teacher effectiveness, which is then either combined with a principal’s evaluation for a total teacher annual evaluation or is used alone as an annual teacher evaluation (Collins & Amrein-Beardsley, 2014). Teachers cannot afford to ignore the real consequences of VAM such as loss of wages or even being forced to move out of a classroom mid-year. Therefore, in a culture of measurement, where VAM and performance indicators rate teacher effectiveness, even the teachers who want to work at a high-needs school will leave a job and relocate to another school located in an area more affluent and a larger White population because those students tend to perform better on high-stakes assessments. Nichols and Berliner (2007) argue teachers at high-poverty schools have to work harder than teachers at low-poverty schools.

Additionally, as Nora explains, the curriculum has been narrowed by high-stakes assessments, which is another consequence of the culture of measurement. Nora points to the practice of teaching to the test (Higgins, Miller, & Wegmann, 2006; Jennings & Bearak 2014; Lazear, 2006; Longo, 2010; Menken, 2006; Popham, 2001a; Posner, 2004; Volante, 2004) as one of the adverse effects of high-stakes assessments. According to Popham (2001a) teaching to the test is the practice of focusing classroom instruction on only curriculum covered on a high-stakes assessment and test preparation. Generally teachers, in all types of schools, focus on teaching to the test; however, this practice tends to be more prevalent in low-income schools (Nichols &
Berliner, 2007). As long as student scores on high-stakes assessment judge schools and teachers than teachers, encouraged by administrators, will focus on teaching to the test (Nichols & Berliner, 2007). Thus, as a consequence of the culture of measurement, the narrowing of the curriculum tends to impact the students who are the most vulnerable the most.

I pepper Nora with more questions, Let me ask you a question: you, we, worked at a school that was considered a very high needs school because of the poverty of our students. That school was constantly under pressure to perform. It’s no longer that way for you because you at a different school now, but did working in a school where the state was constantly looking at what you’re doing to improve test scores make you feel extra vulnerable? Did you feel like you had to do things that you would not necessarily do now because of the pressure on you of that particular school?

I have to say no only because I was not an experienced principal when I stepped into that role. You don't know what you don’t know, Nora answers thoughtfully. Honestly, my opinion is, after me, experienced principals were put [at Crimson Middle School], and they didn't survive any longer than two years. I was there for six. The principal that is there now was not an experienced principal walking in, however, this is her third year and she raised her hand and said, ‘I'm done, it's time.’ My personal opinions is, I get that they want experienced people there but experienced people who know what the reality of running a school like that…no one in their right mind would want to go into a place like that.

Nora’s grim reflection of leading a school where over 90% of the students live in poverty, where the achievement gap is a large divide, where the school grade is a D year after year, and where the state is constantly intervening, supports the literature suggesting principal shortages in schools with high rates of students living in poverty and students of color are more prevalent.
than schools in affluent and White communities (Loeb, Kalogrides, & Horng, 2010; Papa, 2007; Roza, 2003; Tekleselassie & Villarreal III, 2011). Further, researchers argue retaining principals in schools with large populations of students living in poverty and students of color is more challenging for school districts (Béteille, Kalogrides, & Loeb, 2012; Papa, 2007; Tekleselassie & Villarreal III, 2011). The moral responsibility of experienced principals should encourage the opposite, but as Nora continues she reveals one of the reasons why there may be an inability to sustain higher school grades:

*I don't see anybody in line—any experienced principals lining up to work at a tough school like that. It does not matter if the experienced principal moves the school [up a letter grade]. It is not sustainable and if you look at these lower end schools they go up then they come down then they go up and they come down, because it's not sustainable. That is the pattern.*

Sustainability of performance over time at high needs schools is rare (Duke & Landahl, 2011; Giles & Hargreaves, 2006; Leithwood, Harris, & Strauss 2010). Leithwood, Harris, and Strauss (2010) suggest sustainability of school improvement “…is a slippery slope. For every success story, there are many more schools which return to their old ways of doing things once special resources and additional support have been removed” (p. 16). In Florida, the state school board has taken to requiring the removal of principals at schools receiving consecutive years of low grades (Solocheck, 2018). In spite of two of the four federal guidelines for school turnaround including the removal of the principal (The Wallace Foundation, 2018), there is little evidence to suggest replacing the school principal can lead to sustained school improvement (Béteille, Kalogrides, & Loeb, 2012; Peck & Reitzug, 2014). Béteille, Kalogrides, and Loeb (2012) found in a longitudinal study of a large urban school district that principal turnover, no matter whether the move was determined by the district or the principals themselves, more often resulted in
lower student performances on high-stakes assessment. Further, researchers (Béteille et al., 2012; Papa, 2007; Tekleselassie & Villarreal III, 2011) found principals often left schools in high-poverty and/or high-minority communities for schools in more affluent and more White communities.

When Nora said, “I don’t see anyone in line” to become a principal at schools where the majority of the students live in poverty and are of color, I imagined people standing in line in front of the doors of some of the schools around our district waiting for an opportunity to become the principal. I also imagined no one in front of Crimson Middle School, and district leadership calling out to people on the street as they walked by asking them to come. Retaining leadership in high-poverty schools is a reality of our school district. Over 40% of our students live in poverty, which is concentrated mostly in urban areas.

Two years ago, in an effort to address this problem, district leaders created a program to entice assistant principals and principals to work at the schools that did the poorest on the state grading system, which also happened to be the schools where the majority of students lived in poverty and were predominately students who were minorities. If a successful principal or assistant principal agreed to work in one of the schools of concern, the successful principal or assistant principal could be eligible for promotion, or at least obtain an interview more quickly than others who did not work at a school of concern. I have mixed feelings about this move. Students at these schools deserve the best, and I admire district leadership’s attempt at addressing a serious problem. Under the previous district leadership, principals were sent to schools struggling to perform well on the state grading system as a punishment, which was horrendous. Thus, this system is certainly better, but at the same time, school administrators should be working at schools because they want to be there not because they may get promoted
after a couple of years. Plus, none of this addresses the massive societal problems causing our students to live in poverty and oppression in the first place.

As Nora and I recalled our shared time at Crimson Middle School, I was struck by some of the deficit language used to describe students and schools. As I continued on my journey, the use of deficit language was common among my participants, and I even found myself using it as well.

**Testing Season**

*Another dilemma, because you have to use computers to do the testing now, and we all don't have one for every child. Because of the fact that most schools do not have enough technology to test all their kids at once we have to do morning sessions and afternoon sessions. Quite frankly I would rather take a test in the morning than sit all day and then take one at the end. That becomes an ethical dilemma because I would rather go in and be fresh and take it rather than waiting four hours and then going. For example, at my school, we have two technology teachers and an intensive math teacher who can't be in their classroom with computers during testing season. So you're moving them out of their curriculum. We have a thousand kids, and we have approximately four to five hundred computers. Over the last three years, I've been able to build that up. I have two more computers on wheels (COWs) coming for next year, but I built that up over the last several years through Title 1 money. I bought two more COWs for next year so that we're even in a better place. However, so our testing window basically looks like this each week for three weeks straight: Monday and Tuesday, all ESE kids with an extra time accommodations test Monday and Tuesday because they have all day long. I have 300, well almost 300 ESE kids, so Monday and Tuesday extra time. Wednesday and Thursday, for example this week eighth graders testing in the morning we had one extra lab with*
some of seventh grade tests in the morning, but sixth grade tested in the afternoon. Keep in mind every Wednesday and Thursday the kids that aren’t testing are in a holding tank. Now fortunately, the teachers are having fun with this. They’re doing Minute to Win It. Some are giving them extra credit like if you’ve missed any work this is your time to do it. So we’re trying to be very productive and at least have fun with the kids too so that they're not going crazy. For Monday, Wednesday, and Friday, because Fridays only make-ups, we actually have our same schedule Monday, Wednesday, and Friday okay? But Tuesdays and Thursdays there's no regular schedule it's test, test, test. This is for three weeks straight!

The managing of student downtime during testing season that Nora describes is not a topic directly addressed in the literature; however, managing student downtime is heavily considered in literature about classroom management (Del Guercio, 2008; Englehart, 2013; Evertson, 1989; Garrett, 2008; MacSuga-Gage, Simonsen, & Briere, 2003; McLeod, Fisher, & Hoover, 2003; Popp, Grant, & Stronge, 2008; Trussell, 2008; Wong, Wong, Rogers, & Brooks, 2012). The literature is clear: students should not be provided too much free time or downtime without something constructive to do during class time (Englehart, 2013; Evertson, 1989; Wong et al., 2012). Yet, research ignores how teachers or administrators should handle downtime they and their students have when the half the school is testing and their half was moved out of their classrooms. This a nuanced byproduct of testing season, which is directly related to the allocation or reallocation of resources.

Nora suggests she has an ethical dilemma with having to test in shifts because she does not have enough computers for all the students to test at the same time. However, I question whether or not this is an ethical dilemma for Nora, but instead a problem she is resigned to deal with as best as she can given the circumstances. As a by product of testing season, school
administrators must make choices to allocate resources in certain ways to accomplish testing in the time allotted by the state. This is a form of the ethic of justice, distributive justice, whereby the school administrator distributes the “benefits and burdens” to the students as evenly as possible (Beckner, 2004, p. 37).

I found Nora’s use of the term, “holding tank,” a curious way of characterizing student downtime during testing season. A holding tank is a place where liquids are temporarily kept or confined. It makes me think of students as fish in a holding tank waiting for a permanent fish tank. By extension, students are like abstract fish with no purpose other than to hold on until their testing tank is ready for them.

Plus, I have a whole testing team and this brings in the technology piece because we don't have enough desktop labs. I do have we do have a lot of technology but it's laptops, it's COWs (computers on wheels). We have three guidance counselors, my media specialist, my reading coach, my tech specialist and my [senior assistant principal] that's the testing committee that's the team. Now my secretary and I, we review schedules. We are that outside set of eyes and say, ‘this isn’t going to work, or you missed somebody.’ They create it and then we check because we don't want to be a part of that planning because they got to give it to fresh eyes to see, ‘what did we miss, what did we not get?’ But the other part of that is we also bring in, we have two grade level team leaders; we bring them in to figure out with the teachers, the ones who are not the testers. They have to create their own schedule with the buy-in from their teachers of what to do with the kids that are not testing because if they make this schedule and they have the buy-in they don't complain. It's a cohesive group, then there are outer layers of this fruit, but look at all this extra work. After that at the end of every day, my media specialist, my reading coach, and my tech person are in the media center making sure because with a
laptop if you don't shut it down completely it will not recharge for next day. Every single piece of technology has to come back into the media center, the mobile ones, they have to shut every single one and we also have to make sure if when you get to the audio part that the mute button isn't on because then the kids go, ‘I can't hear anything!’

Nora’s story of testing season also tells the story of how schools must reallocate resources, such as space, technology, and other capital resources, as well as human resources during this project management-type period of time when most of the high-stakes assessments are administered to students. Rather than seeming to have an ethical tension with the reallocation of resources, Nora’s story suggests she is resigned to acting as the overseer of the project called testing season. Project management is a process which is time-bound, temporary, scheduled, resource necessitated, “execution of activity,” leading to the production of a service or widget (Schwindt, 2005). In this case, testing season is the project Nora and her team manage. An essential aspect of the managing of testing season, is managing the allocation of resources. Nora’s story exemplifies how schools make do with limited resources during testing season by reallocating the resources. For example, her media specialist, reading coach, and technology specialists are reallocated from their normal work assignments to support high-stakes assessments at the end of every workday. Additionally, the COWs are devoted purely to high-stakes assessments during testing season. These capital resources are not used for instruction or student learning during this period of time, and are, instead, allocated to testing.

I, too, have experienced the reallocation of resources due to high-stakes assessments. The story Nora shares about managing resources to accomplish testing within the allotted state time-frame is much like what I have witnessed at all four of the schools I have worked in during
the last 12 years. Indeed, as I described in chapter one, I was a resource teacher who was reallocated to test proctor during testing season at two different schools.

Too much testing. There's too much loss of instructional time where actually curriculum could have been taught but because if you're at a school of like say for example a thousand and you don't have a thousand computers, which I don't think anybody that I know has like a one on one computer situation. The amount of time you have to put these kids in a computer lab or on a laptop and the loss of instruction is honestly and truly I say at a middle school level at a large middle school which where I'm at come February on there's like hardly any instruction from February to the end of the school year very little curriculum instruction because you can't because they might testing. We had a lot of technology but we still can't get it done quickly. The simplest thing would be this if they took us back off the computers and we could do paper and pencil. We could knock this stuff out in a week.

Perhaps the most precious resource diverted away from students during testing season is instructional time. As Nora’s story demonstrates, the project management of testing season is focused primarily on computer availability because in middle and high schools most high-stakes assessments are computer-based. The number of computers at any given school will dictate how much instruction will be missed or at least truncated for students.

I wonder out loud, how many times during this testing season did you think yourself, ‘this is not right?’

Nora is quick to answer, Every day. Every day, every year. It takes away too much instructional time that we’re manipulating, and spent spinning our wheels, and creating schedules of how to just get kids onto a computer or into a testing lab and what are we going to do with the other kids that we don't have a computer for? Which is somewhat a waste of time and
loss of curriculum time from February to May. Don't forget to give all the kids that need extra
time the whole day to test because they get the entire day. So you need for all the ESE-extra
time kids the entire day just to try and give them the time. Scheduling is a nightmare and
instruction loss of time is ridiculous.

As of 2014, twenty percent of the school districts in the country report their technology infrastructure as poor (Levin, 2014). Technology infrastructure includes internet connectivity through broadband or high-speed wireless, district networks, and simply the amount of electronic devices students have access to. Nora refers to the lack of 1:1 devices available in her school even with her purchase of COWs during her tenure at Blue Middle School. There is surprisingly little recent research on the limits of schools’ technology infrastructure and the requirements of online high-stakes assessments. Although Nora has a great deal of technology as compared to many other schools, she still had to offer multiple testing sessions on testing days, and she brings up a good question: Does the time of day a student tests due to the limits of the schools technology infrastructure impact how well a student performs on a high-stakes assessment? In this case, Nora is not trying to allocate resources toward students close to passing the high-stakes assessment but instead how she can get all of her students tested in the time allotted with the resources she has on hand. Nora discusses reallocating existing resources such as moving teachers and students out of classrooms that serve as high-stakes testing labs. These classrooms do not sit empty awaiting student use solely during testing season. Schools with large numbers of students must utilize all available space throughout the school year, but this requires displacing these students and teachers during testing season. This often requires the closing of a schools media center for weeks or even months on end because the media center also serves as a
computer lab. Nora is frustrated by the loss of instructional time. She recognizes this loss as something morally wrong, but again she seems resigned and accepting.

The move to computer-based testing (CBT) has resulted in longer testing windows, or time frames the state allows the schools to conduct assessments, because schools do not have the technology infrastructures in place to handle testing the whole school testing at once like schools did when the tests were purely paper and pencil assessments. The limited technology infrastructure in schools results in the further narrowing of the curriculum during the testing season because of the limited access to classrooms or media centers.

Moreover, the intended use of resource teachers, counselors, secretaries, administrators, and classroom teachers is certainly not to plan for high-stakes assessments or carry out the logistics. Many researchers have noted the narrowing of the curriculum and the loss of instruction in untested subjects (Amrein & Berliner, 2002; Amrein-Beardsley, 2009; Au, 2007, 2009, 2011; Berliner, 2011; Berliner & Glass, 2014; McNeil, 2005; Nelson, 2013; Nichols & Berliner, 2007; Sloan, 2005; Valenzuela, 2005), but after an exhaustive search there was little research the diversion or pirating of non-monetary school resources that Nora describes, with the exception of focusing instruction on students close to passing the high-stakes assessments (Booher-Jennings, 2005; Neal & Schanzenbach, 2010; Reback, 2008; Reback, Rockoff, & Schwartz, 2011). Perhaps this is due in part to the lack of research on high-stakes assessment since CBT began dominating the administration of high-stakes assessments in schools in the early 2010s. Thus the reallocation of school resources are no longer primarily concentrated on preparing some students for the tests but in getting the testing completed. Students are not testing but not getting instruction because the adults must decide how to allocate limited resources during testing season. Half of their classmates are taking their high-stakes assessments
in computer labs, or the teachers and students unable to access their online curriculum because they have been moved out of the computer lab for three weeks to allow for high-stakes assessments, or having to utilize resource personnel for testing and scheduling rather than for their designated purpose are incidents that occur in most of the schools Nora and I are familiar with, thus it begs the questions if this common throughout the state or the country? The demands of high-stakes assessments on the infrastructure and resources of the schools and the districts throughout the state are likely consistent during testing season; however, there is no existing research to help adequately answer this question.

Many of my staff has said, ‘you can make something that sucks sounds so good because it's your presentation and how you present things.’ For example, here's the crappy thing we have to test but the good news is we're doing it all in four weeks this year. They say, ‘you can make shit sound good you really can,’ but it's how you present it. This is bad we have to test and we are going to do it in four weeks. Of course I always have input from them so there's buy-in. ‘If you have a better suggestion? You think we do something better? Come on we're all ears.’ So they feel like they have a say, too. The bottom line is we can groan all you want but that's not going to change anything. It's our attitude and our attitude is projected upon these kids so if you come in here and you're like, ‘uh!’ What do you think the kids are going to say? You can't have that attitude...the other part I say is if you really feel strongly about this, which I do, I can't stand it, write to your legislators, write to the people where you could be heard and maybe make a difference. Because it's not our district, it's not us, it's not me, and it’s not our district. It's the State. Channel those energies in the right places and that's kind of my approach to anything that we do.
So what you're saying is but you're going to present it to the teachers in a manner they find more appealing and then you're asking them to present it to their students in a way they find it more appealing? I wonder.

Correct. The bottom line is we have to do it. Do your best. That's all I can ask for.

The performance Nora describes, when she presents the plans to teachers for high-stakes assessments as her school enters testing season in the most positive manner possible, borders on emotional labor (Hochschild, 1983). Nora does not blame an other but instead blames the state for setting the high-stakes assessment agenda. Nora admits she does not think high-stakes assessments are right for students, but at the same time, she recognizes she must implement the assessments no matter her personal beliefs. Thus, she decides to frame the presentation to her staff in a positive light. This framing could be considered a form of emotional labor. Although Nora’s staff are not customers she must cater to, she is setting the tone for testing throughout the school for her staff through her attitude. She embraces the “professional ethic” to guide her decision to do what she feels is best for her students (Black, 2008a). The professional ethic guiding Nora’s emotional labor during testing season is part of a pragmatic approach she must take as a public school principal.

As a school administrator, I have an internal struggle with pragmatic approaches to the implementation of high-stakes assessment. I must approach situations every day that require practical, reality-driven solutions. I agree with Nora, we cannot ignore the reality that high-stakes assessments judge the school, the students, and the educators. I would love to tell everyone to ignore it all and focus on good pedagogy, but I do not feel I have enough power to accomplish this. I wonder if enough of us pushed back could we make real change?
The WIDA testing which is for ELL students, students that have limited English language. Out of a thousand students we had to test 110, and we received information from the district with what tests need to be administered to certain students whether they were A, B, or C. Apparently we started testing the first day, which by the way, from January 29th until March 23rd is the testing window, and we have only four adults who can administer this test. Each student has to take four subtests: listening, writing, speaking, and reading. The speaking portion has to be a one on one administered test, which takes 20 minutes per child to do. The three different tiers of students: A, which is speaking no English, B, some English, and C, a little bit of English. When you test these students you cannot have more than 22 at a time. Usually, we will do about two sessions a day and it is my three paraprofessionals and the ESOL teacher who administered these tests. You can imagine the amount of time that it takes. We started testing the first day, the problem is by the end of the day we received information from the district [some of the information they sent us] was printed wrong so 15 students were coded incorrectly and we had administered the wrong test to them. Needless to say, we told the students of the error, which was not on our end. We had a retest these students again, which we offered them ice cream and candy, and they were very happy to do it. But needless to say, the fact that we had incorrect data to administer tests was not ideal. One of the issues with this test is these students who have limited English this is an additional required test they have to take because they still take the FSA when it comes about in the months to come. They have to take those tests as well.

The negative impact of high-stakes assessments on ELL students has a rich literature, from research on the implicit bias in high-stakes assessments (Altshuler & Schmautz, 2006), to the inappropriateness of high-stakes assessments for ELL students (Nichols & Berliner, 2007; Menken, 2006, 2010; Padilla, 2005; Palmer & Snodgrass Rangel, 2011; Solórzano, 2008;
Valenzuela, 2005; Wright & Li, 2008), to calls for high-stakes assessments to be given to students in their first language (Escamilla, Mahon, Riley-Bernal, & Rutledge, 2003). Nora tries to mitigate the negative impact of these 15 ELL students having to retake the speaking portion of the WIDA test due to an organizational mistake by compensating her students with ice cream. Part of implementing high-stakes assessment for those at the implementation level includes dealing with organizational and systemic mistakes caused by the state and/or the district. The fact the state and/or the district caused the mistake does not exempt schools from the state’s requirements, such as testing 95% of the students mandated to take each assessment. Therefore, Nora could not excuse the students who were given the inaccurate WIDA from taking the correct version, but she still felt she needed to compensate the students and encourage them to take the correct assessment. Her leadership approach in this situation is transactional. According to Pepper (2010), high-stakes accountability policies require the school principal to be transformational in some situations and transactional in others. Here Nora needed her students to test again in exchange she provided the students with a treat. She worked within the guidelines of the policy but worked to minimize the negative impact on her students within those parameters.

Further, Nora’s story touches on another resource, time. The time the students, the paraprofessionals, and the teachers use to administer WIDA is extensive. Additional time was taken from instruction because the students were initially given the wrong test. Everyone lost instructional time in this story. Time is most certainly a resource that cannot be replenished. According to Adler (2000), time is both a “visible and invisible” resource in schools. Time for planning, time for testing, and time for instruction are visible or known and scheduled resources,
but testing eats away at instructional time in various invisible ways as noted by Nora in this story.

**Ethics of Implementation**

*At middle school level, there's things I don't like—the pressure that is put on these kids because they do feel it—*

*Clarify for me, I interrupt.*

*Let me give you this one scenario one of my kids, which by the way her sister was the valedictorian at the high school our feeder high school, Blueberry High School...Last year when she took the Algebra EOC, as a 7th grader, she did not pass with a three so she had to come to summer school. She did the remediation class took the test again and passed with a three. So this year she was in Geo as an 8th grader and she had to take the EOC. Well we don't have the results back yet. So we get a red flag alert on child this because she wrote on there, 'oh my god I am—' troubled child alert then we have to do an investigation because what if—'I fail this test my life is over'—yes! 'I'm going to kill myself no no I'm not going to kill myself.' We get these red alert investigations that are sent to us and the day you get it you have to investigate. So we call student G in and I am thinking, 'Oh my God this is my student. That is my student assistant.' So my guidance counselor called her in we're talking and she goes 'listen I only wrote that because I finished the test but I'm freaking out because I remember I didn't get a three on the Algebra EOC last year and my sister's graduating valedictorian at high school Blueberry and literally and I'm sitting thinking 'oh my god I do not want to go to summer school again' but she goes 'I'm writing this down because I'm freaking out for a moment.' She goes, 'But then I tried to delete it and I hit the wrong button and it went through.' We called mom we told her what happened. What's the moral of the story? She says, 'Don't write my feelings down because I*
wasn’t serious but I was panicking at the time’ and she goes, ‘but I hit the little button.’ So this is what these tests are doing to our children. No, this is the pressure that is being put on our students. And we get this information about this student, this red alert, a month later. What if she was serious? How is this red alert helpful a month later?

Test anxiety for students is an unintended consequence of high-stakes assessments. According to Gregor (2005), between 10% and 40% of students experience test anxiety. Additionally, Putwain (2007) found students of color, students living in poverty, and girls were more likely to experience test anxiety. “Test anxiety comprises psychological, physiological, and behavioral reactions that occur in association with concerns about the negative outcomes resulting from failure or poor performance in evaluative situations” (Segool, Carlson, Goforth, Von Der Embse, & Barterian, 2013, p. 489). Nora’s student had good reason to feel anxious about passing the Geometry EOC, and the red alert was undoubtedly warranted, yet, to echo Nora’s concerns, what is the point of a red alert a month after the assessment? If student G was genuinely planning on hurting herself, the state’s response would have likely been too late.

I have witnessed students freaking out over tests many times, but I have also witnessed students “Christmas Tree” high-stakes assessments many times as well. Christmas treeing is when students just guess all the answers. Perhaps this is because I work in high school and at this point many students have been over-tested and seeing little success on high-stakes assessments over the years, they have given up. Or, perhaps, students did this as a silent protest to being over-tested. I have begged students to at least try on high-stakes assessments more times than I can count. Is this wrong? Maybe. If I am being honest with myself, I partly begged students to do their best because I knew the consequences of doing poorly on high-stakes assessments, but also I begged them to do well because if they did well, I did well.
So here is a good question, is the system that we're currently working with in this state anyway, maybe nationally, but more so statewide I think, setting it up so that we are always constantly figuring out how we can game the system? I ask Nora.

Of course.

Why do you say that? I question.

Just from district personnel. I had this person who's in charge of [a large and important department] for the district actually come to a meeting one time and say that the middle schools should not have, for example, geometry or algebra or these industry certifications because we're taking points away from the high schools. I'm taking your acceleration points away. Wait a minute it's okay for you but it's not okay for us? Because middle school we got nothing. So when you have one district person whose questions why I have geometry and how many middle schools have the geometry, which are very few, 'you shouldn't be having that because you're taking away from a high school.'

So we should limits opportunities for students? I suggest.

That is my problem. If I've got kids that deserve this and need to be challenged and moved on you're telling me no because it's points for high school. Come on! That didn't get very far because we still have it, but that was brought up. I heard that loud and clear, and then I bring it up to the middle school director and say, 'what do you mean?' and then to the content person, 'this is what was told to us in a meeting' and they are like, 'What!?!' So do we consider the grade or do we consider the student?

Nora’s story exemplifies the “existence of calculation” (Ball, 2003b, p.215), whereby leaders are focused on calculating how to best game the system, rather than their ethical obligations to students. In this scenario, the district leader is suggesting students not be offered
advanced math courses in middle school because of the negative impact this may have on high schools. Nora aptly describes this practice as unfair and recognizes the ethical implications.

Is there a place for high-stakes assessment or is it just there should be some level accountability? I wonder.

What's the one test that we're measured across the United States?

NAEP, I say.

We do that now. Well, not all schools do it, but we get picked every year. We get voluntarily picked every year to administer it. Here's my thing, if we have that assessment and it takes what? One day? One day! That's a comparison with every state in the United States nationally why can’t we do that? The FSA is money and it's money being filtered to these people because we have no comparison outside of Florida with this the way they grade our schools.

I agree, the people are making the test are funding the lobbyists who are talking to politicians.

Correct it’s all a money game, but it's not for the benefit of our students or us. Not at all.

Indeed, Nora’s suggestion is supported by research conducted by Amrein and Berliner in 2002. They suggest improved student learning as demonstrated by state high-stakes assessment performance are not supported by student performance on the NAEP. Amrein and Berliner (2002) presented data showing increased student performance on state-level assessments but showing stagnant student performance on NAEP during the same period of time. Likewise, Jeffery (2009) reported a large discrepancy between state writing assessments and NAEP writing performance. Amrein and Berliner (2002) argue states distort or inflate student results on high-stakes tests. The practice of changing the cut score, the number above which is passing and below is failing, every year is commonplace in Florida.
Further, Nora questions the big business of high-stakes assessment. Rightly so, according to Reingold (2015) magazine, in the United States, “the business of assessing students through high school has grown 57% in just the past three years, to $2.5 billion.” Companies such as Pearson, American Institutes for Research (AIR), and ETS have profited nicely from creating high-stakes assessments for states since the federal government began requiring it via NCLB. When NCLB was passed, over $400 million per year was allocated by Congress for the creation of high-stakes assessments (Levine & Levine, 2013). This is just funding set aside for the actual tests and does not include test preparation materials or practice test materials, which school districts and states have heavily invested in for the last eighteen years. The very same companies creating high-stakes assessments are also creating the practice and preparation materials. Nora has good reason to question who is profiting and the ethics of the whole high-stakes assessment system.

Conclusion

As the only principal on my journey, Nora’s voice is an important one. Her stories of testing season provide a broader perspective on the impact high-stakes assessments have on the daily logistics of running a school and the motivations and tough decisions she must make about the use of resources. This aspect of high-stakes assessment was something my travels have revealed about the current nature of high-stakes assessment in a Florida middle school. Her perspective on the culture of measurement makes clear she struggles with how students are adversely affected by high-stakes assessments. Nora questions the purpose of testing students to determine school grades when socio-economics seem to be the determining factors, which was something I was aware prior to my journey. Additionally, Nora questions the ethics of those who stand to profit from high-stakes assessments. Nora’s stories have helped
illuminate for me the problematic predicament school principals are in when they are tasked with implementing high-stakes assessments. My journey has only begun. The next stop will further uncover the ethical dilemmas school administrators grapple with when implementing high-stakes assessments.

**Vignette**

‘Mrs. Canady?’ A voice on my walkie-talkie demands my attention.

‘Go ahead.’ I answer.

‘Ms. Conway just called and said she caught Jefferson with his phone out during testing. Room 505,’ my secretary informs me.

‘Copy.’ I reply. I am already walking to 505. I am totally aggravated. I have a ton of things to do and being taken off-task because a student had his phone out during testing was not in my plan. This is going to take a lot of time. I arrive at 505, let myself in, and I am greeted by Ms. Conway, who calmly explains the situation.

‘I collected all of their phones before we began testing. Jefferson told me he didn’t have a phone. Then I was walking around and I saw him staring down in his lap for quite a long time. I went over and saw him texting on his phone.’ Ms. Conway hands me Jefferson’s phone.

As I listen to Ms. Conway, I survey the class. There are maybe 15 students sitting at computers. Some are looking at the computer and some are writing on pieces of paper squeezed into the small space between the keyboard, mouse, and the edge of the desk.

‘Ok, Jefferson, get your stuff let’s go.’ I watch Jefferson get his bookbag and we head back to my office. Along the way, I begin to lecture him in my best scolding mommy tone.

‘What were you thinking? You know there are no phones during testing.’

‘I forgot,’ Jefferson shrugs.
‘How could you forget? Ms. Conway told everyone to turn their phones in.’

‘I didn’t hear her.’

‘Come on Jefferson, this isn’t your first rodeo. You’re a junior. You’ve sat through dozens of tests and you know darn well you cannot have a phone during a test.’

Jefferson just shrugs his shoulders.

‘Now your test is going to be invalidated. I have to go through your phone and make sure you didn’t take any pictures of the test—’

‘I didn’t! I was texting my mom. My grandmother is in the hospital. It was an emergency,’ Jefferson claims.

His defense does not move me. I have been in education for over twelve years, and practically every student caught texting or using their phone when they shouldn’t say they contacting their moms because of a family emergency. We arrive in my office and I motion for Jefferson to sit down.

‘Unlock your phone,’ I demand as thrust the phone at Jefferson. He complies and hands me back the unlocked phone. I look through his pictures, praying that there aren’t incriminating photos of any kind. There are no pictures of the test. I then check to see which apps are open making sure he wasn’t using an app that does math problems from pictures the user takes.

Jefferson’s phone is clean. I hand it back to him.

‘Now I have to call mom, and you have a day of in-school-suspension,’ I exasperate. I will also have to notify the assessment and accountability office. ‘Next time just give the teacher your phone. You can live without it for a couple of hours. The teacher has her own phone, and she doesn’t want yours.’
This story exemplifies how I interact with high-stakes assessments in my day-to-day job during testing season. It is stories like these I wished to expose along my journey. I have dealt with scenarios similar to this many times. Let me unpack this narrative. First, cell phones are prohibited during testing for the obvious reason of preventing cheating. Second, Jefferson having his phone out during testing is considered a behavior problem rather than an academic problem, which warrants a punishment. I think the state and district systems consider it to be both since Jefferson received an academic punishment, the test invalidation, as well as a behavior punishment, in-school suspension. Third, Jefferson’s offense has entitled me to violate his privacy by looking through his phone. I knew Jefferson would not have any pictures of the test or the app on his phone. He did not turn in his phone because it is his lifeline. The phone is an extension of his life and represents his social world. Phones are expensive and not easily replaced by working-class families. In fact, I have yet to find any cheating going on when I have dealt with similar situations. I would have preferred not to look through his phone, but I carried out my responsibility because it is part of my job. The mere fact neither Jefferson or I question why I need to look in his phone, why he should be punished, or why his test is invalidated demonstrates the importance of high-stakes assessments, the maintenance of a secure testing environment in today’s public schools, and the over-reach of testing companies into the lives of students and educators.

Securing testing includes collecting electronic devices from students as well as requiring students to sign an acknowledgment, which states:

During this test, you must not:
- talk to other students or make any disturbance
- look at another student’s test and answer book or planning sheet
- allow another student to look at your test and answer book or planning sheet
- ask for help writing your response
- give help to another student in writing his or her response
- have notes or scratch paper other than your planning sheet
- have any electronic or recording devices in your possession at any time, including breaks, even if you do not use them
- fail to follow any other instructions given

Because the content in all statewide assessments is secure, you may not discuss or reveal details about the writing prompt or passages after the test. This includes any type of electronic communication, such as texting, emailing, or posting online, for example, on Facebook, Twitter, Snapchat, or Instagram. Once this portion of the script is read aloud, students are asked to sign below a Testing Rules Acknowledgment (or check a box for some CBTs) that states, I understand the testing rules that were just read to me. If I do not follow these rules, my test score may be invalidated.  (FLDEO, 2018a, p.37)

Thus, students, minors, who are unable to consent to anything without a parent or guardian present in any other situation in their lives, are required to sign or check a box stating they understand the importance of test security at school. This acknowledgment is more than just a warning. For example, Pearson has monitored students’ social media accounts and has even gone as far to contact state departments of education asking schools to discipline a student when Pearson found the student posted about a Pearson high-stakes assessment on social media (Singer, N., 2015). I question the ethics of state departments of education agreeing to allow testing companies to surveil students in the name of securing intellectual property. Moreover, Pearson sought to have teachers sign security agreements in Massachusetts, which states teachers could lose their jobs or certifications if they breached security (Singer, C., 2015). School-based administrators, at the policy implementation level, are expected to discipline teachers and students who violate agreements made between states and testing companies.

I hate this aspect of my job. If I was not terrified I would be fired, I would refuse to participate. But, there is always this fear in the back of my mind that, if this time, I did not look in the student’s phone it would be the one student who did take pictures of the secure assessment and put it on every type of social media. So I comply with the directive of the assessment and accountability office. I do not agree with this policy, but I am resigned to carry it out.
CHAPTER 5: CHER’S TALE: EMBRACE THE SUCK

My phrase this year was ‘embrace the suck.’ ‘Embrace the suck,’ to teachers. Absolutely, ‘embrace the suck, this is going to suck.’ Nobody likes it. I don’t like it. The kids don’t like it. You don’t like it because I’m like, ‘we don’t really like it.’ My [subject area leader for language arts] she got these little lollipops and she got little, ‘embrace the suck’ on it and gave it to all the teachers. Not the kids but the teachers because that’s just what an emotional roller coaster it is because everybody—their emotions are at all levels you know because nobody likes it. The people who have to proctor don’t like it. The people who have to hold the kids not testing don’t like it. We don’t like it. It’s all interrupted. That's why I said it embrace it you know you don't like it so let's embrace it and put a smile on our faces. That's just moving forward and every year what I just did. I said, ‘embrace the suck’ in other words I said, you know what I was saying is, ‘I think this sucks I think there’s nothing right with this. But I am going to try to make it so it’s livable. So we are all on the same page. We are all making it.

**********

My next stop has brought me to Cher. Cher has mentioned to me over the years the amount of testing we do in schools is ridiculous. Thus, I asked her to meet to share her stories and experiences with me in hopes of gaining some enlightenment about the ethical dimensions of implementing high-stakes assessment for school administrators. “You look fantastic!” I exclaim when I see Cher. She has just returned from a trip to see her husband in Florida’s panhandle. Her blonde hair is even blonder than usual and her olive skin is golden from the sun. She gives me a bright smile and thanks me. We are meeting at a Greek restaurant on a quiet weeknight. I
have known Cher for about eleven years. She is a senior assistant principal at a middle school in the same school district as me. I was enrolled as a new teacher in a professional development course she was conducting when we met. Then five years later, when I was a reading coach, Cher was the district reading resource teacher assigned to be my point person to work with my teachers and I in the use of the district-adopted literacy curriculum. Over time we collaborated on several projects for the school district together and became good friends.

As soon as we order our food I ask her about the importance of students’ high-stakes assessment scores and its relationship to her job as a senior assistant principal. Cher is thoughtful as she contemplates her response. However, her passion and concern for her students is apparent when she begins to speak about using high-stakes assessment scores to place students in courses.

**Culture of Measurement**

*Anything I've ever read about placement of kids in courses you never use one measure to place a child. That doesn't show the whole picture okay? It could be it was a bad day, or 'I just don't test well in this environment,' or 'I don't do well on a computer test, or the questions are biased, or the paper pencil is better for me.' We use one measure to place these kids in and level them in their courses."

*Did someone actually tell you to only use one measure?, I inquire.*

*Because of the position and the school that I was in (as a district-level reading resource teacher) I was to make sure that kids were placed in the correct level of reading classes based on FSA and FSA only. There were no other measures even considered. Now we make them do FAIR who knows what we even use that data for? I don't know of anyone who uses that data. We don't use our data from PSAT in middle school to for placement. Supposedly, our eighth*
graders, the guidance counselors in the high schools are supposed to be looking at that and it's supposed to be considering placement especially advanced placement for those kids based on PSAT that's what the district told us. Today, for instance, I was handed, by the district, core courses. Just the four cores (math, English, science, social studies) and the criteria for placement in those courses was FSA only. That's the way the Automatic Course Requests (ACRs) were set up, and those buttons that you clicked and those courses, boom, and are now in those student schedules for next year based on FSA. We [ran ACRs] today and we will again run it again in June or July. We are mandated to [use ACRs]. As much research as we've read we know is that you never measure by one single measure we never measure achievement by one single measure it's a the body of work, Cher surmises.

Cher is pointing to a body of research (Amrein & Berliner, 2002; Gulek, 2003; Heubert & Hauser, 1999; Hoffman, Assaf, & Paris, 2001; Kohn, 2000; Oakes, 2008; Townsend, 2002) suggesting tracking or programming students into classes based solely on high-stakes assessments is detrimental to students, particularly students of color and students living in poverty. Other data points, such as teacher anecdotal data, school, and district-created assessments are ignored, and the professional opinions of teachers about the placement of their students are not even taken into consideration. Tracking itself is not an effective instructional practice (Marzano, Pickering, & Pollock, 2001). The pervasiveness of regarding students as the same as their test scores and using the scores as the only factor in student scheduling is harmful to students (Amrein & Berliner, 2002). Indeed, this practice suggests students are nothing more than what their individual single test score says they are.

As I listened to Cher’s story, I wondered what if she did what she thought was right instead of following directions from district-level administration? I doubt she would get in
trouble. I doubt that anyone would even know if she placed students in courses using more than one measure. How would anyone know? I know some senior assistant principals do not follow every district edict. I need to think about how I will handle this situation when I am a senior assistant principal. Cher may be frustrated with how students are tracked in our school district, but she seems resigned to carry out the directive.

*It is amazing to me as to how middle school students are just compliant. We just go, it's our time to test, we got our little colored wristband on, we go we sit, and even the worst of the worst kid sits there and doesn't say anything. He probably falls asleep and every now and then you've got to wake them up. That's the worst thing that happens you've got to wake them up and you ask them, 'why are you testing? why are you taking this test?' and they say, 'well, because next year I don't want to take intensive reading. I really I don't want to take intensive math, and I don't want to be in the boom-boom class. I don't want to be in a regular class.' There's a stigma attached to these tests that kids are very very very aware of. So that's why in middle school I believe their complaint. Think about this, the outcome for those kids, level ones and level twos, puts them where they lose electives puts them in remediation. What about those kids that are fours and fives? In our school district do we have things that are enrichment that really pushes those kids who are threes, fours, and fives?*

Cher raises a question about the narrowing of curriculum for both students who are high and low achieving on high-stakes assessments due to the hyper-focus on remediating. The students who are low achieving on high-stakes assessments are placed in remedial classes focused on basic skills. They are often excluded from elective courses, such as the arts, which research suggests might actually improve academics (The College Board, 2008; The Dana Foundation, 2008; Meyer, 2005; Smith, 2009; Winner & Hetland, 2008). Researchers have also
questioned what exactly schools do to enrich students who perform well on high-stakes assessments when the focus in the classroom heavily emphasizes test preparation (Brimijoin, 2005; Moon, Brighton, & Callahan, 2003; Pandina Scot, Callahan, & Urquhart, 2008). Standardized curriculum assumes all students are the same and need the same thing to learn (Au, 2009). However, Cher recognizes not only do the students who perform poorly on standardized assessments suffer from the narrowing of the curriculum but also the narrowing of the curriculum due to high-stakes assessments negatively impacts the students who perform well. Cher further suggests the ethical implications of the narrowing of the curriculum due to the sheer amount of time lost during testing.

Most students I have worked with are compliant when it comes to taking high-stakes assessments. Most adults are compliant, too. Educators and parents alike comply with the requirements of high-stakes assessments. I think high-stakes assessments negatively impact all students, but the students whom I see suffering the most are the students who struggle with reading, writing, and math. These students perform poorly on high-stakes assessments, and they get more high-stakes assessments, more assessments in general, less access to arts-based curriculum, and more test-preparation than any other group of students.

As I mentioned in chapter four, the deficit language we use when discussing our students is part of our everyday way of speaking. Cher refers to groups of students as the number the students received on their FSA ELA assessments, Ones, Twos, Threes, Fours, or Fives. Cher is not the only one. Most educators I work with talk about groups of students in the same way, and I admit I am also guilty. The students do not have an identity beyond their achievement score. We are so unaware of our socialization into our school district’s way of being we do not even realize we are doing it.
In middle school it's, and honestly just from our conversations has made me really be more intuitive about what is really is going on around me. At the middle school-level it's more of a compliance thing. Kids do it because that's what they think is expected of them, and that's what we're cheerleading them to do and because high school is their goal. They see this testing as a step toward being in the classes that they want to be in high school. That's the dangling carrot that they see. I believe that's why they're so compliant they just do it. Matter of fact, when I did make-ups I would have kids coming to me saying, 'oh my god, when are my make-ups? When is reading? When am I making it up? I need it,' and this is like every kid who missed a day. We have a high absentee rating, we have a lot of make ups to do but the kids are just so compliant and so focused because the goal is high school and the classes that they want to be in in high school and they see the FSA as their be-all end-all, Cher explains.

Where do they get that notion from? I wonder.

Their teachers and their parents I'm quite sure. Like, 'you don't want to be in boom-boom classes so you got to do your best on the FSA or you're just gonna be placed in regular classes if you don't do your best,' and so they're like, 'oh no okay that's not me.'

Do you think that's accurate? I probe.

No not really, you know it's the bottom 19th percentile and most of these kids don't fit into that profile at all that are trying to make these tests. They're the high level two, high level threes, most of our kids are. I mean they're not superstars, my kids, none of them. I have very few level fours and fives, but they just have this desire to want to get to high school. It's all about high school, and what they're going to be able to do there.

The compliance Cher speaks of is the “performance agency” discussed by Black (2008a), whereby the students take on the responsibility to perform well on high-stakes assessments.
simply because teachers, parents, and school administrators expect them to. According to Cher, her middle school students have taken on the performance agency because they want to avoid experiencing negative consequences or be socially stigmatized by being in a remedial class. Black’s (2008a) research took place in an elementary school. In contrast, Cher’s description of middle school students invested in performing well enough on high-stakes assessments to avoid negative consequences suggests older students feel a greater level of pressure in a culture of measurement.

I wonder what makes a student a “superstar”? Above average performance on a high-stakes assessment? This is deficit thinking, which I did not recognize when Cher was telling the story, but it jumps out off the page at me now. It makes me sad and angry with myself for not asking what makes a student a superstar, for not even noticing, because nothing will ever get better if we do not interrogate each other about how we refer to students in our daily lives.

Again, it points to my socialization in the school system.

I want to add to the point about FSA and an EOC, Cher continues. In middle school this year what's happening with our algebra students is that we're not going to get the scores back in time to know if they passed the test or not, right? The expectation of these families, and I have a huge problem with this. This is the part, I wanted to tell you today is that okay?

I have a list of 25 kids who had a D quarter 4 in Algebra 1.

They're all invited to the Algebra EOC camp.

They are to sign up.

They are to appear there.

They're there to start the camp.
Then when the scores come in the assistant principal at the camp site will go up to them, ‘you know what you passed! Congratulations to you!!!’”

It cheats families and it cheats kids. It’s ridiculous. They gave students in eighth grade the option now you can go and sit for the camp and then July 11 retake it or you can just take this form and just do a walk in retake on July 11.

In Cher’s story, the students’ outcomes on high-stakes assessments supersede everything, including summer vacations. Again, Cher seems resigned to witness and even assist with a practice with which she disagrees. Cher is describing a common practice: asking students to take part in test remediation and preparation outside of school hours. In this case, students have been asked to take part in summer school based on in-class performance data or class grades without having the actual data from the spring EOC. Once the scores are reported back to the schools by the state, if the students in the summer course failed the spring administration of the Algebra EOC, they retake the assessment in July. If the student ends up passing the EOC, then they just sat through weeks of test preparation during the summer for no reason. Cher also struggled with what exactly did we do with the test results for students other than the ones who score most poorly.

Think about this, instruction isn't even over when we start testing. We'll still have another four weeks of instruction but we are done because of testing, Cher declared.

Testing in Florida begins in February but the most intense amount of testing occurs in April when there are still many weeks of school left. Since most middle and high school high-stakes assessments are administered on a computer, which there are a limited number of, students take these tests in shifts. Thus, one-third or even half of a class is missing instruction at any given time during this month of testing. Further, students are being tested on how much they
learned but they have not even completed the course yet. For example, students will take the FSA Writing assessment the first week of March but still have two months of English instruction after the assessment. Nicholas and Berliner (2007) and Au (2009) describe the narrowing of curriculum as focusing solely on tested subjects, but what about the loss of instruction students miss because they are testing before the end of a school year and discontinuing instruction shortens instructional time. Smith (2000) suggests the school year is essentially over once students take high-stakes assessments. In Florida, this means four to eight weeks of instruction is lost to high-stakes assessments.

_I was thinking about funding, and I think about FTE\(^2\) to make sure that we're on point for our federal funding. This is about… it's all about the dollars that flow from the federal government to the state government who hands it to us as they see fit based on these tests scores and our proficiencies and our accountability. These are business people. I mean think the governor_, Cher passionately explains, her voice rising with frustration.

Cher is referring to the current governor of Florida, Rick Scott. Scott was head of a Columbia Hospital Group prior to his political career (Flgov, 2018). Scott is not the first politician to have a business background, nor is the adoption of business practices in public sector new. The effort to push business practices into public education in the United States has been around since the turn of the 20\(^{th}\) century. Throughout the last 100 years calls to run schools like business have become common, with large school districts, such as New York and Seattle hiring business people to run their large districts in recent years (Howard & Preisman, 2007). Cher is connecting the fascination with business practices in schools and governments with the ever-increasing accountability in schools, which are tied to neo-liberal policies of globalization.

---

\(^2\) FTE (Full-time equivalent) refers to K-12 public education finance. Several times a school year, districts in FL must demonstrate they are in compliance with the FTE rules and regulations, which include how many teaching units each school receives.
Currently, policies put in place by those at the state and federal level emphasize a “human capital world model” approach to standards-based instruction and high-stakes assessments to determine whether or not students, teachers, schools, and districts have met those standards, which are then tied to systems of rewards and punishments and compared to other schools systems around the world to encourage global competition (Spring, 2009, p. 17). According to Spring (2009), the human capital world model is characterized not only by standards-based instruction and high-stakes assessments for determining student, teacher, school, and district worth it also includes scripted lessons and mandated textbooks. Additionally, the human capital world model’s objective for education is measured by economic progress and expansion and to educate a workforce to participate in a global economy whereby the workforce passively accepts the “political and economic structures that operate against their own interests” (Spring, 2009, p.17-18). Thus, Cher’s story reveals the ways in which she witnesses the human capital world model intersect with her role as a school administrator.

Testing Season

Cher takes a long sip of her drink and goes on: The amount of time it takes to plan and organize and all of the implications that you have to think about. For computer-based testing (CBT) you have to have proctors, and you have to pick which proctors are going to be successful.

You have to train those proctors.

They have to go through the FSA [training], and they have to take their little course and get their certificate. That takes time...

and then we have to do a chart based on how many labs I have.
Like little squares so that morning and afternoon sessions, I plug their names in and what room numbers it is.

Then I have to think about how many computers do I have available? My extended time kids, I have to schedule them for the morning session so I have to start with them see how many I have, then I see how many labs I have left over.

Then I can schedule the kids without accommodations, well those are usually in the afternoons and, then you do you have to do grade levels, but then you know you've got to see how long the first test goes. If that's 60 minutes and the next one is 120 minutes well you have to make sure that those are two days in a row.

Then what I do is an implications page, which lists the test administrator and what class periods which teachers are covering on their conference period for the teacher who is administering the test. My principal does not allow dispersals so there has to be teacher coverages, and think about that I mean that's comp time for them.... he feels that [the kids] are not responsible enough to get where they're supposed to be. It's a tough school.

Cher’s job during testing season is to coordinate the people, places, and times of assessment in her school. She is responsible for smooth running of the actual testing taking place in classrooms and the adjustments other non-testing teachers and students to make space for those who are taking the test. Orchestrating the deployment of resources during testing season is stressful; indeed Cher is the project manager of testing season. She is in charge or allocating and reallocating resources throughout the season, which typically lasts about two months. As Cher spoke, her resignation to carrying out the tasks of testing season became apparent.
This week I had to run three sessions a day because of the number of computers. We had to be on it, and we had to be moving at every single free minute so the lunch schedules were screwed up for the kids, too, because the ones they were testing in the middle [session]. I had to have all of those take first lunch, and then I had to move lots of people—it's just the constant movement—it was constant movement all this week...my testing coordinator, he is doing the other nuts and bolts. I didn't rely on the teachers [to make sure students went to the correct testing location and time]. I did three different color wristbands. I did blue for the first session in the morning, yellow for the middle, and then red late for late afternoon....I would say on the intercom anyone with the blue wristband move to your testing location, and we wrote the room number on that the wristband.

Providing students with brightly colored wristbands, as though students were heading into a fair, rather than heading into a computer-based high-stakes assessment, allowed Cher, and those assisting her with the management aspects of the tests, to maintain control over all the moving people. The pressure to execute high-stakes assessment perfectly led Cher to utilize a Neo-Taylor-like system of efficacy and centralize her authority during testing season because she did not trust teachers to carry out her plan (Au, 2011).

Cher continues: It’s crazy. I mean this amount of time that I spent putting together the plan of the day, the plan of the week with who is covering what classes, and who is going to be in which labs being the proctors of the test administrators, and which kids and which armbands are going to go to which kids. What room they're going to go to and to ensure homeroom teachers actually care enough to give them the wristbands in the first place, change the lunch schedule, change the bells schedule, turn the bells off, turn the bells on. The amount of time I spent on just
the organization of the whole things. It's mind-boggling how much time was spent. I can't even start to tell you like nothing else anywhere else exists. Nothing else gets done.

The amount of time spent planning and carrying out the administration of high-stakes assessments during testing season was a common refrain in Cher’s narrative. In fact, Cher told me the story of her catchphrase during testing season:

My phrase this year was ‘embrace the suck.’ ‘Embrace the suck,’ to teachers. Absolutely, ‘embrace the suck, this is going to suck.’ Nobody likes it. I don’t like it. The kids don’t like it. You don't like it because I'm like, ‘we don't really like it.’ My SAL for ELA she got these little lollipops and she got little, ‘embrace the suck’ on it and gave it to all the teachers. Not the kids but the teachers because that's just what an emotional roller coaster it is because everybody— their emotions are at all levels you know because nobody likes it. The people who have to proctor don’t like it. The people who have to hold the kids not testing don't like it. We don't like it. It's all interrupted. That's why I said it embrace it you know you don't like it so let's embrace it and put a smile on our faces. That's just moving forward and every year what I just did. I said, ‘embrace the suck’ in other words I said, you know what I was saying is, ‘I think this sucks I think there's nothing right with this. But I am going to try to make it so it’s livable. So we are all on the same page. We are all making it.’

“Embrace the suck,” suggests a twist on the performance expected of the teachers and the administrators during testing season. The adults are going to pretend to enjoy high-stakes assessments by putting smiles on their faces in front of the students, but when the students are not around the adults put away the pretense and share a common dislike for testing season. These performances are reminiscent of Smith’s (2004) “political spectacle,” whereby humans become actors in roles created by policies. In this case, Cher and the educators she works with
have become actors in the roles created by high-stakes assessment policies. This is a form of emotional labor, which Cher accepts with little question, which highlighted by the use of the catchphrase.

I have been an actor, performing the role of high-stakes assessment administrator. When I proctored high-stakes assessments, I always smiled at the students as I read the script. I encouraged students to do their best on the tests. I would hand out peppermints to help keep students alert. The closest I would come to revealing my true feelings to students was if students complained, I would calmly suggest they vote for state officials that would get rid of high-stakes assessments when they were 18, when it did not impact them anymore, and when they would likely no longer care.

Cher’s narrative includes how she performs the close of the school day during testing season:

_During testing season, every single afternoon I did the afternoon announcements and the first thing I led in with every single day is, ‘Thank you, students. Thank you so much for your hard work and your dedication today. I noticed how hard you're working. I appreciate everything you're doing.’ I said that every single day in the afternoon before they went home and I meant that because I mean bottom line these kids are doing their best._

As Hanley (2015) suggests, an “emotionally skilled educational leader” will cloak the negative aspects of neo-liberal accountability policies in positive language. Cher’s description of her end of the day announcements exemplifies how she attempts to put a positive spin on high-stakes assessments even as she feels negatively toward testing season. Iszatt-White (2009) suggests there are performance norms embedded into the job of the school leader, norms
governing how a leader ought to act in given situations. Thus, the beliefs steeped in neoliberal policies may cause Cher to perform contrary to personal beliefs.

Cher then tells me how testing season impacts her personally: *I know I've been staying late because we get so done so late. I try to get things done in the afternoon like I said I've had three sessions a day. I try to get things done in the afternoon for the next day. I have been there all week I have not left before 6:30 p.m. sometimes it's closer to 7:00. I did work last Saturday trying to put together the civics and the science because I just put together the FSA. The big ones, you know? The ELA and the math. Sent that out and then I'm like, 'oh my gosh I feel like a brand new teacher like staying a week ahead of the kids’*

*Do you do this all yourself or do you have help?* I wonder.

*I have a guidance counselor. We divide and conquer. She does the rosters and the test tickets, and then I do everything else; the implications, I scheduled them. I do all the math.*

*How many guidance counselors do you have?* I ask.

*Two.*

*When she is doing testing who is working with the kids?*

*Just the other guidance counselor and the Student Success coach but most of the time he's testing. I have him testing as test proctor.*

*There's a question about ethics are we taking away other resources not just instructional time but other resources from our students?* I suggest.

*Academic coaches are testing, my media specialist is testing. My media center is closed down for three weeks. It's stressful it's just the stress level of everybody's just elevated this week.*

I completely missed the personal stress placed upon Cher to plan, organize, and carry out all aspects of high-stakes assessment during testing season during this exchange. I focused on
the unfortunate diversion of human resources away from students, but Cher’s story is about more than just resources, it is about a K-12 public school administrator dealing with the stress and the ethical tensions associated with implementing high-stakes assessment. It is about a human being who cares deeply for others. Yet, as Au (2009) describes in systems of high-stakes standardized assessments students are widgets or data points rather than humans likewise the educators, like Cher are merely line workers who collect data points from widgets. Cher bears the brunt of the stress for her school, and I wonder if the reason she has not suffered from burnout is because she has resigned herself to the situation.

Furthermore, the reallocation of human resources away from their intended purpose also stands out as an ethical dilemma Cher faces during testing season. She must pull human resources away from students during testing season and close the media center for three weeks to accomplish the completion of high-stakes assessment within the given time frame.

**Ethics of Implementation**

As we finish up our dinner, the conversation turns to testing season. Since we are in the height of testing season Cher’s frustration is readily apparent as she conveys her story “our intellectually disabled (IND) exceptional student education (ESE) unit grew from fifteen to thirty kids this year with no additional teachers and no additional paraprofessionals (paras), and these Florida Standards Alternate Assessment (FSAA) tests are one-on-one. So the teacher has to take, and the para can't do it, the teacher has to take students one-by-one and scheduled them, no more than one or two per day, and still has to teach and keep the content going. Well, and keep chaos from happening while they're not there. I just feel for the kids who were in the classroom while the teacher is gone doing the one-on-one testing. [My School is] all access points. My IND unit there it's 100% access points….kids with severe disabilities like Down
Syndrome that are non-nonverbal and things like that....I have a Down Syndrome boy on my campus, nonverbal, that will be given the FSAA.

Since the passing of NCLB there have been opponents and proponents of the implementation and requirement of high-stakes assessment for all students with disabilities (Cotto, 2016). Proponents maintain without accountability then teachers and school systems would not provide students with disabilities equal access to education (Katsiyannis, Zhang, Ryan, & Jones, 2007). However, opponents believe testing students with disabilities is cruel and unusual punishment (Smyth, 2008). The contradictions of providing students with Individual Education Plans while at the same time requiring the same students to perform on high-stakes, standardized assessments are not lost. In Florida, the alternative assessment, FSAA, is given to students with significant cognitive disabilities, which is modified for students based on their individual education plans (K-20 Education Code FL § 1008, 1995). Cher is describing the reality of this policy at the implementation level: the classroom teacher cannot be in two places at the same time. Thus, with limited resources school leaders must decide to ask teachers of students with cognitive disabilities to leave their classes to be managed by paraprofessionals while the teacher assesses each student one at a time on the FSAA. She questions the ethics of this policy. Each student requires and deserves at trained professional to administer the FSAA to them in a one-on-one setting, but at the same time, the other students assigned to the class deserve a professional teacher teaching rather than a paraprofessional babysitting. The ethics of implementing FSAA was something Cher struggled with, and she wondered if all students, no matter their ability should be subjected to an assessment.

Yet, I question whether high-stakes assessment is the true cause of diluting the resources meant for students with disabilities. It is easy to remove or reduce human or capital resources
from a group of students who are vulnerable and may be unable to fully understand or explain what is going on in their classroom. The most vulnerable of students are the easiest to ignore or take away services without being challenged. This is an ethical dilemma about how school administrators choose to use resources during high-stakes assessment.

As we contemplate dessert, Cher considers the ethically questionable situations she has faced in her career as an educator in relationship to implementing high-stakes assessments:

*When I was in a different position and right when the scales changed my number one job was to go to all the middle schools that I was to be working with and to sit down with the administrators and print out how all the kids that were just about to move from a 2 to 3 or maybe wasn't 3 right now but would have the potential to move. I had to help them identify these kids. And it really it wasn't about the kid, right? It wasn't about educating that kid. It was about making sure the adults knew who those kids were so they could move them, Cher explained.*

The practice Cher is describing is often referred to as focusing on the “bubble kids” (Booher-Jennings, 2005) or “educational triage” (Gillborn & Youdell, 1998). Gillborn and Youdell (1998) posit educational triage is like battlefield triage where those who can be helped are helped, and those with no hope are left to die. Likewise, Vasquez-Helig et al. (2010) describe this phenomenon as school being “at-risk adverse” whereby students who are not close to performing well on high-stakes assessments are not provided the instruction they need and are often encouraged to drop out of school (p. 562). In the case of education, students are often put into one of three groups: those who are already performing well on high-stakes assessments and have demonstrated they can make the cut, those who are close to performing well on high-stakes assessments and with help can make the cut, and those who are performing poorly and have no hope of making the cut (Gillborn & Youdell, 1998). The bubble kids are the students who are...
close to performing well on the high-stakes assessments and with help can make the cut. In her former position as a district reading resource teacher, one of Cher’s responsibilities was to help schools identify the bubble kids. In this model, the neediest students are always getting remediation, rarely experience advanced curriculum, and are perpetually left behind. Cher’s story implies the practice of focusing on bubble kids is ethically reprehensible, but nevertheless, a common practice in public education.

Another ethical concern for Cher was her English language learners (ELLs). Similar to ESE students, ELL students are partially tested in a one-on-one setting. Cher tells me the story of administering the WIDA, the English proficiency assessment, to her ELL students this year:

*We started with WIDA, and we had 60 kids who tested with WIDA. Well, my first obstacle was that the teacher who teaches the DLA classes, she left February 1st. We started testing the next week, but that was kind of okay because I have this fantastic para that was going to teach her how to do it. Anyway but it was just it made us short one person, and we got a substitute to take her place but the substitute didn't speak Spanish. I have 30 monolinguals so she couldn't help with any of this, and she really wasn't helping in the classroom either so it really put a burden this poor para. Our two paras, plus our guidance counselors had to do all of the WIDA testing so those monolinguals weren't getting serviced in any of their classes because of the testing that took almost a month. I finally did get a Spanish speaking substitute in there about three weeks in and that put them behind. Testing really put these kids behind learning what they needed to learn in class, which is a problem they should have been getting the services they deserve to get, but because I was left with no other choice than to use my paras for testing they fell way behind on their lessons but they did get it finished. The poor kids, you know? It was one-on-one [testing] and then it was small group and they lost a little of their patience...*
because of both ends. They were just testing, and they really weren't learning what they were being tested on and it just wasn't fair to them. I felt it wasn't fair to them. They felt it wasn't fair and the whole thing wasn't fair simply because that teacher moved away and left her job in the beginning of February. It did take about a month but we got them all done, Cher pauses to order a glass of water.

The allocation of human resources away from instruction and into administering high-impact assessments, such as WIDA, speaks to the value placed on gathering data over instruction. Cher describes using human capital in ways not intended: the paraprofessional was never intended to be the primary teacher of this class of students, nor were the paraprofessional and the guidance counselor ever intended to be the administrators of the WIDA. Indeed, the paraprofessionals were hired to support students during instruction. The guidance counselor was intended to guide students academically. Yet, these resources were diverted. Cher believed if she did not allocate her resources this way then the WIDA testing would not be completed by the deadline set by the state. This is an example of distributive justice. Perhaps Cher had more choices and agency then she realized because Cher shared another story about WIDA testing with me where she does use her agency:

After we finished WIDA my principal at my school thought it's a good idea to go ahead and test these students on the FSA even though that they're not required to test. After they've done all of this and gone through everything they've gone through, he says, ‘we can get results on them faster if we get a baseline this year, and not wait till next year. That way they'll show improvement because they're going to learn English better and that will help our school grade.’ I have a real problem with that because they don't know enough English to do well. They're already frustrated and we're going to frustrate these kids even more. I had an ethical dilemma
with that. I don't feel it's fair to them and I told him as much. He said, 'well let me think on it,' and so now three weeks have gone by and he hasn't mentioned it again. I did not because I have a problem with it. I did not schedule them to test, and I won't schedule them to test because it's not fair. He didn't say anything else about it so I'm kind of guessing that he let it go because I let it go. Oh and by the way I think he kind of let it go because it slipped his mind because a week ago he broke his hand breaking up a boy fight. He's had to have surgery this week so he hasn't been bothering me too much. I'm sorry he broke his hand but I'm not sorry that he forgot about his idea of testing these ELL kids to death. I guess it's not going to happen but was a dilemma for me. I just kind of swept it under the rug and didn’t make it happen because I didn't believe it. It's just that's not fair to them, it's just not. I don't know if the FAIR test counts in what you're trying to write about, but that stupid progress monitoring test I really got a problem with that just it's just too much. Especially the third time they take the FAIR test in the spring because they're about to take the FSA. It should be their final measure of their improvement over the year. We're doing FSA now they shouldn't have to keep doing all of these progress monitoring measures but I get it you know it has to do with money and it has to do with state funding, Cher recounted.

ELL students who have been in the country for less than a year are not required to take the FSA ELA assessment, but they are required to take WIDA (FLDOE, 2018a). In fact, if first-year ELLs take FSA ELA their scores are not factored into the school’s grade. Districts and schools can determine whether or not they want their first-year ELLs to take the FSA ELA. Some districts and schools choose to have first year ELLs sit for FSA ELA to establish a baseline. Then, when the ELL students take the FSA the next year, they can potentially show gains, which do factor into school grade. Although there are some testing accommodations
provided to ELLs, such as directions read in heritage language and being allowed to use English to heritage language dictionaries, this test can be daunting for students brand new to the United States (Honigsfeld & Giouroukakis, 2011). Cher finds this practice unethical and unfair strongly enough that she subverts the authority of her principal and refuses to force her ELL students to sit for both tests.

**Conclusion**

Cher’s narrative reveals the contradictions assistant principals face when implementing high-stakes assessment. The story of stress during testing season and the culture of measurement as part of her daily experience provide a unique perspective on my journey to illuminate what it is like to be the instrument of federal, state, and local policies at the implementation level. Cher questions the ethics of implementation while she also motivated to find space for resistance within her role. Cher’s awareness of the politics creating the conditions she must work in every day is a unique perspective on high-stakes assessment policy and implementation. It is the resistance I am most inspired by, this is what I hoped to find on my journey: spaces of resistance. As I move to my next stop, I will continue to explore testing season, the ethics of implementation, and the culture of measurement.

**Vignette**

*So Joe is taking science, Adam is taking social studies, Will is taking English, Jenn is taking reading, Lauren is taking electives, and I am taking math.’ It is late July, and my principal is trying to confirm which administrator will oversee which department for the school year. ‘Speak now or forever hold your peace.’*

*I squirm in my chair. ‘God, I do not want Reading again,’ I think. I have a master’s degree in reading education and was formally a reading coach. Therefore, I have been the*
administrative liaison to the reading department for the last three years. I do not want to do it again, but I also do not want to be difficult. I have a new principal, and I want to please her. Yet, the idea of spending another year working with the reading department mentally exhausts me. The teachers are always new and inexperienced in reading because the good ones are typically siphoned off after a year or two to teach English. Teaching high school reading is not glamorous. Students who struggle with reading and writing and do not like either populate the courses. New teachers often accept reading positions because reading positions are more plentiful. Reading courses exist purely to help students pass the FSA ELA high-stakes assessment. Further, the personalities of our teachers in the reading department are strong and the teachers often fight amongst themselves. Plus, if I plan to be a senior assistant principal, I will need to learn about other contents. I consider the alternatives, ‘English, nope, I may as well stay in reading. Social studies, nope those teachers’ personalities are too difficult. Math, um, no, my principal wants math so that is not a possibility. No way I want electives. Science! Yes, science.’

‘Can I switch to science?’ I finally speak.

‘Joe, are you okay with switching with Jenn and taking reading?’

‘Sure,’ Joe agrees good-naturedly.

The science department gets along famously at my high school. Science has only one high-stakes assessment, the biology EOC. The teachers are mostly experienced in our science department, and they have been working together in professional learning communities for years.

My decision to move departments was for several different reasons, as my story explains, but upon reflection, I see I took the path of least resistance. I have absolutely zero experience in science, and if I view myself as a school resource, I have just misused myself. I should be the
liaison to reading because my expertise is in reading. Reading has the most pressure and the, in my past experience, least experienced teachers. My guidance could have been beneficial. Yet, in the culture of measurement school administrators often make decisions based on what is best for them, rather than what is best for students, teachers, or the school. I must admit, I made the choice to supervise the science department because it was easiest for me. Do I think I was a helpful resource to science, too? Yes, I think my skill set is transferable to all content areas. I cannot help but wonder if I would make the same choice if I did not work in a state and nation with educational policies, which encourage a culture of measurement.

I could not bear witnessing the endless agony and struggle of reading students and teachers as they beat themselves over the head with the same rock over and over again. The teachers teaching the same useless reading strategies designed, not to help students learn to read, but to help the students pass the test. By choosing not to be part of this sick and twisted game, I did nothing to eliminate the problem. The problem still exists, and I have done nothing to change anything. I simply walked away. I left the reading department in good hands because the administrator who oversees reading now is a caring educator. But, what if he was not caring? I did not choose my successor. A school administrator with no awareness of the struggles students and teachers in the reading department contend with could have been assigned to the department. When those of us with social justice orientations leave difficult situations in public schools, we battle the ethical tensions associated with having to choose between what is best for us versus what is best for students.

According to Padilla (2005) in the culture of measurement humans are abstract objects whose performance on high-stakes assessments are the only true measure of their learning, teaching, or leading. I was certainly aware my perceived effectiveness would be viewed through
the success of my department’s student performance on high-stakes assessments when I asked to oversee science. FSA ELA assesses reading and writing for all 9th and 10th graders as well as 11th and 12th graders who are retaking the 10th grade FSA ELA because they have not passed. All students must pass the FSA ELA to graduate versus the biology EOC, which assesses students at the end of the school year they take biology. Students in Florida do not have to pass the biology EOC they just have to take it. Most students in my school district attempt to perform well on the biology EOC because it is calculated as a final exam score in the student’s final grade, but, clearly, the stakes are much lower for the administrative liaison to science than reading. Mintrop (2012) found managing the forces of accountability policies and keeping the best interest of students and personal integrity at the forefront of practice is a difficult balance for school administrators. I would not characterize my decision as unethical; however, if I am being honest with myself my choice was undoubtedly self-serving.
CHAPTER 6: MICHELLE’S TALE: IT’S A FISCAL TESTING YEAR

Schools game the test at the disadvantage of students because we have limited resources and so how do we spend those limited resources you want to spend your resources in a way that gets you the most value for the resources you're expending. So that's just business 101 and so when we look at students we are going to pull out or we’re offering tutoring services to who do we go to? We go to the students that are going to you get us the most bang for the effort that we're putting into those kids. We immediately go to –

The bubble kids, I suggest.

Right we go to those kids that are about right on the edge of moving to another bucket, and that's where we focus all of our energy, and it's at the disadvantage of the other kids that are following falling in the middle of these buckets. So again by doing that are we gaming the system? Are we working to the advantage of you know school grade? Absolutely, we are, and it's at the disadvantage of kids because at some point when you're making those decisions you are choosing not to provide certain services to kids.

***********

The next stop on my journey brings me to Michelle, a petite woman who is always smartly dressed. I met Michelle when I was a brand new assistant principal at Orange High School (OHS). I was overwhelmed with the responsibilities but Michelle, another assistant principal, kindly mentored me and showed me the ropes. Michelle is a senior assistant principal who works in the same school and district as I do. She has been an assistant principal for nine years at OHS, a Title 1 school, and has been an educator for fifteen years. She was a junior
assistant principal prior to becoming a senior assistant principal, a job she has held for five years. My hope is Michelle’s stop on my travels will provide me insight into how one works within the tensions of the organizational requirements of enacting high-stakes assessments and personal belief systems.

Tonight I am meeting Michelle to talk about her feelings and experiences related to high-stakes assessments. Michelle has openly questioned the purpose of high-stakes assessment in conversations with me over the last three years. I am eager to delve deeper into her perspective on high-stakes assessment from her unique position as a senior assistant principal, plus Michelle is not afraid to share her opinion. Indeed, Michelle is often sought out by central office administration to sit on committees because they know she will share her candid and honest opinions no matter who else is in the room. I am counting on Michelle’s candor to provide me with rich data for my pilgrimage.

**Culture of Measurement**

She greets me warmly as we sit down in a quiet café. We order and we briefly discuss the ongoing of our school district before we switch to the topic at hand. I briefed Michelle about my purpose and goal of this interview, and then before I could ask her a question she began discussing a teacher who just recently transferred to OHS:

*My heart breaks for this particular teacher. She was a former reading coach at her previous school and because the school is a D school, and it is on the state turnaround list, anyone the state views as not having an effective VAM score cannot work at that school. Unfortunately for this particular teacher, she being non-classroom personnel, as often we do, she was assigned to classes where there was a teacher vacancy. So because there was a vacancy no instruction or limited instruction was happening in that classroom and, unfortunately, those*
kids did not perform. Her state VAM score got calculated based on those classes she was assigned to, that she wasn't actually the teacher for, and so she lost her job and had to be transferred to another school because the state would not allow her to stay at a turnaround, “state turnaround school” with a less than effective VAM scores. So this poor woman, who from everything I have seen is fantastic at her job and should not be in this position, has had her reputation tainted because of things that we do in our district. We assigned non-classroom personnel to vacancy lines because it helps to make our technology work. Also because of what the State is doing. They're calculating VAM scores differently than our school district calculates our VAM scores because her district VAM score is highly effective. How can the district VAM scores be highly effective in the State’s VAM scores be ineffective?

All based on high-stakes testing results. It is OHS’s gain to get this teacher, though, I add.

Oh, it is our gain, but it does make me wonder, our VAM scores as administrators are based on school-wide data so it does make me wonder what our state VAM Scores are? Would we be losing our jobs if our school dropped to that level? We look at what has happened recently in the school district where they had to come in, and they had to assign new principals.

Michelle’s story of the teacher who lost her position at a D school because of her state VAM scores is indicative in a culture of measurement where human beings are merely abstract (Padilla, 2005). Darling-Hammond, Amrein-Beardsley, Haertel, and Rothstein (2012) explain, “researchers have found that teacher effectiveness ratings differ substantially from class to class and from year to year, as well as from one statistical model to the next” (p. 9). Michelle asks why the disparity between State and District VAM scores? Indeed, Darling-Hammond et al. (2012) report even within the same content area different tests can have very different results,
and VAM does not account for the backgrounds of students assigned to teachers. For example, students’ poverty, attendance, English language ability, or whether or not the student has a learning disability are often not controlled for statistically (Collins & Armein-Beardsley, 2014; Darling-Hammond et al., 2012). As evidenced by the teacher who came to OHS from the D school, her rating on the State VAM scores was poor, but her school district’s VAM scores were outstanding.

Further, Sorenson (2016) explains VAM does not indicate causality between student test scores and teacher effectiveness, rather VAM is a correlation. Predicating a teacher’s job on an inexact correlation did not sit well with Michelle, but at the same time, OHS benefited from this teacher’s plight, which Michelle acknowledged. Nevertheless, due to the value placed on VAM via high-stakes assessments in the culture of measurement, which is pervasive in the public school system in the United States, the teacher’s reputation was still tainted even if those at OHS think getting her at their school was a boon.

*This is something that I’m going to share: we went to data updates, student data updates, and in the past I guess I really haven’t paid attention. I would go to the file I would withdraw the kids who withdrew during the testing season and get rid of them, but I never downloaded the state report that really said which kids were required to test and which kids were not required to test. This year I did that because as I get more experience in my job I want to better understand every aspect. The first thing I noticed when I downloaded the report was there were a ton of kids on my report for subjects we were ‘required to test’ that were not sitting in those classes. We didn’t test them because they're not sitting in EOC classes, therefore they should have never ended up on my ‘required to test report.’ These are all kids counting against me for participation, and when I did further investigation what happened is that a lot of these kids were*
retaking the course, possibly for credit recovery, possibly they signed up to take the course in virtual school and then dropped it, whatever. The state captured them and put them on my file as required to test so had I not gone in—it took an entire week an entire week and two APs to go through the report going kid by kid analyzing why they were on the State's ‘required to test report.’ But we didn't have them on our to be tested report so that's a big concern, how inaccurate has my data been, right? Here's the other thing I learned, I think this is just high school, but another thing that I noticed some kids their testing status wasn't 'required' it was 'optional' so I asked how is that? How is this kid optional? What I was told is that kids that are in, let's say for example, AP US history. It's optional. If they take the test it counts in the school grade if they don't take the test it doesn't count in the school grade. It doesn't count against the school. Just because I'm this type of person, and I try to like think of how I can strategize to make a situation to my advantage I was thinking to myself, ‘if I were an unethical person’...because we know a US history test is really it is a literacy test so if the students do not have literacy skills they cannot do well on the test. In theory, if I was an unethical person I could schedule all of my Level 1 and Level 2 kids into AP US History and then not have those kids take the US History EOC and guess what? None those kids count against me and so 10% of my school grade would be calculated on just the kids with level 3, 4, and 5 on FSA ELA. If these students I had take the US history EOC. If that is not an issue for the state to examine I'm not sure what it is because if this little southern belle has figured this out I'm sure someone else has, too.

This is an example of “gaming” by “reshaping the testing pool” of students to improve the school or district’s overall outcomes (Figlio & Getzler, 2002, p. 1). Gaming state accountability systems has been a practice since the implementation of accountability policy
(Armein-Beardsley, 2009). Although Florida has closed loopholes others remain, such the one described by Michelle.

The opportunity is sitting right there out in the open for any school administrator to use to game the system. It seems cruel to subject students to advanced placement courses when their previous performance on high-stakes assessments suggest otherwise. However, Michelle’s story left me wondering why their previous performance on high-stakes assessments dictates a student's access to advanced placement courses. Although Michelle was discussing the opportunity to game the system by reshaping the testing pool, why we have barriers in place for any course for any student? I know advanced placement courses are challenging but the students and their families should be the ones who decide whether or not a student takes a course.

The language Michelle used when speaking about groups of students struck me as I transcribed the interview. Much like with Nora and Cher, I am so steeped in the culture of my school district and the way of speaking about students as abstract scores on high-stakes assessments, I did not notice this during the actual interview. Michelle named the groups, Level 1 and Level 2 kids, based solely on their scores on FSA ELA high-stakes assessment. Michelle and I also refer to students on the cusp of passing the FSA ELA as bubble kids:

When I sit with my instructional coaches those are the conversations we have. When district-level personnel come in they want to specifically know what our plan is to move our bubble kids. Those are the conversations they have with us, then we turn around and we have with our instructional staff. And your high twos, where you can move to a level three those are your golden kids. Yeah, because not only are you going to get a learning gain on them you're also going to get a point in proficiency. And then your other golden kids are your kids that are in your bottom quartile because you're going to get a learning gain on them and then you're going
to get a bottom quartile gain on them. Any kids that you can move you get two points instead of one that's who you are focused in on.

Back in the AYP times it was even finding the students who were hitting on a four or five different areas, I reminisce.

The game plan changes depending on you know the point in the game that you're at. So, for example, every year I go into summer and I start off you know thinking 'I'm going get the best teachers providing the best instruction in front of all of my kids.' Okay, well then August comes around and we're about to start school, and I look at where I have vacancies and then I start to think and start to shift instructional lines based on where I have vacancies. If I have a reading vacancy, well guess what? That long-term sub I'm sticking with my juniors because that's the area that is not going to this year hurt me on school grade. Alright, because my ninth grader is calculating to school grade, my tenth grader is calculated in this school grade, and my 12th grader sitting in a reading class better pass our reading test otherwise they are going against me in graduation. So the way they calculate in school grade if I have a vacancy position, my juniors are getting screwed. That wasn't my intention at the beginning of the summer. My intention at the beginning of the summer was every kid was going to have a quality teacher but when reality sets in I have to make some tough decisions and every manager every business owner has had to make those tough decisions. What the problem is, is that we should not be treating schools and children like corporate America however our State is dictating we treat students in schools like corporate America because that's how they're treating us.

Accountability, I comment.

Yep! Well, but it goes beyond accountability. It's the accountability in a corporate system there has to be a difference between accountability for a for-profit organization and
accountability for a service nonprofit organization. That's the difference that they are not distinguishing because they want to run us like a for-profit corporation because that's what makes sense to them. Fortunately for me I have a business background that's what I was educated in and I own my own business so I can look at the situation this way and I can make these tough decisions that I never wanted to be in a position to have to make but I have to make them because, ultimately, I have to do what's going to help those kids. My best reading teachers in those situations are going to my 10th graders, and then my second best reading teachers are going to my 12th graders, and then my third best reading teachers are going to my 9th graders, and then the worst of the worst are going to my 11th graders. Unfortunately, that's how I have to do it because of how it's impacting me on school grade. Quite frankly, you know school grade is also whatdictates whether or not kids are choosing to come to my school or choosing to go to a charter school. The state has set this up as a for-profit business. It shouldn't be that way but it is that way. If they really want to do what's best for kids they need to change the way they're treating schools, and they have to recognize there is a difference between how you treat non-for-profit service organizations versus for-profit business organizations.

As previously mentioned, gaming the system has been in play in public schools since the dawn of accountability (Armein-Beardsley, 2009). This is prevalent in a culture of measurement, where scores on high-stakes assessments are more prized than the humans taking or administering them. School administrators focus on doing whatever it takes to perform well on high-stakes assessments, but gaming requires the administrator and/ or teachers to sacrifice their own ethical or moral principles and oftentimes student well-being (Armein-Beardsley, 2009). Michelle sees the opportunity to game but resists taking it. She recognizes students who struggle with reading would not find success in AP US History and placing them in the course to
achieve a 10-point gain on the school grading system could be harmful to those students. However, she also acknowledges if she found a loophole then someone else likely did, too. Indeed, Armein-Beardsley (2009) claims those gaming the system of accountability know what the rules are but choose to find ways around the rules and are better at gaming since they are familiar with the rules in the first place. The temptation to game is ever-present in public schools where the culture of measurement is omnipresent, and, in some cases, those willing to game are rewarded with recognition or even promotions. At the same time, Michelle realizes she must at least negotiate the ethics around which students receive the more accomplished teachers based on which students will be testing. It is clear she does not like to compromise, but, given the lack of qualified teachers applying for positions at OHS and the culture of measurement she works in, Michelle is doing what she thinks is right for the common good, which is an example of the ethic of the profession by doing what is best for the collective instead of the individual.

Michelle comments on the use of business tactics in public schools and notes how the purposes of public education and for-profit business practices within public schools do not align with one another. This misalignment of purpose is rooted in neoliberal and neoconservative values focused on privatizing public schools and adopting corporate practices (Giroux & Saltman, 2009). The goal of education via a corporate model is, “the capitalist imperative for the growth of financial profit at any cost” (Saltman, 2009, p. 11). The corporatization of public schools views students as commodities, products, or widgets (Au, 2009; Saltman, 2009). Thus, the high-stakes assessment serves to determine the success of the public school system and the individuals within the system. At the school and district-level grades based on high-stakes assessments represent the growth or financial gain or loss of the public school system in question. At the individual level, high-stakes assessment determines the worth of the student and
teacher. In a corporate model of schooling high-stakes assessments determine profit margins and individual successes. This model is encouraged by the culture of measurement because the statistical measure represents profitability and encourages gaming the system because those willing to compromise their ethics and participate in gaming the system are often rewarded, and those that do not engage in gaming the system are often punished.

**Testing Season**

*Here’s another thing,* Michelle says, *You're always testing the ESE kids in the morning, and this is the issue we ran into, one of the ethical dilemmas that I had this week, is my science coach came to me and she said, ‘you had my 9th graders out in the morning for testing’... because we're testing by teams and she's on the ESE team. She said, ‘you had my 9th graders out in the morning two days this week. Next week when they take their Algebra EOC they're going be out two days in the morning.’* She tells me, ‘I've got half of my kids that I will not see for a total of four days, and the following week we do the Biology EOC.’ *So because of the way the test is structured, we are almost forced to test our ESE kids in the morning, but with our neediest kids, it's requiring us to take them out of the same classes for long periods of time. If that's not an ethical dilemma the state needs look at I don't know what it is.*

Michelle calls out what she considers an ethical tension. Michelle is referring to the amount of time away from instruction due to actual testing, which greatly impacts students who receive extended time testing accommodations due to IEPs, 504 Plans or ELL status. Most research focuses on test-preparation during instructional time as impacting instructional time and narrowing the curriculum. However, the time students, especially students who need extended time for testing, lose out on instruction is a sparsely studied aspect of the loss of instructional time, although an ethical dilemma of implementation of high-stakes testing for Michelle (Simon,
Nelson (2013) found students in grades three through nine can spend as much as 55 hours in testing, and that number did not include hours spent sitting in a testing location before and after actual test administration nor test-preparation. In general, as Simon (2010) points out, Florida has a state statute forbidding test preparation in place of instruction; however, the statute also lists the many loopholes school can utilize to prepare students for state-wide high-stakes assessments (FL K-20 Education Code, 2018). For example, one of the loopholes allows for additional diagnostic testing for Level 1 and Level 2 students (FL K-20 Education Code, 2018). Many students who perform at Level 1 and 2 on high-stakes assessments are also the same students who have IEPs, 504 plans, or ELL statuses. This statute does not address the need for compensatory instructional time to make up for the amount of time the, as Michelle suggests, “neediest” students spend in high-stakes assessments or preparing for high-stakes assessments. The much-needed extended time accommodation for students with IEPs, 504 plans, or ELL status should assist the student instead of creating a more significant disadvantage, such as what Michelle described happening at OHS.

Michelle adds to her concerns about the loss of instructional time due to testing: *So it's completely unfair to everyone involved. It's unfair to students because we are pulling them out of their classes time and time again. For my ninth graders—my ninth graders take a two-day Algebra 1 test, they take a three-day FSA ELA test, they take a Biology assessment. What is that? Six days that they're out of class but because I can't test all of my ninth graders at once and could only, at max, test half of my ninth graders at one time. They have lost not six days of instructional time but they have lost twelve days of instructional time. We start off with it's unfair to students because that's almost three weeks of instruction that they lost. Twelve days of instructional time it almost equates to three weeks of instruction that they lost. That's the*
students that it's unfair to, then you got to think about the people that you’re asking to proctor, and in my case I'm a Title 1 school. Now I fortunately in some ways have more resources like non-classroom teachers than some other schools do. I'm put in the tough situation of do I pull classroom teachers out of the classroom to test kids or do I lean on my non-classroom personnel to test kids? Counselors, coaches things like that, and if I lean on them then what am I pulling them away from doing? Which by the way they're not evaluated on so I’m telling people, ‘guess what you have to go,’ and again…I literally test every month of the year. I test for I would argue maybe as much as fifty percent of these people's time I am pulling them away from the job that they're evaluated on and I'm asking them to do testing, which is something that doesn't hit them in their evaluation at all. How is that fair to them?

The allocation of resources, human resources in this instance, is an important aspect of high-stakes assessment implementation during testing season. Michelle highlights the ethic of distributive justice when she allocates resources during testing season. Indeed, testing season can be likened to project management because it is a time-bound, temporary, scheduled, resource necessitated, “execution of activity,” leading to the production of a service (Schwindt, 2005). Testing season is time-bound (albeit reoccurring), temporary, scheduled, resource necessitated activity, which leads to the production of test taking by students. During testing season, resources are reallocated from their original purpose. Michelle explains to successfully complete a full season of high-stakes assessments administration she must utilize her human resources in ways never intended. As her story indicates, instructional coaches are not evaluated on test proctoring but since they are not tied to classrooms she uses them for testing half the year.

This practice is common in the school district we work in, and having been an instructional coach myself I can attest to this because I spent an exorbitant about of time
proctoring high-stakes assessments. If I make a rough estimate of the time I spent as a proctor at the high school level, I would have to say it was close to half of my work time. That is a lot of time to pay a resource teacher their hourly wage to proctor a test. It was definitely not the best use of my time, and it was not part of my evaluation. It never occurred to me to refuse.

Michelle explains further:

It's not fair but if I don't have them doing it then I'm pulling teachers out of their classrooms where they can't be with their kids and instruct their kids. Not fair to teachers. Now let's talk about how it's not fair to me and how much time that I'm spending. Because I have to orchestrate it I have to plan it out and guess what it's not part of my evaluation either. If it's not part of my evaluation but it has to be done because all my other things in parts of my job that aren't part of my evaluation still have to get done. Guess what? I'm doing this on my personal time. Now let's talk about how it's unfair to my husband and my son who gets what? In the springtime for sure they do not see me at all how is that fair? And I don't get compensated for that, they don't get compensated for that, it's completely unfair to everyone involved in this process directly and indirectly.

From February through the end of May, I work every Saturday I do it so that my kids can go work with teachers and get extra tutoring support, but I also do it so that I can do all my testing plan, my testing schedules, my testing list get done. It is anywhere from five to six hours each Saturday, and it's away from my family. So it was funny, I was just telling my reading coach today that I used to really like my birthday, which is coming up in two weeks now I hate it because I can't take a day off. I can't even take the weekend off surrounding my birthday and to spend time with my family. I do all of this for my students in my school that I love very much and considered to be part of my family, but it definitely takes its toll on you know that other part.
So you have to weigh out what you're doing in your personal time with what you're doing in your professional time because you are you taking away from your own life. Are you? I ask.

Yeah, I got another story for you so this week because my husband, he travels for his business, and this week kicked off a lot of testing for us. My own son had testing, but my husband was in the Keys on business but in order to make sure that I got to work to get testing started for my students and my teachers I had to wake my son up an hour early. He also had testing this week, and he went to work with me and we were at school by 6:30 every morning. He sat in the conference room and ate his breakfast half asleep as I prepared for testing. Then, when it was late enough for me to drive him to his school, I drove him to his school and then he took his own test. Yes, my own son had personal sacrifices and it's due to my commitment for testing happening at my school.

Testing season is crazy, I state emphatically.

I want to say is that testing season never ends all right because I currently have students coming to school that will take their PERT test in June and then retake their Algebra and Geometry EOC's in July. Testing season has not ended we are just going into a new fiscal testing year, Michelle says half-jokingly. I'm personally feeling guilty about like the amount of time that I spend on these testing plans. It's required but it isn't part of our evaluation. There's nowhere that people are saying, ‘yeah you did an exemplary job of testing you deserve a performance pay.’ When are you doing this? When are you preparing for these tests that we have to we have to do, and if we don't do them we will be F school because if you don't get a 95 percent participation rate you are an F school, so when do we do this? Michelle asks no one in particular.
Michelle experiences testing season as an ongoing, yearlong endeavor. Yet, Nelson (2013) found testing took only an average of 62 hours for a school administrator. For guidance counselors and paraprofessionals, testing took up approximately 100 hours (Nelson, 2013). According to Nelson, “test administration required substitute teachers to proctor tests or supervise the classrooms of teachers who were engaged in other testing tasks. Across all districts in the sample, 1,021 substitute teachers facilitated testing or supervised students” (p. 19).

Although Nelson’s (2013) study seems to be in-line with what Michelle describes for students, teachers, and other resource personnel, it seems to be way off in the approximation of the time Michelle describes spending on testing herself. This may be due to differences in in who oversees testing in different school districts.

Further, Michelle connects her evaluation to testing and testing season. The assistant principal rubric in our district does not mention testing in any capacity. Nevertheless, it is understood she is responsible for planning, scheduling, and carrying out the administration of a plethora of high-stakes assessments. The State penalizes schools for not testing students in the school-grading system. Schools must test 95% of students required to take a particular high-stakes assessment or the school may receive a grade of F (FL K-20 Education Code, 2018). Again, there is nothing written about what would happen to the administration at a school receiving a grade of F because they failed to test 95% of the students, yet there will surely be a consequence of demotion or an unwanted switch to another school.

**Ethics of Implementation**

Our conversation turns to the ethics of implementing high-stakes assessments:

*I don't necessarily disagree with us having state testing. I disagree with how we're using the state test. We all know as educators’ grades are very very subjective, and I think that you do...*
need some measures as to whether or not students are actually meeting standards so that you know how to provide remediation, if it's needed, and so there is that measure to be fair from one teacher to the next. However, you know these measures should be formative they should not be for the purpose of evaluating whether a school is effective or not effective or whether a student is capable or not capable. You know we have very very capable students that don't necessarily perform well on these high-stakes tests, but you put them in a real-world application, and they have skills our other are very high functioning kids don't have, which will make them very successful in life if it's channeled in the right direction.

Wayne Au (2009) argues that high-stakes, standardized tests create an atmosphere where students are merely widgets and have no value as human beings; rather the students’ scores on the standardized tests are their true value.

Really, I mean, really smart kids, I mean when you think about the obstacles they overcome in their lives, and the things they're able to handle and deal with...amazing, amazing kids. We have just beaten them down with these tests, and we essentially suggested that they're failures by testing them over and over again.

Michelle echoes Au’s (2009) observations: “students’ lives, home cultures, histories, educational differences, and socio-economic conditions mean nothing within the logics of high-stakes standardized testing” (p. 43). Michelle clearly struggles with the, “the objectification-through-quantification of students” (Au, 2009, p. 41) because she works with and knows the real students behind the low test scores and does not view her students as ineffective widgets. Yet, Michelle does not fully reject high-stakes assessments. She finds some ethical congruency with having a level of teacher accountability and student mastery of content:

What if there were no tests? What if we were just teaching? I ask.
That’s a problem, Michelle stated emphatically.

Why?

Because we know that there are certain teachers in this profession because they truly are teachers. They are trying to move kids. They want students to have the 21st-century skills students need in order to be successful beyond high school and into college and into careers. We also know that there are teachers that either they're in it for a paycheck or they just don't know what kids need because maybe this is a second profession for them. You have to have an expectation. There has to be that expectation. For each course are their standards and you have to have some form of points, you have to have some form of assessment, it doesn't necessarily have to be a single assessment at the end of the year. I think that's what your study is about is it is a single assessment at the end of the year the right tool in order to assess whether or not students have mastered standards. I would argue no but there has to be assessment and criteria demonstrating students have mastered standards. Now in my opinion, that doesn't necessarily mean that a kid passes or fails a grade. It means that you know they need to continue to work and build.

If there were no high-stakes assessment then you would have time to spend in classrooms observing teacher practice, which is another form of accountability, right? I push.

At some point, you still have to have an assessment and you have common criteria, and, quite frankly, you have to have common assessments and benchmarks so you know a student or students’ performance or even a teacher's because we all have personal biases that get in the way. You have to have common criteria. You have to have common measurements. My issue is that it doesn't necessarily have to be one and done. These should be continuous. They should be throughout the year and we have to spend the time training teachers to really understand what
the criteria is, what their standards are, what those standards are asking kids to be able to demonstrate. If they can demonstrate that and we have evidence to prove that it doesn't only have to be proven on one test.

I actually got into this conversation with another group and we were talking about it—it was our instructional coaches. We were really talking about what I'll be using these assessments for. Are they formative assessments to help us better instruct students or are they summative assessments? Do standardized assessments need to be there so that we can judge ourselves and the quality of education that we're providing our students against other schools in our district, other schools in our state, other schools across the nation? Absolutely because education is a very subjective thing. We all look at our curriculum and our standards and teachers we treat them as professionals and they make some subjective decisions as to how they are going to assess it or how they are going to teach that and then what area is they're going to place emphasis on. You have to have forms of assessment in order to compare yourself to other people and so that you're trying to get a standard, but again for what purpose are you using that? Is the purpose to inform and to make educational decisions or are you using that you know assessment to make judgments and to penalize people and penalize schools? Sadly, right now we are using these assessments not to inform but to penalize and judge people and sadly we think we're penalizing and judging school systems schools and teachers but who we are really penalizing and judging are ultimately our students.

Burger and Krueger (2003) posit high-stakes assessments are useful for determining if a student is learning. “It is also important to view tests as supporting curriculum standards to help teachers develop the necessary content competencies in students that will lead to curriculum mastery” (Burger & Krueger, 2003, p. 16). Michelle is advocating for using high-stakes
assessments as tools to determine whether or not students have mastered content standards but also as a measure of teacher effectiveness. Popham (2001b), a long-time researcher of assessment and evaluation in education, argues standardized testing can be appropriately used to determine student achievement but is not appropriate for use in determining school or teacher effectiveness. Further, Armein-Beardsley (2014) states, “teacher effectiveness is simply a difficult construct that is not easily reduced and quantified for such simplistic purposes unless of course a series of heroic assumptions are used as means to justify oft-political ends” (p. 89). Thus, Michelle makes clear the contradictions within the ethics of implementation of high-stakes assessments. She finds value and congruency in the intentions of accountability but has ethical concerns with the impact the tests have on students.

Michelle returns me to her ethical dilemmas with testing:

_The biggest dilemma I have with testing comes with the retesting and the retake tests. At high school, once the student takes the test and does not perform at a proficient level we have to just keep retesting them until they get to that proficiency mark for those at graduation benchmarks, like the tenth grade FSA Reading and Writing and the Algebra 1 EOC. My dilemma with retake testing is that when you have your level 1 students, who you know they've been in school for maybe a month, they haven't had the remediation time that they need in order to—I don't want to say I know that it is absolute, but there is a high probability that if I test them they're still not going to get proficiency. But you have this dilemma because you're being asked to test them because they could potentially be showing growth, and we want to measure that. You're putting this kid in front of this high-stakes test. There's all this anxiety build up from this kid they have over the last ten, twelve, thirteen years have sat and taken tests and gotten low ratings on those tests, and they are they so beaten down. I mean they just work themselves up to_
a frenzy, and it's just heartbreaking because they're coming to you and saying, 'I don't want to take this test. I know I'm not going to pass it,' and because supposedly the data is important you're expected to test them. You're trying to encourage them to test and take it because maybe, just maybe, this time they're going to pass it even though there's a high probability they won't.

I agree with Michelle. The system abuses the students who take and retake high-stakes assessments. The state does not care that they are broken down by over-testing and repeated failures. This is a high-school phenomenon because retaking high-stakes assessments only occurs for the assessments required for graduation. However, the same kids who did not reach proficiency on their third grade reading FSA are also the same kids who are retaking the tenth grade reading FSA when they are seniors. Janesick (2010) is correct this is child abuse.

It was especially difficult this year after a hurricane because we lost a week of instructional time. I actually sat with my reading coach, and we made the decision to give our juniors the opt-out. If they did not feel like they were ready, if they felt like they were under too much emotional stress, because we still had a lot of families that we're dealing with flooding and power outages, they did not have to test, and we were not going to make them feel pressure to do so at all.

Did you get any push back? I asked.

Michelle continued, I didn't tell anyone! I'm the senior assistant principal at my school so I told my principal and he was very very supportive if it's right for kids he's going to support it. But, I didn't ask for permission, and if I don't think our school district would have a problem with it, but if they did then I'm okay with that because I feel like I have justification for what I did. Ultimately, I did it for the kids.
So nobody said ‘oh you only tested this many of your juniors this year so what’s going on with that?’ I inquire.

Because it was a retake situation, but had it been 9th and 10th graders first-time test takers for that test, if we wouldn't have hit that 95 percent mark, and you know if you don't hit a 95 percent tested mark in any one subject area your school is automatically rated F.

Michelle touches on an aspect of high-stakes assessment she struggles with, but she has also found space to resist implementation. Obviously, it frustrates Michelle to witness students’ repeated failure on high-stakes assessments. Horn (2003) suggests the inclusion of high-stakes assessments as graduation requirements perpetuates a “cycle of failure” (p. 30). Students who have continually failed to pass high-stakes assessments, which serve as state exit requirements for graduation often have been less than proficient on the reading and/or math high-stakes assessment their entire school lives. Students who make repeated attempts at taking exit-required assessments more likely to drop out (Nichols & Berliner, 2007, 2008).

Yet, Michelle used the space within retesting as an opportunity to resist high-stakes assessment implementation. She was fully aware she could resist without the consequences from the state she might face had it not been a retake assessment. However, providing students who had just faced the trauma of a hurricane the ability to opt-out still could have been risky within the school district itself. Thus, Michelle chose not to ask for permission. Anderson and Cohen (2015) refer to this as “counter-conduct” which means, “attempting to work subversively and productively within the constraints of the current policy and cultural context.” (p. 9). Michelle subverted the implementation of high-stakes for graduation requirement retake assessment within the constraints of the State’s policy.
Michelle reiterates, *The challenges that you have to overcome with the kids that are your level 1 and level 2 kids that are attending these high need schools, and for us when they get to high school it really is instilling that hope in them. Instilling that belief they need to continue on and persevere. For the entire time they have been in school and they've been testing, they have gotten the results back...what's the terminology on the result? I mean it's so derogatory the descriptions of the levels—*

*Inadequate.*

*Yes, inadequate, Michelle agreed, the state verbiage, yes, state verbiage inadequate.*

*The kids know they are inadequate. They know that they're going into boom-boom classes. They know that they're no good and so by the time they get to high school they have now seen this and heard this for 10 years. There's 10 years of data to go back and reinforce this, when I say data I mean test results, and the biggest thing that we do in high school is cheerlead and bring the kids up again so that they don't give up and drop out.*

Michelle adds, *let's put numbers to that. Last year for us ten percent of our school population earned a certificate of completion. They did everything that they were expected to do. They sat for every test we gave, which so for the math, the Algebra test we give that four times a year, okay? For the reading test we give that twice a year, as well as we give five PERT attempts to meet the math benchmark, and then there are multiple SAT at ACT attempts. That was 10% of our population last year that earned certificates of completion. The only thing that prevented them from graduating was passing either the FSA ELA test or the Algebra 1 EOC either or both. 10% which equates to about 36 kids our graduating class was about 370, okay? That's a tough thing to say to a kid and a parent, 'yeah your kid came in all these 13 years and did all the work but you didn't pass the test and so we're going to sort of give you this, attaboy!*
Good job! Thanks for showing up,’ but you know it really doesn't equate to anything. I heard at one point the state was looking into a portfolio for graduation or for meeting state graduation benchmarks it was in a report that I read months ago but I haven't seen anything else about it. Then as if the recorder was a direct line to state lawmakers, Michelle added, “yeah state legislators where are you at on that? I got to show you 10% of my kids that work their asses off. What are you doing for them? Because I got evidence to prove that they deserve their high school diploma.

Let’s talk about what these tests and the school grading really does –it really creates segregation, Michelle points out what is obvious to both of us but seems invisible to most others. Up until this point our school district’s solution for grade issues has been in what they call the school choice program where they're choosing through magnet programs or to charter schools. Middle-class kids into these high poverty areas and so the solution is desegregation, but the problem is the stigma of school grades based on test scores is creating more and more segregation as we talked about earlier when we are talking about kids that should be feeding into our schools are choosing away from our schools because they don't want the stigma of going to a low performing school. Low performing schools will continue to be low performing schools.

I agree, It's called creaming. When charter schools take only high-performing students into their schools, and then everybody that’s left behind in the public schools are struggling students (Lacireno-Paquet, Holyoke, Moser & Henig, 2002). So the school continues to struggle. In fact, the school can only ever struggle because once that starts happening there’s no way to rebuild. You're not going to get that student body back that can help you improve your school grade by performing well on high-stakes assessments. The school rating system, which is based purely on high-stakes assessment results, helps perpetuate creaming by ranking schools. I think
what you’re seeing are parents seeing OSH poor school grade and using choice options to move their students to the area’s charter schools.

Michelle nods her head in agreement, *Sadly, when a school has a reputation of being high needs or is low performing our high performing students choice away from us. This was very evident this year when I had an AP track team so students that were looking to go into AP Human Geography. They would be taking all honors courses, and I was projecting 150 kids to show up on day one to be part of that team. I had less than 70 kids show up so that population of students was choosing to not come to my school because of the reputation. The reputation based on the school grade and the school grade based on test scores.*

Again, Michelle addresses an unethical consequence of high-stakes assessments: resegregation due to school choice. In 2002, Boger predicted how federally mandated high-stakes assessments would impact segregation:

> When "high-stakes" accountability measures are imposed upon, and interact with, school systems hampered by growing racial segregation, they threaten instead to worsen the plight of schools that are disproportionately filled with nonwhite children from low-income families as middle-class and White parents, along with better trained, more highly qualified teachers, abandon those schools. In short, the convergence of racial resegregation and statewide, high-stakes accountability measures is likely to increase the racial segregation and economic isolation of some public schools, whose students will disproportionately fail state accountability tests, thereby entrenching broad patterns of grade retention, student demoralization, and teacher flight. (p. 1375-1376).

This prediction has become a reality throughout the United States. Several researchers have found that school choice, particularly charter school choice, has resulted in resegregation in the
United States (Bifulco & Ladd, 2007; Booker, Zimmer, & Buddin, 2005; Clotfelter, Ladd, & Vigdor, 2013; Cobb & Glass, 2003; Frankenberg, Siegel-Hawley, & Wang, 2011; Furgeson, Gill, Haimson, Killewald, McCullough, Nichols-Barrer, Bing-Ru, & Verbitsky-Savitz, 2011; Garcia, 2008; Glenn, 2011; Michelson, Bottia, & Southworth, 2008; Nathanson, Corcoran, & Baker-Smith, 2013). Indeed, school choice seems to be the latest incarnation of White flight. Michelle is experiencing this consequence at her school, and she concludes that school choice coupled with school grades based on the schools’ collective student performance on high-stakes assessments results in resegregation.

Michelle used the term “choicing” to refer to families using choice options to attend other public schools or charter schools. This new present tense of the verb to choose is reminiscent of googling. Neither are words you find in the dictionary, but neither requires explanation. I knew exactly what Michelle meant when she said choicing. It means parents and students are choosing a school other than their neighborhood school. This act has become so common that Michelle can parse it down to one present tense verb. Consider what that means in the context of public schooling in the United States and Florida: school choice has a major impact on public schools.

**Conclusion**

It is evident from Michelle’s narrative the job of senior assistant principal at OHS was wrought with both personal and professional ethical dilemmas. The ethics of implementation are made real through her story of subverting implementation by allowing recent hurricane victims to opt-out. Michelle’s ethical dilemmas about the time she devotes to planning and overseeing high-stakes assessment at her school present the reader with a true story of how school administrators are required to focus on the professional over the personal, even when managing high-stakes assessments are not on their evaluation. Perhaps the most difficult for Michelle was
telling the stories about the culture of measurement where she demonstrates students and
teachers are not valued as people but just statistics. Her experiences with ethical struggles
surrounding high-stakes assessment are present in her stories of implementation, living through
the testing season, and the persistent culture of measurement. Michelle’s honesty and her
willingness to discuss serious topics many school administrators would avoid, in particular,
resegregation, clarify and name real-life problems in the current atmosphere of high-stakes
assessments and accountability in public schools. On this stop of my journey, I learned from
Michelle’s stories some of the ethical dilemmas school administrators face when living and
implementing high-stakes assessment. My next stop brings me to Dave’s tale.

Vignette

‘Mrs. Canady, I need to talk to you,’ Daniel tells me as he walks quickly to catch up with
me. I am heading to lunch supervision and am running late. Daniel is a senior and has been
one of my favorite students since his freshman year. Daniel is a smart and funny kid who is
popular with his peers. Daniel has not been the perfect student. He has gotten into some minor
trouble over the years but has begun to mature into a better version of himself. I stop and look
at Daniel. He looks pensive like he has the weight of the world on his shoulders.

‘Today is my last day,’ he claims.

‘What?!? Where are you going? Did you move?’ I demand.

‘No, I am going to be an apprentice for my brother. He’s an electrician. I am going to get
my GED,’ Daniel explains.

‘But why? You are too smart to drop out.’

‘I am never going to pass that test. Why waste this year in school, when, at most, I will
get a certificate of completion.’
‘Your retake results came out?’

‘Yep, I didn’t pass.’

‘What about the SAT or the ACT? There is still a lot of time and chances for you to pass left,’ I plead.

‘Miss, this is the thing; I am not even close to passing. I am not a good test-taker. I am tired of working so hard and failing. Thank you for believing in me all this time. My mom wants you to keep in touch with her.’

I wish my story of knowing Daniel and his struggles with achieving a passing score on FSA ELA or a concordance score on SAT or ACT was an isolated one. Unfortunately, I have known too many Daniels and Danielles in my work as an educator. I want to preface this section by stating that vocations such as electrician are fantastic careers. I do not believe every student needs to go to college to have a fulfilling and productive life. However, I do not believe a student should have no chance of obtaining a high school diploma when they have done everything but pass a high-stakes assessment. Students like Daniel feel they have few options; in fact, Daniel has more options than most students in his situation. Unlike most high school dropouts, he has real earning potential as an electrician. Most high school dropouts are jobless or work in low wage jobs (Robison, Jaggers, Rhodes, Blackmon, & Church, 2017).

As a high school administrator, I find it difficult to watch kids struggle and repeatedly fail. My job as a reading coach was 80% cheerleader, and I still adopt that role as an assistant principal. Many students like Daniel have been failing the FSA ELA and its predecessor, FCAT, since third grade when they took their first high-stakes assessment. These students are often the same ones retained due to poor performance on the FSA ELA in grade three or retained in kindergarten, grade one, or grade two as a preemptive measure. They never scored a passing
score on the reading high-stakes assessment in their lives. My ongoing ethical dilemma associated with high-stakes assessments is mainly caused by watching young people with promising lives beaten down by failure for years upon years.

However, there is a solution to part of this problem – portfolios. Currently, Florida offers portfolios as an alternative to passing the third grade FSA ELA. This option is not available for students in high school. Every high school student must pass the 10th grade FSA ELA to graduate. Only some students with disabilities can obtain a waiver and earn a diploma without passing the high-stakes assessment (FLDOE, 2018b). ELL students should be provided the test in the language of their choice or given a waiver similar to what ESE students receive. Yet, these solutions lie in the hands of state policymakers, which is well above my power and current situation. Again, I feel powerless and have resigned myself to witness the reality many students experience.
CHAPTER 7: DAVE’S TALE: THE CYA MENTALITY JUST ROLLS DOWNHILL

I kind of felt like a lot of what they had me doing was unnecessary work. I felt like it was just triple and duplicate it was beyond what it needed to be. I mean I believe in having contingency plans ready anticipating what might happen but it was like off-the-wall stuff like, ‘what if we have a school shooting that day?’ Well, I’m not even worried about testing to be honest with ya.

Was that actually a question? I ask in disbelief.

Dave responds in the most matter of fact tone, that was a question like, ‘how do they secure the testing materials if we will be going to lockdown?’

Still in shock, I clarify what I am hearing, how will we secure the test if we went on a lockdown because we had a school shooter?

Dave simply nods his head, mm-hmm.

************

My next stop on my quest brings me to Dave. Unlike my other participants, I did not really know Dave prior to beginning this pilgrimage. Dave is a fellow Ph.D. student in the same program as I, but he started a couple of years later. I met Dave a few times over the years at doctoral program mixers, but we never really spoke to one another. However, I heard through several of my committee members Dave was expressing similar frustrations to mine with high-stakes assessment, which aligned with my research and might help me on my journey to my sacred destination. Thus, I reached out to Dave and asked him to take part.
I met up with Dave on a bright Saturday morning on campus. We sat in a little gray room for a long time as Dave shared his stories and experiences as a testing coordinator in a large school district in the area. A testing coordinator in Dave’s school district is a teacher employed to manage and oversee the school-based implementation of testing. The testing coordinator is responsible for planning, scheduling, and carrying out the implementation of high-stakes assessment. My hope is Dave’s narrative will shed light on the answers I seek on this journey. Dave started his career in education as a high school math teacher and then worked as an administrative resource teacher, which is a teacher with administrative responsibilities, such as student discipline, before becoming a testing coordinator. Dave served as a testing coordinator for two years, the first year at Silver Middle School and the second year, during the time of my interviews with him, he was at Bronze Middle School.

Although there were some small differences in titles between the districts, Dave’s narrative rang familiar to me. One of the differences between Dave’s narrative and the other participant’s is the frequency throughout his story I stop him to clarify. This was not because his school district and my school district’s context were very different, but rather because I was in disbelief and wanted to make sure I was hearing what I thought I was hearing. At the time of this writing, Dave has been promoted to an assistant principal at a high school in his school district.

Dave sips coffee as he begins to tell me his story: *My first year brand new to [testing coordinator] my principal didn’t know anything about it because it was a brand new administrative team. They were all brand new. They were still trying to figure out what meant to be principal let alone knowing everything behind it. Nobody knew anything about testing and then on top of it I didn't have a network manager to help with all technology stuff so it was an*
exciting first year. Then I decided to get little bit closer to home and also get with a very experienced administrative team. I was hoping that I could learn because at that point I was in the new assistant principal pool. I was looking to move up.

**Culture of Measurement**

Dave dives right into a narrative about the culture of measurement at Bronze Middle School. The thing...was my principal was really nervous as the past two years we were that school was about two - one point away for next letter grade up so she wanted everything to be administered perfect so if they weren't close to that school grade they couldn't blame it on what school didn't do so that she could say with confidence everything was given exactly as it was supposed to be given, Dave emphasized.

Wouldn't it be more about what the instruction was going on in the class? I suggest.

You would think.

I'm just trying to understand.

I know it's hard to explain unless you're there, and it was like the culture of the school. Like for instance, my previous year I did one day of training for FSA Writing right before the writing test and wasn't a full day. It was planning periods where teachers will come in during their planning period, and I would give them a 50-minute instruction on what to do. Then we got closer to the other half the administration I'd pull them in and give a refresher during a planning period. This year I can't even tell you how many hours it was. I did a full eight-hour professional development day of training on the test. For all teachers they were required to do it a full day--

* A full day? I shockingly interrupt.
of training on security, like my test security, like how to maintain test security, going over the manual to the point we were highlighting, telling them what to highlight in their manual, like, what words to highlight of the manual and the scripts.

I am so sorry, this is just beyond, I can't even understand that... I again interrupt.

Dave continued, We did a breakout session of how to organize the room. What the administration was going to be looking for because they did room checks before the test administration to make sure all the rooms were prepped. They would tell them exactly how to set the room, give an example, and make them do it like an example classroom. Another breakout session was how to give accommodations: what's allowed and what's not allowed for oral presentation, which was ridiculous because most all the tests were on the computer. If you have an oral presentation the text-to-speech function will take care of it all. The teacher didn't really have to do anything but we went into a lot of what if this happens but the power goes out then they have to read. Before we started math and reading they made me do a whole other planning period day of refresher and then made me do separate training for civics and a separate training for algebra and a separate training for geometry so I trained more than I've ever trained in my life and these teachers were so--

I intrude on Dave’s story again, What is the rationale for having a different testing training for every test?

They felt like I needed to go every over every single test script like with the highlighter with the teachers.

The script?

The script.

The scripts are the same.
So the teachers knew exactly what they were supposed to do and if they had a problem they can't come back and say they weren't trained. That's what [the teacher] did or what they didn't do. I did the full day of training plus a planning period day of training plus three other trainings, and then not only that. I then had to train all my proctors separately who were paras that weren't necessarily test administrators. They had to be trained separately for every single test as well. They won't let me just do one general ‘this is what a proctor does’ they made me train them every time. I was trying to get ready for a test while training people at the same time.

Dave describes a school culture of measurement dominated by an administration focused on the perfect administration of high-stakes assessments. A culture of measurement is characterized by a focus on the assessment students take versus the instruction they receive daily or, in other words, “the overemphasis on the process…and the test-taking competence” (Padilla, 2005, p. 250). Dave’s story demonstrates his principal is less concerned about the day-to-day teaching and learning and much more about teachers properly administering the tests to the students – the process of testing and test-administration competence. Dave was resigned to deal with his principal’s warped ethic of justice, which is focused on following the rules.

I try [to make things less stressful for the teachers], like the script I mean what's more boring than having to highlight a script like they're making us do. So I would try to interject a little humor like, ‘what do you all think we've highlighted the same thing in three of the other scripts so do you think we are going to highlight here? Oh you’re so brilliant.’

During the training where I was like frustrated because I'm like why are we taking my time to do this again for the 15th time when the teachers are grumbling, and I'm grumbling on the inside. I'm like somebody has to be the professional calm presence to get us through this reassurance sort of thing. Even though I was frustrated I put on I'm happy face to be here trying
to be a little bit perkier than I am normally face. I'm an introvert, of course, by nature so my natural tendency is just to be not very verbal and to be kind of monotone. I was forcing myself to be upbeat and perky for lack of a better word especially in those trainings. On the other side of it I had to be a calming presence. I was a wreck like the morning of the first day of an administration for a test when we'd switch whatever grade-level we were testing. Teachers tended to be more nervous that first day than the second day because they have more experience by the second day, and they would come in all stressed I'd have to be make sure that I was calm, and that when I was talking to them I wasn't being short because I was feeling rushed, like feeling anxiety myself, and I have to make sure this is done that's done or that class is covered, or that class moved, or that sub knows to go here versus over there. That sort of thing. I was afraid they were going to feed off of it, and they were already feeling that way themselves. I just didn't want them to feel anything for me to make their situation worse.

In his role as a testing coordinator, Dave had to perform as though he was calm and relaxed even when he felt stressed and anxious. Dave’s behavior is an example of emotional labor, whereby an employee’s job requires them to pretend to feel positive and upbeat when they actually have negative feelings (Hochschild, 1983). Dave’s performance includes making light of unnecessary trainings in front of teachers, pretending to be calm so teachers feel less anxious, and refusing to let his principal see him flustered. Hanley (2015) argues neoliberal policies within education, such as the implementation of high-stakes assessment, and the pressure educational leaders feel when these reform policies are put into practice “create a demand for emotional performance” (p. 105). Dave’s job exists solely because of neoliberal education policies. Dave’s job is to enact neoliberal education policies. Yet, to carry out this job Dave needed to regularly put his emotions aside and perform the role of testing coordinator.
The other thing for me also was my principal said a couple times throughout the year that testing was her baby and she was gonna push me to my breaking point that I was gonna break down before the year was over, and I refused to let it show that it was getting to me which frustrated her because she'd be like, 'I'm really nervous about testing because I don't see you being stressed about it.' I said, 'well outward appearances are different than what I'm feeling on the inside.' I said, ‘Of course I'm feeling anxiety about making sure that it's done and it's done right. There comes a lot of work with it. All that anxiety that comes with it.’ I'm like, ‘but how you respond. It's that's not necessarily how you should respond by feeling that way doesn't mean you should respond that way. I said, ‘and I want to maintain a calm presence.’ It did it pissed her off because she felt like she wasn't stressing me out enough like I hadn't reached my breaking point.

So was she trying to do more things to get you to the breaking point?

Yep, she kept adding stuff. Yeah, adding stuff to try to break me. I mean I wasn't disrespectful to her at all, but try to be professional, but I'm like I'm not gonna be stressed and be outwardly stressed. I don't think it helps anybody or helps the situation. We'd be sitting in the meetings going, ‘what still needs done? what's not done?’ and they start writing on their whiteboard. I would work systematically I've got to do this before I can do that before I can do that so I'm not gonna think about Y, Z if I haven't done X yet. Right? Well, they'd be like feeding off each other like you got do X you got to do Y you got a do Z and now you got to do one two and three after that, and what about four or five and six? They did get all worked up. I'm like I'm not even thinking about that. I gotta get this done before I can do that and that would aggravate the hell out of them. It was almost like, ‘if you're not stressed, we haven't done our
job. You’re not worried enough about my school grade, about how this is gonna go.’ Almost questioning my integrity and my ethics.

The culture of measurement permeated Bronze Middle School so much Dave’s ethics were questioned because he did not outwardly display the level of concern about high-stakes assessment preparation his administration felt he should be. According to Biesta (2017) in a culture of measurement educational professionalism is “distorted” by “systems of inspection and control” ensuring the professional complies with predetermined standards (p. 321). For Dave, his lack of visible stress suggested he was not complying with his administration’s expectations. The feeling of being inspected and not meeting standards was common throughout Dave’s stories.

I was definitely feeling more pressure than I ever felt. Like I mentioned a couple of times, I felt like I was constantly being watched, and I was gonna be written up for the most minor thing happening because nobody’s perfect you’re gonna make a mistake on something. I felt like I had to be perfect all the time.

According to Rowold and Schlotz (2009), the transactional leader typically provides only negative feedback, which adds to the employee’s stress level, but this behavior beyond transactional. This leadership style may not be effectively defined as transactional but instead, Dave’s principal’s leadership style might be better defined as toxic. Lipman-Blumen (2005) considers a leader toxic when she imposes grave and long-term maltreatment on her employees by using power strategies, which are exceptionally punitive and spiteful. Dave’s principal was a dysfunctional leader who used high-stakes assessment as a tool to embody “abusive supervision” (Tepper, 2000). Dave’s principal’s need to break Dave is indicative of her “indifference” to her employee’s well being (Jha & Jha, 2015). Indifference can be described best as, “taking
purposeful action to cause emotional, social, or psychological hurt as a means to elicit desired levels of performance” (Jha & Jha, 2015). To counter his principal’s toxic behavior, Dave’s negative thoughts and private feelings were suppressed (Tracy, 2000). This leadership style coupled with a federal and state mandated high-stakes assessment resulted in a culture of measurement.

As I listened to Dave’s story, I became grateful I had never worked in such a toxic work environment. I, too, work in a culture of measurement, but never for a toxic leader. I have worked for transactional leaders focused on maintaining high levels of performance for students on high-stakes assessments, but I never once worried I would get written up for messing up the implementation of high-stakes assessments.

The shooting at Stoneman-Douglas impacted the students’ response to testing when they staged a walkout at Dave’s school during a test administration:

_We did have a walkout. It was one of those it was rumored it was going to happen but it was mostly coming from of the high schools. We decided since the test that day was our 6th graders and they're not talking about at all. We're asking teachers if they have heard anything about it everybody was like, ‘No, no, no, no.’ We decided to go ahead and proceed with giving the test because there wasn't much room in the schedule to adjust and just not give tests that day. Right in the middle of the test at 10:00 in the morning students just decided to get up and they walked right out in the middle of the test. At that point, I didn't know who we were going to have to invalidate. The whole grade? If we would have to invalidate a whole room? I do not know what we were going to do so I contacted the district. What they wanted me to do was track who talked to who, who had access to a cellphone, which none of those kids had access cell phones because we collected them all in advance and they stayed with the test administrator the entire_
time. Basically at the end of the day after consulting the district, who consulted with the state, they determined it was no different than like the fire alarm had gone off in the middle of the test. They decided they weren’t going to invalidate the test so everything was good, but it required hours and hours and hours of work for me just to figure out: who went where, who walked out, who stayed, who had possession of cellphones, who didn’t, which everybody did not have access, but it took time to gather all that information and figure out. It was a nightmare for me at the time like I was pretty beyond stressed thinking I’m gonna get written up I’m gonna get fired for this not that I had anything to do with it.

Dave’s story is peppered with terminology used in education during testing season. For example, Dave uses the terms invalidate, test security checklist, and scripts in his stories. All this terminology Dave uses comes from the Florida DOE’s (2018a) assessment guide. An invalidation occurs when there is a breach in test security. Merriam-Webster (2018) defines invalidate as “to deprive of effective or continued existence.” In this story, Dave wonders which student he will have to invalidate not which student’s test he might have to invalidate. As Au (2009b) argues, students become abstract widgets whose only value is a test score in a culture of measurement. No matter his intentions, Dave refers to a student as an “invalidation,” which suggests he has begun to think of his students in the abstract.

**Testing Season**

Dave’s story turns toward his administration actions during testing season, *I mean my biggest frustration comparing first year with the second year [as a testing coordinator] was I feel like the administrative team at my second school was very, very knowledgeable about testing overall, but they made it a lot more difficult for me and for the teachers because they wanted*
everything to be absolutely perfect. It was to the point where the morning of testing teachers were actually a nervous wreck when they picked up their testing materials.

When you say perfect what do you mean? I ask.

In my first year as a testing coordinator, I did the state-mandated security checklist. I mean I didn't do the one that gave you in the [state provided testing manuals] because it doesn't have enough detail on it. But I'd make my own and I would have what accommodations were offered for instance versus what accommodations used. Well, in my new school they wanted this legal side sheet of paper and it had like exactly what was written in their IEP. So I would have to go pull every single IEP for every single student and put that into an excel file. As you can imagine the IEP changed throughout the year so it was a nightmare trying to keep track of those changes. I mean I was keeping up the changes all the way until the morning of testing. I had one giant sheet filled with every possible accommodation, right? I've actually typed in their accommodation likes, Johnny needs 60 extra minutes format, only whatever accommodation they have in [the IEP], they wanted it in there where the state didn't really require you'd have specifics. [The state] just want to know, do they need extra time? did they need a flexible presentation? [My administration] wanted me to actually put in the actual checklist. It went to the teacher. Then that teacher had something that was a lot more detailed than they ever need to know about the student.

Further, Dave accepts the extreme ethic of justice his administration exemplifies by attempting to comply with what is perceived as a perfect test administration. Dave having to document every aspect of each student’s testing accommodations should be unnecessary, assuming students are testing with their everyday classroom teachers, because the teachers should already be familiar with their students’ given accommodations. IEP or 504 testing
accommodations should be consistent between the daily classroom environment and high-stakes assessments (Bottsford-Miller, Thurlow, Stout, & Quenemoen, 2006). Yet, research suggests schools are not consistent with providing students with IEPs or 504s accommodations in the classroom or on assessments and most certainly not across both settings (Lazarus, Thompson, & Thurlow, 2006). Perhaps, Dave’s task was in the best interest of students in spite of his administration’s intent of compliance. The juxtaposition between the variability of IEPs and the compliance and standardization of high-stakes assessment is striking in Dave’s story:

How many students in the school? I question.

There was 1226.

How many ESE students?

I remember there was a little over 200 about a third of the school had either an IEP or 504.

So you’re keeping track of 504 as well IEP accommodations, okay?

That was kind of surprising to me when I transferred from my last school to this school. I figured higher socioeconomic-status is probably not as much labeling but the more I got into it more I realized that’s more the parents knowing what their rights are, and how to get my kid an IEP was about power and influence, and getting their kids on this IEP so they could either get waivers or special accommodations. It was a ton of work like I think I worked every single Saturday from it was basically in January till the end of the school year.

So for all the testing season? I clarify.

For all the testing, Dave confirms. The ESE Specialist was new, and my administration told me not to depend on her. I had to look at every IEP. I mean we had reports we could pull, but, of course, sometimes it would miss an accommodation like if the person didn't check off that
it was a testing accommodation when they were doing the IEP when I pulled present accommodations they wouldn't pull. I had to pull every single IEP or 504 and read them to make sure I wasn't missing anything.

I find interesting from this is that they were telling you not to depend on that person, but they were a hundred percent depending on you so you if you messed up it would still be on them, I conclude.

I was just terrified to be written up. I was terrified of a kid not getting their accommodation because I missed the change. I mean I doubled check that triple check like it created more work for me because I was paranoid that I was missing something, Dave added.

Obviously, you know students with accommodations need to get their accommodations and I am a proponent of a student getting their accommodations, but I have a question about what does that mean about the administration's trust of teachers and teacher ability? I inquired

I mean my first school in my experience had me create per class one eight and a half by eleven sheet of paper [listing all of the students’ in the classes accommodations]. This school it was a full legal-sized piece of paper with writing so small because we had to put so much information into it. It was hard for me to balance [how I feel about high-stakes assessment and being a testing coordinator] because I'm like it is my job, and my job is this because of testing. On the other hand, I felt conflicted like we're putting teachers through the stress, we're putting students through the stress, parents were stressed. I had parents calling me directly saying, 'little Eliza doesn't feel good in the afternoon. She can't test the afternoon, and she needs the test in the morning,' and they were trying to get me to make changes. I'm like, ‘well that's the only day.’ That doesn't mean give in ok, and I really do not have the option or enough resources to schedule every kid according to their personal testing preferences.
First, the push and pull of the tensions between compliance with standardized testing regulations and the variability of individual student needs as outlined in an IEP are apparent in Dave’s narrative. The goals of federal and state mandated high-stakes assessments conflict with the goals of federally mandated IEPs (Sackel, 2006). On the one hand, all students are required to take high-stakes assessments, which measures how much a given student has mastered reading or math standards established by the state for all students in each grade level. On the other hand, students with disabilities have Individual Education Plans, which are distinct documents outlining accommodations for each student with a disability in a variety of educational settings including the high-stakes assessment setting. This contradiction in federal education policy plays out in Dave’s story. His school attempts to reconcile this conflict by listing out each accommodation students receive for teachers to review during each high-stakes assessment administration.

Second, in this story, Dave portrays his school leadership’s “bureaucratic orientation” with centralized control over policy implementation (Tschannen-Moran, 2009) and the environment of “threat rigidity” (Olsen & Sexton, 2009) he worked within. Tschannen-Moran (2009) defines this type of leadership as authoritative, with the administrators using “compliance with organizational directives” to maintain disciplinary control over teachers and “without close supervision” teachers will not carry out their responsibilities or perform their duties. (p. 221). This leadership style also, “serves to reinforce an implicit distrust” of teachers (p. 221). Throughout Dave’s narrative, his school administration, particularly the principal, micromanaged every aspect of high-stakes assessment implementation in various ways. Indeed, Dave then tells another story of the ongoing control his administration maintained over him:
My administration wanted everything perfect like in October they were making me not just come up with when are we giving the test, a draft schedule, they're making me go down to this granular level as what teachers gonna test in what room. Like having basically having testing that was planned out and of course lots changed between October and test time. I constantly had to go back and change things I just felt like it created so much more work for me, Dave recounted.

It sounds like you got a lot less autonomy this past year than you did the previous year so and that's the site-based administration not district administration? I suggest.

District administration wasn't involved. I was site based myself. Only if I had a testing issue, like a kid, for instance, had they submitted their tests early before the requirement. All of our students waited until they were told at the end of timing to submit their tests and just sit there and recheck their work or whatnot. If I had a kid submit when they got finished they submitted it instead of following the directions and waiting, my administration would make me email the district to get written permission to make sure that was okay. Where I knew it was okay before I ever contacted the district administration. Basically, it was like if he read the script everybody was told to do option A well that student just did option C it’s not against the testing rules for that kid to submit earlier than what they told they were supposed to. I think that part of the issue was I knew a lot about testing but they hadn't done it themselves for so long that they were reverting back constantly to old rules. So not understanding the changes that have happened since they were more deeply involved in the testing. And they didn't want to be educated they just wanted black and white black and white, okay? So you know I would have a conversation with a district person the district test administrator saying, ‘hey’ and they probably think I know less my second year than my first year. Then she's like, ‘I understand that school.’ I mean that
was basically she understood that they were doing more of the cya and that necessarily wasn't coming from me, okay? But like as far as like a detail when I was talking about the changes or to have to keep up with IEPs and how much detail there was in the security checklist, and I would get those back from teachers at the end of the test and the teachers would have like volumes of information written on the security checklist like the student fell asleep from 1:05 to 1:06, this student went to the bathroom. I mean like very very detailed any time a student sneezed or anything it was like this is incredible that they were this nervous. The cya mentality just rolled downhill right from the administration.

Indeed, this bureaucratic style of leadership thrives in an environment of threat rigidity (Olsen & Sexton, 2009). Threat rigidity is an organizational state, which occurs when the organization is under attack from the outside, whereby the organization responds in “identifiable ways” (Olsen & Sexton, 2009, p. 15). In this instance, the community and school district have Dave’s school under scrutiny because the school missed increasing the school grade by just a few points the prior year. The identifiable ways the organization responds to the threat include tightening control over subordinates, pressuring subordinates to conform, individualism is frowned upon, and the adoption of “accountability and efficiency” methods (Olsen & Sexton, 2009, p. 15). Dave’s description of his administration’s bureaucratic orientation within an organization under threat includes pressures for teachers to conform, tight control over teachers, no room for teacher individualism, and high levels of accountability. Olsen and Sexton (2009) suggest the subordinates within the organization often have high levels of work-related stress and an impeded ability to think creatively, which lead to feelings of not being valuable to the organization. Dave described the stress he, the other teachers, the students, and the parents feel in his story:
Every morning was like, ‘where you at? where you at?’ We’d have to come together at least once a week to ask, ‘where you're at?’ because they would get nervous not knowing what was going on and they would want like down to the detail of like who's given the test. It was mostly that the principal and the assistant principal for curriculum. At least once a week at least once a week to talk about what my plans were, and I'll come up with all these plans. My [senior assistant principal] would be like, ‘well I think this is better,’ and she make a switch, and now I got to go back update this list and that list. It would just create like a domino effect of stuff I had to do. They also would make me come up with all these contingency plans like, ‘if this person calls out who would you have them replace them?’ or, ‘if they were in the room how who would cover their class?’ like down to I would have to figure out who would be the sub for the sub. Not even related to testing. Then we didn't want any room close to the testing environment with students in it because they may disrupt the testing environment so I would have to relocate classes with students who weren't even involved with testing and figure out where they were going to go. When you're in the middle of testing and every morning I would have think about testing materials and teachers, I was stuck I couldn't leave the room unattended. I would have people calling out it was very difficult for me to enact things or let people know that they need to do here or there when I'm in the middle of trying to hand out stuff to the teachers. The other thing too is the APC would always like make last-minute changes like, ‘I think it’d be better for this person to be the test administrator for this person.’ ‘why?’ ‘well because they let two students go the bathroom at one time and almost caused an invalidation.’ It was almost like that the teacher made any questionable moves, not really mistakes, if they just felt like the teacher was not 100% on their game they would just switch them or take that responsibility away from
them without even asking them about it without even coaching them on it. There was no trust that that person could do their job from that point on.

*Sounds exhausting.* I say sounding sympathetically exhausted myself.

This style of toxic bureaucratic school leadership has its roots in Neo-Taylorism, whereby the worker is merely meant to carry out the manager’s orders. “In [Taylor’s] system the judgment of the individual…was replaced by the laws, rules, principles, etc., of the science of the job which was developed by management (Callahan, 1962, p. 27). Neo-Taylorism takes the managerial control and efficiency aspects of scientific management and combines them with notions of capitalism and market-driven forces as well as systems of rewards and punishments (Terry, 1998). Dave’s principal’s tight control exerted over Dave, and her desire for the perfect, most efficiently run administration of high-stakes assessments combined with the omnipresent threat of writing faculty up for the smallest infraction exemplifies the Neo-Taylorist aspects of her toxic bureaucratic leadership style, except Dave’s stories do not indicate his principal rewarded anyone. Thus, the pressure the principal at Bronze Middle School felt and perpetuated on her staff was due in part to the school grade, which is a market-driven force that creates completion between schools. Thus, to be more competitive Dave’s principal enacted methods of control.

I take a deep breath and surmise, *It sounds like almost like your administration had ruled from fear and with fear. I’m sure was in other things too but in relationship to testing certainly that’s a lot of stress so it was probably a lot more stressful this year.*

*Yes, I got called to a room, I mean it wasn’t just me it was the teachers as well, I got called to a room and a kid had finished her test had already submitted that kid had gotten out of the test browser and they were playing games that weren’t even online but the teacher was*
freaking out she's going to lose her license. The kid's playing games on the computer but he was already submitted and didn't affect anybody else. Just sitting there, right, and the teacher was worried they're gonna lose a license because they were instructed not to do anything except sit there. But they were done. I kind of felt like at the end of the day we were losing common sense and our ability to make decisions trying to follow the letter of the law exactly as it was written or told or perceived by administration.

The tight control Dave’s administration maintained over testing seemed like overkill. The micromanaging of every aspect of the test administration itself was aligned with the controlling aspects of an organization in a state of threat rigidity. The focus on compliance and conformance deprofessionalizes educators and suggests they are incapable of making any decisions (Ingersoll, 2009). Dave especially felt micromanaged during testing season.

My first year testing when teachers were done they would buzz the office and say, ‘I'm done,’ and they’d call us and say, ‘so-and-so is done,’ and we would take turns going and picking up a test materials. Now I had to have pick up teams. They wouldn't allow me just to go pick it up if a teacher was done and they were the only one done. I had to do a color-coded map of the different people available to pick up bins from teachers once they were done to verify material at the room. It was verified at the room, and the bins were not allowed to leave that room unless it was verified hundred percent ready to go. If you signed it and it made its way down to my room and I found the mistake when I was verifying it on my end then that person would be written up. For example, we had a Dean who usually does discipline, and she was on the pick up team and she picked up a bin. I think the teacher forgot one signature on something that could have been very easily taken back, ‘hey sign this.’ She was punished she wasn't written
up she was punished like she had to label stuff for like a day that was her punishment. I bet you'll never do that again.

The actions of the principal of Bronze Middle School as depicted by Dave are those of a toxic leader (Lipman-Blumen, 2005). The toxic leader is like the bureaucratic leader in that she manages subordinates using rewards and punishments to maintain control, and the toxic leader focuses on monitoring the details and the carrying out of day-to-day regiments of their employees without the rewards. As noted earlier, Dave’s stories do not suggest his principal had any system of rewards, but his stories frequently include punishments.

It definitely makes me more stressed at home and I tried to keep it separated from home as much as possible but I know like I was short with my wife and my kids. My in-laws live with us too so they were getting on my nerves even more than normal.

That's another interview, I joke.

The thing is when I got home they had expectations like my father-in-law is elderly. If he has something to do around the house, he can't do a lot on his own, they were expecting me to stop what I was doing. Just because it's after hours doesn't mean I don't have things I need to do for my training I have to do tomorrow or make sure that I've double-checked all the accommodations are turned on properly in the test platform for the student or stuff along those lines so kind of led to a little more frustration to my personal life. During testing season if I had a class on Saturday then I did work on Sunday, and there were some weekends where I work Saturday and Sunday, but my thing was is they weren't gonna get me on anything because I was really honestly afraid I would be written up.

Research suggests educator stress is higher when the principal has a transactional leadership style (Eyal & Roth, 2011). One can only imagine the stress educators, like Dave, feel
when their principal has a toxic leadership style. It is not surprising to learn Dave took his work home, or that he put in extra hours. According to the NEA (2018), educators often work at night and on weekends to keep on top of their jobs. The role of a testing coordinator is much like that of a project manager as testing season can be likened to project management because it is a time-bound, temporary, scheduled, resource necessitated, “execution of activity,” leading to the production of a service (Schwindt, 2005). Dave’s job has some unique time-bound challenges he faced during testing season, but perhaps more importantly, Dave’s principal approached her overseeing of high-stakes testing as an opportunity to not only look for opportunities to provide negative feedback but also to be cruel to her staff.

**Ethics of Implementation**

*When I was a testing coordinator my first year you know there’s an ELL student, the English language learner, they’ve been in the U.S. less than a year there’s no reason for them to take the ELA portion of the [FSA] test, Dave explains.*

*Its up to the administration whether or not they take this test, I interject.*

Dave agrees, *Right they’re not required to take it but that's what I was strongly encouraged to do. You know to test these students it's not gonna count against us or for us but it gives them experience so for the next year. I had kids sitting in front of that computer just breaking down an in the middle of a test because they felt like they were complete failures. I had decided at that point I'll never be doing that again so this year when it came up it's like, ‘you’re going to do that right?’ ‘Yep,’ and because [the ELL first years] never magically made it on the testing assignment [my administration] didn't know they hadn’t tested. It wasn't so much my school administration it was the district administration that was telling us to do it.*

*The district administration was saying, ‘You need to give it?’*
Yeah, basically we would say it's a baseline right? Like you're creating a baseline for the ELL student so the next year they'll show growth and that'll give the school wherever they are the growth points right.

In unison both Dave and I exclaim, You're gaming me the system!

Dave nods, Exactly! But I mean on the flip side I knew the administration that I was working with this year I knew that there were rule-followers. The district said to do it we were gonna do it because we weren't gonna get caught and get in trouble. I never really asked for permission I just didn't do it and it never came up. I'm not having kids crying over a test that shouldn't count for them to begin with so that was one of my frustrations.

Every day I had to give an update on like what's our percentage tested look like. I understand getting to the 95 percent, but it was like every day until we hit that 95 percent and then once I had that 95 percent it was almost, it wasn't told to me explicitly, but it was implied we're gonna stop testing because we don't want those students who have been absent there anyway. We don't want their test scores. I ignored it I tested them anyway.

How much did discipline impact who was not tested? ‘oh this kid is gonna be suspended during this test.’ I think out loud. I remember an assistant principal I worked with many years ago purposely suspending certain students for ten days so the students would not be able to take the high-stakes assessment. I have only known one assistant principal to do this, but I suspect there were many more. It has been several years since I witnessed this practice, and I hope it has been eradicated from schools completely.

‘We'll make sure they're gone during it. They need to be gone.’ Was it strategic like who didn't do it, or was it just because they're absent a lot and you assume that because they had a discipline history? They weren't necessarily suspended at the time, but [administration] knew
based on their history in the previous year the school would be better off not testing that student. Now my first school I was at I was told to not the test students once we hit that percentage that 95 percent. Really? I'm just like, ‘I can't believe somebody asking me to do this.’

Coltrane (2002) argues that high-stakes assessments do not test the academic ability of non-English speaking students new to the United States but instead tests their language acquisition skills. All ELL students are assessed on the English language acquisition assessment used in Florida, referred to as WIDA. However, schools and districts have the discretion to test first year ELL students on the FSA ELA test, and the state will not count the students’ scores in the school’s grade calculation. Districts and schools benefit from having the students take the FSA ELA because it provides a baseline the school can demonstrate growth from against next year. The assumption is the student as second year ELLs will perform better on the FSA ELA the second year. The growth students demonstrate between the two years counts as gains and does impact school grade. This practice is something Dave finds unethical. He feels strongly enough about this he subverts the authority of his district and most likely his principal, if she knew, and resists testing his first year ELLs for the FSA ELA.

The desire for schools and districts to collect baseline data from first year ELLs, no matter how unethical it may seem, was viewed by both Dave and I as gaming the system (Figlio & Getzler, 2002). This practice is allowed not via a loophole left in the policy but was created by the policy itself. The state’s policy encourages schools and districts to collect the baseline from first year ELLs by not holding the students’ performance against the schools. The state testing policy allows for 5% of students to not test. The policy for testing 95% of students was created to prevent schools from only testing certain groups of kids while also taking into account it is almost impossible to test 100% of students. However, the 5% wiggle room can be taken
advantage of, as Dave points out, due to poor behavior or excessive absences low-performing students often make up this 5%. Dave’s administration attempted to exclude the remaining 5%, but again Dave subverted his administration’s implications and tested the students anyway. Dave found it unethical to leave 5% of the students out of the testing pool, and rather than comply with the expectations of his administration, Dave ignored the directive. Although Dave’s stories suggest he is resigned to compliance, Dave found spaces to push back.

The aspects [of high-stakes assessment] I disagree with most I mean the teachers, more so this last school but even my other school, I felt teachers were stressed during that time okay I feel like there's so much accountability on them to show results. ‘Hey if Johnny just put his head down on this test.’ then I can see that stresses them out. They don't like that student's scores hurting their data. For me, I have a problem with you judge that teacher based on one day on one test you gotta look at the whole picture you got a look at their observations their evaluations what's the principal saying about their teaching when they go into that room.

As Berliner and Glass (2014) posit, a single day of testing is not an accurate measure of a student or a teacher. One day of testing says little about any students’ accomplishments or the teacher’s work in helping students achieve those accomplishments (Berliner & Glass, 2014). Although teachers in most of Florida are only evaluated based on principal observation and yearly evaluations, student test scores on high-stakes assessments are unofficially, and unethically, used as a measure of teacher effectiveness and to inform the principal’s final evaluation of teachers.

I think my whole viewpoint, especially as a testing coordinator, I always felt like my responsibility was to make it as painless as possible for the teachers. I think that's how I justified doing testing overall without really believing in it, but if I can be that person that makes
it as easy as possible and stress-free as possible for teachers and students then that's what I'm gonna do as the testing coordinator. If not me then who else? Maybe that's how I rationally justify it. If it isn’t at me it's gonna be somebody else, and they're not gonna at least find little places to resist. I think that's why my outlook has been on teachers a lot like this past year because I feel like some of the stress especially this year was completely unnecessary. They tell me they’re gonna lose their certificate over minor stuff. I mean accidents are gonna happen I've never heard anyone losing her certificate over something that happened as an accident. You intentionally lie your way to help a student or cheat? Yeah, you lose your license, but not over accidentally missed timing a student by five minutes. Yeah, they will invalidate that student’s test but you’re not gonna lose your job. But on the other hand, too maybe it wasn't just that losing a certificate it was, ‘am I gonna get written up by the principal because I saw that guidance counselor getting written up because they missed a signature on something.’ I think the culture of school played a lot into how teachers and administrators handled their outlook on testing.

As Dave’s narrative suggests, teachers feel a great deal of stress during test administrations, mainly due to pressures from school leadership (von der Embse et al., 2015). Dave feels ethically compelled to be the person to ease the stress felt by teachers during testing. He points out if he was not the one to do his job it is possible the person doing it instead may not care as much or at all about teachers and students. Dave notes how fear appeals from teachers and parents impacted students:

Kids would get physically sick. I had more kids to be physically ill during the tests who’d have to leave when they come back later on in the day and finish it when they fell better but my administration’s flipping out making me email to get permission to make sure that kid and sit for
the test we don't have to invalidate them when I knew they were gonna be allowed so they didn't trust my ability to make decisions that's what was frustrating.

Research suggests fear appeals to students have adverse effects on students’ test performance (Putwain & Remedios, 2014; Putwain & Symes, 2011; Symes, Putwain, & Remedios, 2015). Test anxiety fueled by fear appeals can result in physical reactions from students, and reports of students exhibiting physical reactions such vomiting, illness, and vomiting during test administrations have been widespread (Koretz, 2018). Yet, the focus is not on the child but rather on completing the test. If students begin a test they must finish it the same day or the test is invalidated. This counts against the school grade and the student is often pushed into remedial classes if the student’s test is invalidated. Dave’s concerned for the student’s well being is contrasted by his administration’s concerns about the student completing the high-stakes assessment. “The reduction of people to databases and the constant effort devoted to fabrication empties institutions of authentic practices and relationships” (Anderson, 2009). Again, Dave’s administration’s ethics are questionable with the importance of student performance on high-stakes assessments above student safety. Additionally, Dave’s story further highlights the lack of trust the school administration had in his ability as an educational leader in his school to carry out simple tasks and to have adequate knowledge about the state’s rules and policies guiding high-stakes assessment.

Conclusion

On this stop on my journey, Dave’s narratives outline an unethical autocratic leader focused solely on the performance of students and teachers on high-stakes assessments, too often at the expensive of the people in her care. Dave enacted his own ethical beliefs by working to find small spaces to resist high-stakes assessment policies he disagreed with through subverting
district and school-level authority, which is what I have continued to search for throughout my journey. Dave’s testing season was tightly controlled by his administration, yet Dave chose to be a calm and reassuring presence in spite of the circumstances. The stories Dave shared with me about his experiences implementing high-stakes assessments at Bronze Middle School were, at times, disturbing. Dave’s narrative exemplifies how high-stakes assessments policies can lead to perverse and misguided interpretation at the implementation level. Next, my travels lead me to consider the collective narrative findings.
CHAPTER 8: TALES OF MY REFLECTIONS ON THE NARRATIVE COLLECTIVE

“Seasons change with their scenery
Weaving time in a tapestry
Won’t you stop to remember me” (Simon, 1966).

**********

As my journey is drawing to a close I am reflecting upon what I have learned from my participants’ narratives, including my own. I was expecting to hear stories about students’ being stressed out and over-assessed on high-stakes assessments and the ethical tensions we, as school administrators, feel when we implement these assessments at the school level. I was right, I did hear those stories on my journey, but I learned about ethical dilemmas I would have never considered prior to undertaking this study, such as the use and misuse of school capital and human resources. On this stop of my journey, I will discuss what the collective narratives in this study reveal about the ethical tensions associated with implementing high-stakes assessment and the contradictions felt by administrators who are implicated advocates (Carpenter & Brewer, 2012).

Purpose

The purpose of this journey was to explore the stories of the ethical tensions K-12 educational administrators navigate when implementing high-stakes assessment policy in a culture of measurement during testing season. Some educational leaders, in particular K-12 school and district administrators, struggle with the tensions existing between their own personal...
belief systems, organizational dilemmas, and the requirements of enacting high-stakes assessment policies.

**Research Questions**

My research questions are: 1) In what ways do K-12 educational administrators navigate potential ethical dilemmas between personal ethical obligations to students versus organizational directives associated with the implementation of high-stakes assessment policy? 2) How do K-12 educational administrators experience and live the testing season? 3) How does the culture of measurement play out in the stories of K-12 educational administrators?

**The Implicated Advocate**

The implicated advocate is defined by Carpenter and Brewer (2014) as a public school leader who is at once implementer of neo-liberal policies of high-stakes accountability, and, at the same time, the guardian entrusted to be the advocate of students. In many instances throughout my participants’ narratives, they describe situations where they are implicated advocates. In this chapter, as I focus on the collective narrative of my participants and commonalities among their stories, I address their roles as implicated advocates. Aspects of the role of the implicated advocate fit within several ethical paradigms. The implicated aspect aligns to the ethic of justice. The advocate aspect aligns with the ethics of care and critique and social justice advocacy. Yet, the contradictions outlined by Carpenter and Brewer (2014) align nicely with the contradictions outlined in the ethic of the profession. Educational leaders are at once called upon to be social justice leaders and at the same time the implementers of high-stakes assessment policies.
Culture of Measurement

Accountability

“Accountability is thus not to the state but rather to a citizen public and the communities that schools serve” (Valenzuela, 2005, p.12).

Echoing throughout all of my participants’ narratives is the unequivocal agreement there needs to be an assessment of student learning and teacher effectiveness. There must be some type of accountability was something they all said. Even when I suggested our jobs, as administrators, were to hold teachers accountable via classroom walkthroughs and coaching cycles, they all resisted this notion. This leaves me wondering if the ongoing repetitive political narrative about holding teachers and schools accountable, the politics of high-stakes, standardized assessments, has poisoned the minds of the educators, too? We are, as Lipman (2004) suggests, assimilating to the culture of measurement. The public has ignored the testimonies of psychometricians (Linn, 2000) warning of misusing high-stakes assessment outcomes, and even those of us on the front lines seemed to have fallen for the promise of high-stakes assessments as the panacea of teacher ineffectiveness. Yet, since the initiation of NCLB has there really been any proof? At this point in my doctoral journey, I feel as though I am an expert on high-stakes assessment, and I am in my 12th year as a practitioner in the public school setting, but I have never seen any high-stakes assessment results clearly pinpointing teacher effectiveness. Indeed, the educational leaders in the schoolhouses are more equipped to determine teacher effectiveness than a test possibly can. We know which teachers struggle, which ones need more support, and which ones shine, why do we need a test to tell us that? All poor test scores do is drive teachers away from high-poverty schools, avoid teaching vulnerable
populations, and misplace blame for poor assessment achievement on teachers and students and their families.

What could this phenomena, school administrators entrenched in the idea there must be some accountability via high-stakes assessments, be embedded in? I would argue the roots of public school accountability in the United States are deeply seeded in the perception teachers are not professionals and must be monitored by outside, more knowledgeable, forces (Bietsa, 2017). In fact, since the inception of compulsory public school, there have been outside forces working to control teachers, the majority of whom have traditionally been women (Ingersoll, 2004). Thus, using a supposed objective measurement to manage what and how teachers teach is a means of control.

Prior to national accountability via NCLB, Cuban (1990) argued historically school reforms, “largely [left] schools and classrooms unchanged” (p. 10). In light of the culture of measurement, the ethics of implementation, and day-to-day struggles my participants felt during testing season, it is hard to agree with Cuban (1990), yet one must ask: have these changes improved the lives of students who live in poverty or students of color? Indeed research suggests the achievement gap has persisted in spite of NCLB and its successors (Cohen et al., 2009; Haycock, 2001; Ladson-Billings, 2006; Lee, 2002; Reardon, 2011, 2013; Williams, 1996). I would argue that things have changed but not for the better.

All four of my participants were adamant high-stakes assessments was necessary for holding teachers accountable, yet they also felt a great deal of compassion for teachers. Indeed, Dave, Nora, Michelle, and Cher all discussed how hard high-stakes assessments are on teachers. For example, Dave described minimizing the stress teachers feel. In this way they, are implicated advocates.
High-Stakes Assessment is Big Business

High-stakes assessment is big business for companies like Pearson and AIR. Burch (2006) found the largest categories of educational contracting were test preparation and development as well as data reporting and management. Contracting a company to create and grade the high-stakes assessments and then a company to manage all of the data related to the high-stakes assessments, and report the data is done at the state and local levels. Burch (2006) suggests the outside contacting of educational services such as these are quiet avenues of privatization that get very little press. Yet, all of my participants mentioned the connections between high-stakes assessments, politics, and money. Perhaps, those of us at the implementation level feel as though we can see the connections to money.

Narrowing of the Curriculum and Loss of Instructional Time

Narrowing the curriculum is defined by Padilla (2005) as instruction focused on basic reading, writing, math, and test-taking skills. I argue in addition to students experiencing the narrowing of curriculum they are also losing a great deal of instructional time due to high-stakes assessment. Any student loss of instruction due to high-stakes assessments is the loss of instructional time. Dave shared how teachers spent multiple days out of their classrooms as he trained them on the testing scripts. Every time a teacher is pulled out of their classroom a substitute takes over for them. Typically, a teacher leaves plans for substitutes to keep the students busy. The content of these plans is not normally challenging or new learning for students. Thus, these days away from students narrow the instructional time for students. Nora, Michelle, and Cher all describe how the excessive computer-based high-stakes assessments remove students from teachers and classes for multiple days. Students who are not pulled from
class are left behind with a partial class where the teacher may or may not move forward with instruction since many students are testing.

Moreover, Michelle and Cher describe how students with disabilities lose instruction because their teachers are one-on-one testing with other students leaving the remainder of the class with paraprofessionals. Cher and Nora also describe this experience occurring for students enrolled in ESOL programs. In each case, students were denied a teacher for multiple days due to high-stakes assessment. This is clearly a loss of instruction. I wonder if administrators have more options open to them when planning for high-stakes assessments for students who are ESE or ELL then are typically considered. It is easy to take resources away from the most vulnerable populations if the students are unaware or unable to communicate the disparities.

Since each of the four participants were instrumental in creating the conditions for students to experience the loss of instructional time because of CBT or one-on-one testing over the course testing season they are implicated advocates (Carpenter & Brewer, 2014). They each discussed how unfair this practice was and how they try to make the best of the situation, but at the same time, they were mostly responsible for the way their schools enacted high-stakes assessments. They each discussed feeling powerless over the amount of instructional time lost, but they each had some amount of power in the implementation, thus they are implicated advocates.

None of my participants considered the loss of instructional time for students or students being denied curriculum as positive. On face value, losing instruction due to high-stakes assessments is not a good thing. However, not all instruction and not all curriculum is in students’ best interests. Perhaps, some students are experiencing curriculum and instruction that is oppressive, in those cases the students are better off testing or playing Minute to Win it.
Labeling

Labeling students based on how they perform on high-stakes assessments did not necessarily come out in all of my participants’ stories, but in the language my participants used to describe students and schools as a whole based on how the majority of students attending the school performed on high-stakes assessments suggest labels embedded in the culture of measurement. Throughout the interviews, the use of deficit terms such as bubble kids, level ones, level twos, low-babies, low-performing students, struggling students, struggling schools, and low-performing schools are common terms used by my participants and, disappointingly, at times myself. Dave even describes a student as an “invalidation.” Michelle’s describes the pain her students feel because of their poor performances on high-stakes assessments has rendered them failures in her narrative. State assessment labels describing students as “inadequate” was a source of frustration for her. As Michelle stated, these labels are unfair and do not truly define who students are and what they are truly capable of doing. I argue that we need to do better ourselves. Educators do not need to adopt deficit language to describe students just because the state does. These deficit terms serve to perpetuate beliefs about what students are capable of academically and allow the culture of measurement to flourish.

Resegregation

The only narrative to address one of the most significant issues connected to high-stakes assessment, resegregation, was Michelle’s. Michelle recognized the relationship between high-stakes assessments and privatization and how Florida’s education policies are contributing to the resegregation of schools. Michelle’s argument was students’ scores on high-stakes assessments result in school grades, and schools with lower school grades are predominately in neighborhoods with high rates of poverty and people of color, which is bore out in the literature.
School grades present a one-dimensional view of schools, but more affluent White families choose where they live or choose to send their children to charter schools rather than have them attend a neighborhood school with a low school grade. Then charter schools cream the students who perform best on high-stakes assessments leaving the neighborhood public schools with students who have not performed well on high-stakes assessments, thus perpetuating the cycle.

**Experiencing the Testing Season**

It was interesting to see in follow-up interviews how much my participants distanced themselves from testing season once it was over. Almost in the same way the negative aspects of winter fade away once spring and then summer arrives. We still know we will have to shovel snow and scrape the ice off the windows of our cars, but that took place in the past and will not be back until the distant future. For now, we are happy to ignore the realities of winter when we think of it in June. In much the same way, once Cher, Michelle, Nora, and Dave were no longer in the thick of testing season, and they did not appear to be as concerned with the stresses of testing. They were aware of the realities occurring during testing season and could speak of them, but they did not have the same level of passion or frustration as they did when I interviewed them during testing season.

**School Administrator Performance**

Attempts to appeal to faculty and students as though they were content with the status quo of high-stakes assessment implementation were evident throughout my participants’ narratives. In reality, it was difficult or arduous as school leaders to perform in ways that contradicted how they felt. It was a major source of frustration for my participants, as evidenced by Cher’s mantra, “embrace the suck.”
Performing the implementation of high-stakes assessment. Nora, Cher, and Dave described how they performed in ways contradicting how they honestly felt during the testing season. Dave’s story of how he made redundant and unnecessary professional development bearable for teachers by pretending to be outgoing and funny, as well as his story of performing calmness in the presence of his principal and teachers, were both examples of emotional labor (Hochschild, 1983). Dave felt as though he had to perform in these ways during testing season because teachers were looking to him as a model for how they should feel and his principal wanted to terrorize him emotionally (Leonard, 2014). Further, Cher explained she used the term, “embrace the suck,” as a way to perform positivity during a stressful testing season. Cher’s narrative includes how she performed for students by thanking them for testing during the end of day announcements at her school. Additionally, Nora describes performing for her staff by putting a positive spin on the testing season. In contrast, Michelle denied performing for teachers and students during testing season.

In each of these cases, my participants acted out the role they felt were needed by those around them. Hanley (2015) suggests, an emotionally accomplished school administrator offers sensitive reactions in affirmative language when enacting neo-liberal policies, such as high-stakes assessments, requiring them to perform contrary to their own belief systems, which results in being implicated advocates (Carpenter & Brewer, 2014). I do not think any of my participants felt they could show their true feelings. As leaders, we are expected to implement policy and have others follow us in the implementation, and it is difficult to inspire followers to get a job done if we are completely honest about our feelings.

Emotional labor. Three of my four participants described faking feelings during the implementation of high-stakes assessments. As part of implementation, Nora, Cher, and Dave
put on happy faces and pretended to be happy about high-stakes assessments. Nora took pride in her ability to make something bad sound good. Dave felt he had to act as though he was extroverted during teacher trainings, and he had to act calm around his toxic boss because Dave did not want her to think she “broke” Dave. Cher’s mantra of “embrace the suck” is the epitome of emotional labor. This aspect of the implementation of high-stakes assessments is ripe for further scholarship.

**Compliance and agency.** Another aspect of experiencing the testing season, which came up throughout my pilgrimage, was the feeling of helplessness my participants described in our stories. They expressed frustration with the systems maintaining the juggernaut of high-stakes assessment, they were aware they were part of the machine but felt paralyzed to do anything about it. There are consequences for not testing, not completing testing by state deadlines, and for not securing the testing environments. The need to follow the rules fits within the ethic of justice, and these things impact the school grade and the individual teacher, administrator, students, or all.

Yet, there are always spaces within policy implementation for interpretation. The state has rules and guidelines, which are interpreted by district-level administration and filtered down to the school administrators who then further interrupt the guidelines and rules. No doubt the rules constrain the implementation, and there are pressures to complete testing within the rules and timeline with limited resources, but there are still spaces to for those at the implementation level to make some decisions and have some agency.

I suggest the amount of agency each participant’s narratives described is tied to their position in the school system’s hierarchy. Nora’s narrative did not seem as helpless as the other’s narratives, and this may be due to the fact she is a principal and has more agency. Nora
was frustrated by the amount of testing and the impact on her students, but she rarely came across completely helpless. Instead, I would characterize her more pragmatic. She identified what needed to be done to implement high stakes tests, and she did it. Michelle shared few feelings of helplessness. Michelle seemed to question the overarching policies and practices related to high-stakes assessment implementation more than the other participants’. I would chalk her agency up to the amount of freedom her principal gave her to implement high-stakes assessment.

On the other hand, both Dave and Cher’s stories expressed feelings of powerlessness over the implementation process. Dave’s principal micro-managed him as well as the implementation of high-stakes assessment, which may be the reason his stories are replete with examples of compliance. I think it is important to note, Dave is the only participant who was not technically a school administrator, but a teacher. Thus, his role in the school’s hierarchy left him with less power than an assistant principal or a principal, which resulted in his narrative illustrating experiences of helplessness. Although not powerless to the extent of Dave, Cher’s stories reveal constraints surrounding her decisions in implementing high-stakes assessment. Cher readily admitted she never really questioned high-stakes assessment other than to express frustration with the amount of work it required prior to this study. I imagine Cher never really considered the ethics of high-stakes assessments until I asked her to during her interview.

The contradictions between compliance and agency for a school administrator are examples of implicated advocacy (Carpenter & Brewer, 2014). At once they must comply with state-mandated high-stakes assessment, which implicates them, yet they have some agency for implementation, which opens up space for them to be advocates. My participants felt powerless to advocate, but three of them did find the space.
**Resignation.** Throughout their narratives, my participants seemed to recognize the unethical negative aspects of implementing high-stakes assessments but were resigned to it anyway. Resignation is distinct from compliance in that resignation suggests we have accepted some responsibility and follow through on completing our job duties even if we do not agree. When we are complying, we feel powerless to change something that it out of our control. Resignation is accepting that some things are out of our control, but we take some responsibility for the way it is. For example, Nora and Michelle were resigned to implementing high-stakes assessments. Nora’s resignation came up when she was discussing the students who were not receiving instruction because they were “in a holding tank.” She does not shirk her role in putting kids in the holding tank but instead has accepted that is a necessary evil of implementing high-stakes assessment during testing season. Likewise, throughout Michelle’s story of handing students a certificate of completion instead of diplomas at graduation because the students had not passed a high-stakes assessment required for graduation there is a sense of personal responsibility and acceptance.

The sense of responsibility coupled with acceptance is resignation, which is at the heart of being an implicated advocate. The implicated part is the acceptance that implementing high-stakes assessments are out of our control, but the sense of responsibility speaks to the understanding that school administrators are responsible for being advocates for their students. In this sense, resignation is the act of school administrators accepting the ethical tensions between personal belief systems, requirements of jobs, and obligations to students as part of life. Resignation is the act of living as an implicated advocate.

**Resistance.** In spite of sharing stories of feeling powerless, both Cher and Dave found ways to demonstrate subversive resistance to implementing high-stakes assessments to ELL
students. Both worked for principals who wanted their first year ELL students to take the reading high-stakes assessment in addition to the English proficiency assessment and both resisted. Dave just ignored the directive, but Cher argued with her principal. In either case, the students did not take the FSA ELA. Additionally, Michelle chose not to test a group of 11th grade retakers. Michelle chose to keep quiet and not tell anyone but her principal, who supported her decision. At any rate, my participants’ stories exposed little in the way of resistance.

**Obligation versus inspiration.** My participants’ obligations to students, to their jobs, to their families often were at odds with the aspects of their jobs that inspired them. For example, Nora sadly outlined how her teacher used to be able to do more creative teaching before high-stakes assessments and now she sees teachers teaching to the test. Students who were thriving despite the odds inspired Michelle, but she was frustrated by the system that did not recognize their efforts. The source of inspiration for my participants was not in the scope of my study. I regret not asking more questions about where they find joy in their jobs.

**Student Performance**

Throughout the stories I collected from my participants concerns about students and student performance abound. Ethical concerns about how students experienced high-stakes assessment implementation during testing season and the aftermaths students continued to experience during and through years of schooling came up with all four narratives. My participants shared stories of students feeling pressure to perform, of students being over-tested, and the consequences for students who do not perform well on high-stakes assessments.

**Pressure to perform.** Perhaps the most tragic story concerned with the impact of high-stakes assessments on students was Nora’s narrative of her student assistant typing into the test
she wanted to kill herself. As Nora makes clear, after interviewing the student, the student never truly intended to kill herself. Nevertheless, the student was compelled to write during a high-stakes assessment that she wanted to kill herself because she did not think she could be successful on the test. The pressure this student felt to perform permeates from messages students receive from teachers, school administrators, parents, and peers throughout the entire school year. As Cher’s story illustrates, students’ placement in classes is based almost entirely on student test scores, which becomes real for students during testing season. Cher explained her students do not want to be in “boom-boom” or low-level classes, thus they put effort into the high-stakes assessments. Additionally, Michelle shared her frustration with the graduation requirements the state of Florida has and questioned the policy. Michelle saw this pressure for students to perform unnecessary and devastating to her students’ futures.

**Amount of testing.** Nora, Cher, and Michelle lamented the number of tests students must undergo throughout their careers. For Nora and Cher, many high-stakes assessments were redundant and unneeded. For Michelle, high-stakes assessments were unending. In high school, students who do not perform well on high-stakes assessments get more high-stakes assessments. As Michelle suggested, it is not a testing season but a testing year. As I shared in chapter one, when I was a test proctor I would see the same students over and over again. I wonder how it stands to reason a student who did not pass the FSA reading retake in October would be ready to achieve a concordance score on the SAT three weeks later.

**Management Versus Leadership**

My participants’ stories were replete with examples of management. Managing the testing season and all the decisions associated with it were apparent in Nora, Michelle, Cher, and Dave’s narratives. Indeed, there are few examples of leadership in any of their stories except for
Cher, Dave, and Michelle’s decisions to not test groups of students. Perhaps this is due to the overwhelming responsibilities accompanying testing season. The management of testing season is the implicated part of the school administrators’ role in high-stakes assessment. The leadership of testing season opens the space for advocacy. Unfortunately, my participants’ stories suggest they did not take advantage of that space.

**Ethics of Implementation**

The ethics of implementation of high-stakes assessments brought me to this journey. I personally struggled with implementing high-stakes assessment, as I do not believe high-stakes assessments are in the best interests of students or public school systems. My participants’ stories illuminate the tensions we as school administrators deal with when implementing high-stakes assessments. However, the ethics of implementation go deeper then personal ethical tensions as the narratives in this study have revealed. These stories include questions about the ethics of implementation for our schools, our districts, other administrators’ decisions for implementation, and the state accountability system as a whole. These stories of implementation reveal the implicated advocacy qualities of school leadership.

**The Closing of Loopholes and Gaming the System**

Most of the research on high-stakes assessment occurred from 2002-2009 since that time states began the closing of loopholes in expectations. Many loopholes allowing for schools and districts to game the system during the beginning years of NCLB are gone. For example, schools kicking kids out of school so they do not test was occurring (Figlio & Getzler, 2002) but now graduation rate is calculated as part of school grade. Another area ripe for gaming back in the day of NCLB, was finding students eligible for special education services to avoid having to include them in the test scores for the school (Figlio & Getzler, 2002). Now schools in Florida
must test at least 95% of the students required to take a particular high-stakes assessment, or the school will receive an incomplete as a school grade, which will eventually revert to an F after an investigation. Students with disabilities must take the FSA and EOCs or FSAA. Students with disabilities can obtain waivers for the two high-stakes assessments they are required to pass to graduate, 10th grade FSA ELA and the Algebra 1 EOC, but they must attempt each assessment at least once to qualify. Thus, schools can no longer hide kids.

Nevertheless, schools and districts still find ways to game the system, as Cher and Dave both discuss in their narratives, first-year ELL students can be used to help schools obtain points by creating baselines one year to show gains the next, thus boosting the school grade. Both participants grappled with the ethics of asking students who already struggled to complete the WIDA test to take FSA ELA. In both cases, Cher and Dave subverted authority and resisted testing the first year ELL students in their schools. At the time of this writing, Dave reported to me his district is now requiring all schools to test first year ELLs on FSA ELA, in spite of the state allowing schools districts flexibility on this issue.

Michelle also discussed gaming the system when she spoke about schools manipulating which students took which classes when, thus impacting which students took which tests when. For example, some schools have chosen to have some 9th grade students take liberal arts math rather than algebra 1. The students will still need to take the algebra 1 EOC in their sophomore year, but for one year these students no longer in the required to be in the test pool. The same gaming happens with the biology EOC, low-performing 9th graders are placed in physical science instead of biology. The students are still required to take and pass biology and sit for the biology EOC, but for one year these students are no longer in the testing pool. It may not seem like much of a game since the schools are only keeping the students out of the testing pool for a
year, but for a school with a D or an F grade, this practice can mean the difference between a state takeover or not. For example, one low-performing high school in danger of state takeover had struggling 9th graders take liberal arts math and physical science, which resulted in huge bump in test scores that year and prevented the takeover. The following year, when these students were included in the testing pool, the scores dipped, but the school takeover process is so complex and requires school grades of Ds or Fs for consecutive years the school did not face takeover.

It could be argued this type of gaming is not inherently unethical. It might be beneficial for struggling students to have another year to mature and master content before taking the high-stakes assessment. Preventing a state takeover might also be in the best interest of the students and the community. The very notion of gaming the system implies implicated advocacy. The school administrator may be implicated gaming the system for personal benefit within state systems of accountability, but they may be gaming the system as a form of advocacy.

**Students with Disabilities**

The contradiction of students with IEPs taking standardized, high-stakes assessments pointed to the contradictions in federal education policy. According to Cotto (2016), “if students with disabilities require individualized education programs, then perhaps a blanket policy of inclusion or exclusion into testing systems is contradictory. A simple rule that all students with disabilities are included or excluded without regard for individual students and the context will continue to be insufficient” (p. 26). The goals of federal and state mandated high-stakes assessments conflict with the goals of federally mandated IEPs (Sackel, 2006). Dave’s narrative touches on this contradiction as he attempts to provide teachers with every possible accommodation of every student with disabilities in each testing room on one piece of paper.
**Project Management and the Allocation of Resources**

As students of Taylor’s (1911) *Scientific Management*, Henry Gantt and Henri Fayol are considered the founders of project management, thus the roots of project management are embedded in business/production efficiency (Seymour & Hussein, 2014). Project management is a process which is time-bound, temporary, scheduled, resource necessitated, “execution of activity,” leading to the production of a service or widget (Schwindt, 2005). I would argue high-stakes assessment season or testing season fits the definition of project management. High-stakes assessment is time-bound, as it must be completed within a time frame designated by the state. High-stakes assessment takes place in testing season, which typically lasts for one or two months, thus it is a temporary endeavor. A school administrator or testing coordinator, the project manager, create a schedule to ensure timely completion of the project and oversees it. The resources allocated toward high-stakes assessment administration are limited but necessary for carrying out. The execution or administration of high-stakes assessments to students is the purpose of the activity, which leads to the production of students’ assessment results.

One of the most critical aspects of project management is the allocation of resources. Throughout my journey, my participant’s narratives often referred to how they would allocate human and capital resources during testing. For example, Nora discussed how her testing committee, consisting of her media specialist, guidance counselors, reading coach, tech specialist, and her senior assistant principal, managed testing together during test administration. One of them was always at the ready to run a laptop to a classroom. Cher’s story includes how her success coach and guidance counselors help her manage testing. For both Cher and Nora, managing the project of high-stakes assessment required reallocating human resources to assist with the project. In Cher’s narrative, she describes helping schools determine who the bubble
kids were to help direct human resources toward them for test preparation. Another participant, Dave, laments multiple days of professional development on high-stakes assessment administration that teachers were required to attend instead of being in their classrooms teaching. Cher and Michelle’s stories also address the use of paraprofessionals for the purpose of high-stakes assessment. All of these examples result in the diversion of human resources away from students.

Further, Dave, Cher, Nora, and Michelle’s narratives all refer to relocating of students to other classrooms or parts of the school to make way for test administration, which diverts the resource of physical space away from students. All participants discussed the amount of time, the most precious resource, students are pulled out of classes to test in our stories. The enormity and priority of high-stakes assessment implementation necessitate resources are diverted away from their intended purpose. The project management of high-stakes assessments during testing season requires school administrators to allocate resources among competing projects, with the priority project always being testing. Unfortunately, this leads to the diversion of needed resources away from students elsewhere in the school.

The districts in the region I work in do not frown upon this misuse of school resources, and in fact, it is encouraged or at least an accepted way of work. There are enough Florida Department of Education employees in and out of schools throughout the region for me to assume the state must also be aware of how schools divert resources to complete high-stakes assessments within the given time frames. Furthermore, the state of Florida only funds the direct costs of high-stakes assessment expecting schools and districts to figure it out how to manage the actual costs, which suggests the state is fully aware of how schools allocate and reallocate resources during testing season (Jakee & Keller, 2017). Nevertheless, schools are faced with not
only competing with one another through avenues of choice and privatization but also within the school exists completion for scarce resources.

Consider the use of the success coach for assisting with testing. The success coach is meant to be a resource for students whom the district or state deems at-risk to graduate. These students have risk factors such as previous grade-level retentions, repeated failures on high-stakes assessments, histories of disciplinary consequences, low school attendance, and usually at least two or more of these factors. The school district created a position for someone to work with these students to help them get back on track to graduate. Then school leaders decided this was important enough to hire someone to be a success coach at their schools only for school leaders to turn around and divert this human resource away from the students the success coach was hired to work with in the first place. High-stakes testing presides over all creating instances of implicated advocacy in decisions school administrators must make during testing season.

**The New Normal**

Little research has focused on the negative impact of high-stakes assessment in the last decade, save the work of Wayne Au (2009, 2011, 2016). Why? Perhaps because in the wake of NCLB many researchers had a great deal to say, but as the years have gone by and NCLB became RTTP then ESSA, researchers shifted their focus to other education policies. The amount of high-stakes assessments given to students has been normalized. Except for parents and public school employees protesting and opting out in 2014, high-stakes assessments are no longer news. Yet, much has changed in the way students are tested since the heyday of anti-accountability research. When NCLB was adopted, high-stakes assessments were administered on pencil and paper over the course of five or six days for the entire school for math, reading, writing, and science tests. This remained the way students were tested in Florida until 2013. In
the days of paper and pencil, school resources were all completely focused on high-stakes assessments for those five or six days.

However, with the widespread adoption of computer-based testing, high-stakes assessment now happens in shifts because public schools do not have enough computers for all students to test simultaneously, as described by Michelle and Nora in their stories. Thus, many resources are diverted from their true purpose to assist with testing, while the rest of the school attempts to operate as normal as possible. Normal means students being moved from the classroom they are assigned to meet in to another room for a few weeks or a month to make room for students who are testing. Normal means all available resource teachers and academic coaches are test proctors because the classroom teachers are still teaching. Normal means middle and secondary classroom teachers must teach the same lesson multiple times because on Monday half the class is testing and on Tuesday the other half the class is testing, but only in the morning. Most testing is scheduled for the morning. Thus, afternoon classes have all the students, which means teachers can move forward with the lesson but now the afternoon classes are many days ahead of the morning classes. Normal means every computer in the school is unavailable to anyone who is not testing during testing season. Normal means the media center is shut down for weeks at a time to students because there are banks of computers housed there being used for testing. Normal means school administrators spend more time planning and carrying out testing than any other task. Normal means everyone, students, teachers, parents, and administrators, believes they must take or administer the test. Normal means students who do not make the cut spend years and years in remedial classes. As I write this, it sounds absurd, yet this is normal in the day-to-day workings of middle and high schools in my region. It is
accepted as the way it is. It is anything but normal. This normal way of work has created implicated advocates out of all school administrators in public school settings.

Of all of the stories I collected from my participants, Dave’s story of his principal asking for contingency plans in the case of a school shooting is the only one that genuinely depresses and disgusts me. The school shooting at Stoneman-Douglas occurred during the testing season I was conducting interviews for this study. Discussions around school safety and shoring up schools to protect students and teachers from possible school shooters were happening in the media, social media, schools, districts, and Florida’s legislature. I can only imagine this principal was responding to these recent events when she asked Dave what his plan was for securing the testing materials in the case of an active shooter. This is the result of Neo-Taylorism gone to a perverse extreme. This principal prioritizes the control and efficiency of high-stakes assessments over the safety of the children and adults in her care. For her, high-stakes assessment, and all the plans and contingency plans take precedence over humanity. This principal’s behavior goes to questions about ethical school leadership as well. The ethical dimensions of school leadership, no matter what ethical paradigm studied, all place human safety at the forefront of the school administrator’s ethical responsibilities. Her question is morally reprehensible. I cannot think of one person I know who would knowingly put their child in the care of this principal.

An Examination of Systemic Issues Using the Ethic of Critique

The ethic of critique, as outlined in chapter two, begs the answers to the following questions: “who makes the laws, rules, and policies? Who benefits? Who has the power? Who is silenced?” (Shapiro & Gross, 2008, p. 25). I would be remiss if I did not examine the systemic issues surrounding the participants in my study using a critical lens. Cher, Nora, Michelle, and
Dave all identified connections between high-stakes assessments and those in power, or lack thereof.

Who does make the laws, rules, and policies? While there may seem to be an obvious answer, the laws, rules, and policies guiding the implementation of high-stakes assessments are created and enacted by layers of bureaucracy. The history of accountability in the state of Florida and the United States presents us with the sources of high-stakes assessment we continue to live with today. It is easy to blame Governor Jeb Bush and his brother, President George W. Bush, and undoubtedly, they are to blame, but there are many cogs in this wheel. Many lawmakers have contributed to the system of accountability over the years. There is a federal Department of Education and Florida Department of Education filled with politically appointed bureaucrats who interrupt the laws and filter them down adding guidelines, rules, and policies along the way. School district bureaucrats then further filter down the laws, guidelines, rules, and policies and make interpretations to the school level. School administrators take all the information they have been given and implement high-stakes assessments following the laws, guidelines, rules, and policies provided to them from the bureaucrats above them.

Who benefits? As noted by Cher, business people benefit. The makers of the assessments, the canned curriculum, the test preparation materials, and the keepers of the data all benefit from high-stakes assessments. The politicians, who are elected and reelected based on their education platform, benefit. The lobbyists benefit. The bureaucrats, who must justify their jobs, benefit. Charter schools benefit. Who has the power? Nora and Michelle pointed to the state having power. They are right, but who is the state? I think they meant the politicians and the bureaucrats.

---

3 See chapter one for a history of high-stakes assessment nationally and in Florida.
Who is silenced? The most vulnerable are silenced. Students with disabilities, students who are English language learners, students who live in poverty, students of color, and every student who has been subjected to test preparation, remedial reading and math courses, and students who have been retained. Educators, too, have been silenced. Rarely are educators asked what they think about a law, a rule, or a policy impacting them or their students. Families and communities are silenced. When the Opt Out movement was at its height, the state and schools districts sought to silence them and discourage them from opting their students out. Thus, those who are impacted most by high-stakes assessment laws, rules, and policies are the ones who are silenced the most.

To what end? To accurately critique the system of high-stakes assessment one must consider why and what is to gain. First, there is a lot of money being made by those who benefit. Second, it is social control of the masses is maintained those in power (Au, 2007; Berlak, 2000; Bowles & Gintis, 1976). As I described in chapter one, the use of, so-called, objective measurements to assess and assign students to groups has been in practice since the turn of the 20th-century. I suspect my participants and I feel so powerless because of everything I have described in this section is out of our control.

Reflections on My Ethical Obligations as the Researcher

As noted in the Limitations section in chapter three, there were times I struggled with thoroughly critically interrogating my participants’ narratives. This was partly because I was embedded within the same school district as Nora, Cher, and Michelle, but also because I was friends with them, especially Nora and Cher. As I reflect on the study I recognize I have an ethical obligation to not only question the beliefs and practices shared with me during my research, but also directly address them with my participants. Thus, I am an implicated
advocate, too. The nature of this study suggests advocacy, but by ignoring these practices and beliefs, I am implicated. Some practices and beliefs must be questioned, and I plan to bring it to their attention in future discussions on my journey.

Indeed, there were many times when I could not “see” or recognize that Cher, Michelle, and Nora were resigned or complying with directives because I was entrenched in the same compliance and resignations. The practices I did not recognize as problematic since I assumed these practices were common or typical or even that I have never experienced different approaches to implementing high-stakes assessments. Thus, not only do I have an ethical obligation to question my participants’ practices and belief systems, but also I have a responsibility to interrogate my own, which I have attempted to do throughout my journey and will continue to do in future practice and research.

Conclusion

Throughout this journey, I sought to collect narratives illuminating the motivations of school administrators as they negotiated the ethical tensions between their own personal belief systems and the requirements of their jobs. The collective narrative of the ethics of high-stakes assessment implementation suggests there is much school administrators struggle with during implementation. The culture of measurement and the need for school administrators to perform was pervasive in the collective narrative. Also, persistent in the collective narrative was the intensity of the testing season for school administrators. The collective narrative points to the implicated advocacy nature of school administration in public schools in high-stakes accountability conditions. Next, on the last stop of my journey, I will discuss the narrative possibilities of high-stakes assessment implementation.
CHAPTER 9: THE TALE OF NARRATIVE POSSIBILITIES

“Something good comes with the bad

A song’s never just sad

There’s hope, there’s a silver lining

Show me my silver lining” (Söderberg & Söderberg, 2014).

**********

“Dr. Canady! Please just listen to me! We cannot get it all done unless we use the media center,” Julia, the senior assistant principal, begs.

I have just been appointed the principal at Craven High School, and I am working with my senior AP to create a testing schedule, which accommodates the students rather than the state.

I shake my head and tell her, “Julia, I appreciate your concern, but we have been over this many times. We are not shutting down the media center for two months. I don’t care if we have to move every computer out of the media center and relocate them to the least utilized rooms on this campus, but we are not closing down the media center.” Julia is already frustrated with me for limiting which resource teachers she can use to proctor high-stakes assessments. I am sympathetic to Julia’s position and recognize our school’s limited resources but to am steadfast about this and refuse to bend.

**********

My pilgrimage does not end here; it remains unfinished, as was Chaucer’s, but not as abruptly. Indeed, my journey may never be complete, but I chose to end my writing here with
narrative possibilities. Chapter’s four through seven of this quest presented the narratives of four school leaders’ experiences implementing high-stakes assessments in a culture of measurement during testing season in Florida. At times the stories shared were downright depressing and even frightening. My participants are what Carpenter and Brewer (2014) call, “implicated advocates,” school leaders who at once advocate for their students and are implicated in accountability policies including high stakes assessment. School administrators of public schools in Florida live within the ethical dilemmas between advocate and implicated. On the eighth stop on my travels, I reflected on my participants’ collective narratives, in this chapter I offer possibilities for practice, possibilities for scholarship, and possibilities for educational leadership preparation while considering the role of the implicated advocate school administrator. Throughout this stop on my journey are not only my hopes of the future of practice and research but also my daydreams about what could be in the future.

The narrative possibilities I suggest are temporal stances working within the existing accountability policies. It is highly unlikely public education policies in the United States will abandon high-stakes assessments, nor the overarching accountability policies set in place over twenty years ago. Perhaps, the preeminent scholar focused on high-stakes assessment policy, Wayne Au (2007, 2008, 2009a, 2009b, 2011), recognizes this because many of his critical works does not call for solutions or remedies. Thus, I wish to consider ways to address the ethical issues surrounding high-stakes assessment by pushing against and finding spaces within the existing policy’s boundaries. For me, this approach is simultaneously hopeful and realistic in nature.
Narrative Possibilities for Practice

I enter a classroom and look around at the beautifully decorated room. The pale blue walls are calming, and the furniture is arranged in four groupings of desks and chairs. The teacher, Mrs. Smith is walking around helping students with a chemistry experiment. The students in the class represent diversity in every way – students of all abilities, colors, multiple languages, and test scores are together. This is how all of the classrooms and students at Craven High School, an experimental school, are grouped, completely heterogeneously. We have de-tracked students and try not to stress about high-stakes assessments, which continue to loom large over public education in the state of Florida.

Opt out, come back!

My story of narrative possibilities starts with school leaders who demonstrate advocacy for democratic public schooling for students and families by supporting the return of the Opt Out Movement. The Opt Out Movement was a very loose mixture of parents and educators supporting students’ rights to Opt Out of high-stakes assessments. It was a strong force in education during 2014-2015. In spite of the strength of the movement at the time, Opting Out petered out shortly after it began. Opt Out saw some success in Florida with reducing the number of tests students were required to take. School districts began to use the results of high-stakes assessments in place of final exams, and the Algebra 2 EOC was eliminated completely. This quelled those protesting, but there is still work to be done. It begins with parents and educators continuing to work together to diminish the amount of high-stakes assessments. Thus, the future of high-stakes assessments needs families and students who choose to opt out and school leaders who support and advocate for these choices. School administrations, in partnership with families, must resist the forces of accountability policies and prepare to fight
against the consequences students and schools face when students do not test. Parents and educators should advocate and demand the state allow portfolios to be submitted as proof of reading proficiency in all grades levels (Au, 2016). If it is possible for the state to allow portfolios to establish reading proficiency for 3rd graders then it is possible for all grades. I personally believe supporting an ongoing opt-out movement is ethically the right stance for leaders in education whose responsibility is to advocate for their students and families.

**De-Tracking**

The second possibility is the story of public schools de-tracking students. Tracking students has been a practice since forms of academic or IQ tests were administered to students around the time of World War I (Callahan, 1962). I remember the school I attended tracked us. I realized which group I was in by the time I was in third grade. There were four classes per grade-level. The above average smart class, the almost above average class, the just below average class, and the below average class. I was in the just below average class. Although not necessarily a painful memory for me, realizing I was below average definitely impacted me. Imagine students today, labeled as low performing and tracked into remedial classes, which they refer to as boom-boom in my region. Although the high-stakes assessment will still have its stigma, de-tracking can go a long way to minimize the labels attached to students based on their test performance. De-tracking is no small undertaking especially for large school districts that have been tracking students for decades. Nevertheless, school administrators should advocate for student de-tracking because it is possible and has been successfully done elsewhere (Rubin, 2006). De-tracking students success or failure on high-stakes assessments will limit the labels associated with being a Level 1, inadequate, or in the boom-boom classes, which is the most ethical stance.
Paper and Pencils

School administrators can advocate for a return to paper and pencil high-stakes assessments in all grade levels, which would eliminate much of the diversion of resources referred to throughout the narratives presented in this study. Stories of schools shutting down for five or six days to administer high-stakes assessments does not sound like a solution; however, five or six days versus months of disruptions would at least get the testing over with more quickly and prevent schools from reallocating classrooms, computers, and teachers for months at a time as well.

Experimental Schools

School administrators should begin to advocate for a few schools to operate with zero high-stakes assessment preparation. The schools would still be required to test, but what if a few schools in districts throughout the state could focus on teaching culturally relevant literature-rich curriculum (Ladson-Billings, 1995). The experimental school could even be track-free and utilize culturally relevant pedagogy with absolutely no test-preparation of any kind. An experimental school, such as the one I advocate for, aligns with my ethical stance.

Bring Back the Local

Despite the “decline of the local,” local communities still have some control (Foster, 2004). For example, as I was writing this conclusion I heard from Dave who shared with me his district has mandated all ELL students in their first year in the Florida will be required to take the FSA ELA assessment even though the state does not require it. Local school boards and administrations are allowed to make some curricular and assessment decisions about their schools, but it is limited. It is difficult to implement a culturally relevant curriculum if the state is making curricular decisions rather than local persons. I doubt misusing resources would be
tolerated in a community with more local control. Thus, we at the local level must advocate for our students. As a school administrator, I can advocate for less high-stakes assessments and more student-friendly implementation of high-stakes assessments.

**Give Students a Purpose**

Once high-stakes assessments began to dominate the focus of public school in the early 2000s, the arts and vocational classes began to fade out the curriculum students had access to, particularly at schools where large amounts of students did not perform well on high-stakes assessments attended (Nichols & Berliner, 2007). This narrowing of the curriculum focuses purely on the tested subjects. Unfortunately, students who already struggle with reading and math are left with nothing but reading and math leaving the student no positive reason to attend school. As advocates, school administrators should allocate more resources toward classes students want to be in and classes offering other options after high school in addition to tested subjects. This may result in students having positive reasons to attend school and advocating for our students to have access to arts and vocations is ethically in students best interests.

**Small Acts of Resistance**

Three of the participants I met on my journey found ways to resist the implementation of high-stakes assessments to students who are monolingual during the implementation of high-stakes assessments and to 11th grade FSA ELA retakers. More of these spaces of resistance exist and it is the responsibility of school administrators to find these spaces and advocate for their use. For example, many teachers I know have told me over the years they do what they want when they close the door to their classroom and it is just the teachers and their students. However, this has become increasingly difficult when test-preparation and mandated curriculum are the expectation for what school administrators see when they walk through a classroom.
Although it is the school administrator’s job to surveil teachers’ practices, perhaps school administrators could encourage teachers to plan for and teach something other than mandated curriculum or test-preparation. Advocate for students by resisting the expectations of canned lessons and provide teachers the autonomy and the respect to plan lessons focused on what the teacher knows students should learn.

**Social Justice Advocacy Leadership**

Perhaps the most ethical approach to advocacy a school administrator can have is one of social justice. Social justice leadership is defined by Theorharis’s (2007) as school administrators focusing on, “issues of race, class, gender, disability, sexual orientation, and other historically and currently marginalizing conditions in the United States central to their advocacy, leadership practice and vision” (p. 223). Further, Anderson (2009) argues for advocacy leadership, which calls for school leaders to partner their politics with a strong social justice orientation. School administrators should be the advocates for students of any ability, students who are English Language Learners, students living in poverty, students of color, and students of any gender or sexual orientation. The inclusion of all students in all curriculums while providing extra supports for those who need it is the true definition of equity, and high-stakes assessments serve to exclude students and narrow curriculums. It is the ethical responsibility for school leaders to adopt a social justice orientation and look for spaces to include and expand opportunities for all students.

**Narrative Possibilities for Scholarship and School Leadership Preparation**

“What are you thinking of focusing your research on?” I ask Erin, one of the most brilliant students at the University I teach courses in ethical educational leadership at part-time.
“Well, I am not sure there are so many things that interest me,” Erin replies. “I am interested in state systems of accountability for one and, obviously, ethics. There are so many issues in education and for educational leaders to tackle.”

“That is true,” I agree. “There’s a great deal of work to do. Focus your research on your passions. Ethics in relationship to educational leadership is a broad subject but if you zero in on ethical issues in educational leadership surrounding accountability you will find a plethora of openings for scholarship.”

**********

Scholars who work in the field of educational leadership should advocate for more research in high-stakes assessment policy implementation. Indeed, scholars of educational leadership should be advocates for democratic public schools and should be committed to researching the tensions illuminated in this study. My participants’ narratives exposed multiple avenues ripe for new research or additions to existing research. Scholars in the field are also the teachers of aspiring school administrators, and, in that capacity, I argue for purposeful preparation focused on social justice and advocacy of students and families.

**Allocation of Resources**

More research needs to be conducted at the intersection of high-stakes implementation policy and the use of school resources. As my travels revealed, in the region I work in oftentimes people, places, and things were not used as intended. At this writing, I found no existing research on the allocation of resources by school administrations for the implementation of high-stakes assessment. It is my opinion ethical practice and scholarship would benefit from researchers investigating how school administrators allocate resources, especially during testing season.
Researching and Teaching Advocacy and Social Justice Leadership

Stories of change in high-stakes assessment implementation begin with scholarship intent on social justice and advocacy leadership. More empirical research focused on how social justice is enacted in public schools in spite of high-stakes assessment would benefit the field. Researchers in the field of educational leadership could focus on the negative impact controlling leadership has on the school community and inclusion. Empirical research on school administrators who have de-tracked did not use mandated curriculum or test preparation materials and found spaces for resistance could be valuable for both further research and practices. I recommend future research on the impact of testing season on school administrators in a public school setting in other states as well as in the elementary setting.

Furthermore, scholars of educational leadership are charged with teaching aspiring school administrators. Developing a cadre of practitioners with social justice orientations prepared to apply that vision to schools depends on programs of educational leadership recruiting and retaining scholars with social justice leanings. Aspiring school leaders being provided the opportunities to wrestle with case studies in ethical leadership by professors of educational leadership would benefit the future school administrators and the schools they will someday lead. Exposure to scholarship on social justice and advocacy leadership would also be beneficial, as would avenues of resistance. Indeed, direct instruction on ethical paradigms and ethical approaches to practice with an emphasis on social justice leadership would have a positive impact on public school communities.

Students with Disabilities

One of the most frustrating stories relayed during this study pertained to how students with intellectual disabilities lost out on teachers and teaching during testing season. Research
that tells the story of the negative impact of high-stakes assessment on all students with disabilities would positively impact the research field, practice, and even laws protecting students who are vulnerable. Indeed, part of the irony of high-stakes, standardized assessments is the dichotomy it presents when applied to students with disabilities who have Individualized Education Plans. Perhaps more empirical research denoting this dichotomy would help students with disabilities.

**Students who are English Language Learners**

As with students with disabilities, stories shared by my participants about students who are ELLs were also frustrating to hear. Research concerned with the negative effects of high-stakes assessment and testing season on students in their first year in the country and all English Language Learners could assist advocates for ELLs and their families with negotiating the public school system in the United States.

**Scholarship in Ethical Educational Leadership**

More research surrounding the ethical implications of school administration in relationship to the implementation of high-stakes assessments and testing season in public schools would be valuable to both practitioners and scholars in the field of educational leadership. School administration is an implicitly moral endeavor, which requires practicing and aspiring leaders to consider what is best for students when making decisions carefully. Theoretical research and empirical research on ethical tensions surrounding policy implementation are currently lacking in existing research.

**Ethic of Care**

From test anxiety to labeling, my participants’ narratives suggest student well-being is not the first consideration during testing season. Empirical research concerned with the ethic of
care (Gilligan, 1982) during high-stakes assessment implementation would benefit all students and, hopefully, impact policy and practice. Certainly, research focused on how school administrators enact the ethic of care in their day-to-day work, as school leaders, would benefit the field.

**My Implications for Future Research**

If I were to conduct this research a second time, I would ask my participants questions about their inspirations and the positive aspects, if any, of implementing high-stakes assessments during testing season. I would ask my participants to consider where they have agency and where else in their practice they could find opportunities for resistance. I would ask my participants about the push and pull between implicated aspects versus the advocacy aspects of their roles. Going forward I plan to explore the deficit language more closely, perhaps employing discourse analysis methods, as well as reexamine the performativity and emotional labor described by my participants, and further connect their experiences with my own. Further, I plan to further research resource allocation and reallocation during testing season on a deeper level.

**Conclusion**

*It is a beautiful day at Craven High School, the spring breeze flaps the school flag back and forth in a memorizing dance. These are the days we live for in Florida. Daniel, the junior assistant principal, gives me thumbs up from across the courtyard to let me know testing has started smoothly this morning. It has not been easy being a new principal at this high school. I have had to work hard to demonstrate we, teachers and administrators, have more agency to make decisions than we often realize. This is true in many situations, including high-stakes assessments. Although we are in the midst of testing season, the general vibe is calm here at*
Craven* because we refuse to allow testing to rule our students and us completely. Decisions made based on our priorities, which are guided by the following quote, “the question concerning what is most important is distinguishable from the question concerning what is morally right” (Frankfurt, 1988). We recognize our students must take the state-mandated high-stakes assessment, but it does not have to be the most important thing happening on our campus.

**********

The purpose of this journey was to explore the stories of ethical dilemmas K-12 educational administrators navigate when implementing high-stakes assessment policy during testing season. The participants’ narratives shared in this study reveals school leaders struggle with the tensions existing between their own personal belief systems, organizational dilemmas, and the requirements of enacting high-stakes assessment policies. My hope was through learning their stories I raised questions about the purpose of high-stakes assessment and the impact of policies at the implementation level and illustrate how school administrators negotiate these tensions. Further, this research illuminates how educational leaders work within the boundaries of high-stakes assessment. I hope I have shined light on how they rationalize or reconcile the differences between their own personal ethics and the requirements of their jobs while navigating these tensions.

My wish is that by sharing the ethical struggles associated with implementing high-stakes assessment policy others who are leading education may be encouraged toward deep reflection and consideration of belief systems, which has the potential to encourage dialogue about the purpose and nature of high-stakes assessment. For those who accept accountability at face value, it is more important to expose them to narratives of those who struggle with accountability, such as the ones presented in this study. By using a narrative inquiry approach to analyzing the ways
in which educational leaders deal with the dilemmas between accountability and a sense of ethics they bring to their day-to-day jobs may be beneficial to those aspiring or veteran educational leaders struggling with similar dilemmas. Furthermore, using the metaphor of a pilgrimage to organize the findings into pedagogical tales may be useful in the preparation of educational administrators. Finally, I conclude with hope for the possibilities of finding spaces within the dilemmas of high-stakes assessment policy implementation for school administration to push back both in theory and in practice.
EPILOGUE: INQUIRY AS PILGRIMAGE

“And your Volvo lights lit up green and white
With the cities on the signs
But you held your course to some distant war
In the corners of your mind” (Schultz, Felice, & Fraites, 2016).

**********

This dissertation chronicled my pilgrimage’s movements through nine stops on my path to my sacred destination: the attainment and creation of knowledge. At each stop I gathered knowledge from existing literature and methods, then I stopped to meet others along my path to gather their stories of implementing high-stakes assessment during testing season in Florida, and I connected their experiences with knowledge collected by previous travelers. My participants were the people and places on my journey. It was their motivations, which propelled me forward on the path to my sacred destination. I stopped to reflect on the collective narrative and to provide temporal stances for future practice and scholarship before ending this account of my inquiry as pilgrimage journey.

The Journey

My travels have taken me a long journey toward the attainment and creation of knowledge. I began on this path years ago, when I was in my first year as a doctoral student and a full-time reading coach in a high school. I felt as though I was proctoring high-stakes assessments every single day. During this time, I distinctly remember creating a poster for Dr. Karanxha’s ethics and educational leadership course where I drew a stick figure straddling both
sides of a line. I was the stick figure and one side of the line represented the requirements of my job and the other side was my own ethical beliefs. The line was the tensions I felt and still feel today.

My journey continued throughout my coursework, as I attempted to understand ethical tensions educational leaders in public schools deal with throughout their careers. I tried to understand the purpose of accountability policies by learning more about the policies and studying the scholarship about accountability. As I completed my coursework, my travels became more focused on a narrow path of scholarship as I focused in on what would become my research questions. Through the dissertation process, the path I took on my travels became very narrow, and at times, unbeaten. Although I have completed my dissertation, my inquiry as pilgrimage will never end because I will always seek answers to my questions for the rest of my life.

The Motivations

My pilgrimage journey has been motivated by my desire to find the answers to my research questions and to reach my sacred destination: attainment and creation of knowledge, in other words, a quest for knowledge. Yet, I end this tale of my journey without reaching my sacred destination. A pilgrimage follows the story archetype of a quest journey (Diaz, 2010; Frye, 1951). According to Woodside and Megehee (2009), it is the journey itself that is at the center of the travel story rather than the completion of the trip. In retrospect, I wonder if my journey was also motivated by what Kociatkiewicz and Kostera (2012) refer to as a quest to find an ethical or “good” leader (p. 861). My aim was to understand how school administrators grappled with ethical dilemmas associated with implementing high-stakes assessment in a culture of measurement during testing season, and I now have a more profound knowledge of this
phenomena. Further, I have contributed knowledge to the field of educational leadership by taking this journey. However, I have to consider the possibility I may have been searching for the ideal educational leader who did what was ethical in all situations no matter the cost, as well as the answers to my research questions.

In retrospect, I think I was also motivated to learn how to live with ethical tensions associated with the implementation of high-stakes assessments. I live in a constant state of ethical struggle between what I think is right and my obligations. Perhaps, this is the life of a school administrator with strong beliefs about social justice working in a public school in an age of accountability. At this writing, I have begun to consider looking for another job. I have become worn down by the questionable ethical practices encouraged by the state’s school grading system. Students are not centered in these decisions, but instead unattainable goals for schools.

My inquiry as pilgrimage has been a journey toward not only understanding how school leaders grappled with ethical dilemmas associated with implementing high-stakes assessment in a culture of measurement during testing season, but also a journey to understand my place, as a school administrator, in this ethical conundrum. My childhood experience as a struggling student figures prominently into my motivation for this quest. I recognize if high-stakes assessments existed when I was in K-12, I would likely have been retained and possibly have dropped out of school. It is hard to reconcile this hard reality with the requirements of my job, which often implicate me as an agent of state whereby I do not always do what is right for students. Thus, learning the stories of other administrators at least provides insight into how others have resigned themselves to the things they cannot change and resist the things they can.

The People and Places

236
The people and the places I have visited along my journey helped me gather tales to deepen my knowledge about the purpose and nature of high-stakes assessment. I needed to learn from Nora, Cher, Michelle, and Dave’s stories. I needed to know where my experiences overlapped with theirs. I needed to gather their narratives together in one place to try to make sense of our individual and collective experiences with implementing high-stakes assessments during testing season. Their motivations, dilemmas, resignations, and resistance propelled me forward on my path of inquiry.

**Implicated Advocates**

My participants and I are all implicated advocates. The resignation and acceptance of how things are may be indicative of being a school administrator in a public school during this period of accountability. In order to lead, one must be both a leader and a manager, one must have the agency to resist and be resigned to accept negative aspects of our jobs, and one is at once implicated in systems of accountability and a social justice advocate for students. We are walking contradictions strategically accepting what we cannot change and changing what we can. This is a messy business with few clean cut and easy answers. Accepting my role as an implicated advocate may be the best way forward.

As I reflect on my inquiry as pilgrimage, and my sudden desire to leave my job as a public school administrator, I feel as though the negative, implicated aspects of our jobs outweigh the advocacy aspects. I find myself implicated more than advocating. There is no balance between the managing and leading, no balance between resignation and agency, no balance between being implicated and being an advocate. I am past the tipping point. I am not doing what is best for students but I am doing what is best for me and my career. I am finding it hard to live with myself. I am more implicated because I recognize things are wrong, and I am
doing them anyway. Thus, the way forward may not only be accepting my role as an implicated advocate, but to find more spaces to engage in student advocacy.

**The Sacred Destination**

I will spend the rest of my life on my journey to my sacred destination: the attainment and creation of knowledge. Now that I have completed my dissertation, I am closer than before because my findings have enlightened me. I do know I have a moral purpose for my role in public schools. I have a new understanding that leadership is partly accepting I cannot do the right thing for everyone all of the time. I am not giving up trying to do what is right, but I realize a leader should feel conflicted because most ethical dilemmas are not choosing between what is right and what is wrong, but between what is more right and what is less right. I will never agree with the nature and purpose of high-stakes assessments, but I will look for places and spaces to resist. I endeavor for all the narrative possibilities outlined in chapter nine will come to pass.

**Conclusion**

I will never completely finish this journey toward my sacred destination: the attainment and creation of knowledge. Yet, the part of my travels chronicled in this dissertation represent my path toward the culmination of my doctoral degree. I have indeed attained and created knowledge on this quest but still have much to learn and share with future pilgrims. I hope my inquiry as pilgrimage and the tales I have studied, collected, and presented in this dissertation will inspire other veteran and aspiring school and district administrators to become advocates in their roles in the praxis of high-stakes assessment as ethical practitioners. I hope my journey will inspire other scholars to take similar journeys following my path and creating their own paths of knowledge.

******

238
“The long and winding road
That leads to your door
Will never disappear
I've seen that road before
It always leads me here
Lead me to you door” (Lennon-McCartney, 1970).

*******
REFERENCES


doi:10.3102/0013189X07306523

doi:10.1080/01425690802423312


doi:10.1080/00220272.2010.521261

doi:10.1177/0895904815614916

http://ecommons.aku.edu/pakistan_ied_pdcn/3 adresinden erişilmiştir.


http://www.fldoe.org/core/fileparse.php/7764/urlt/GradRequireFSA.pdf
https://www.flgov.com/meet-governor-scott/


Frick, W. C., Faircloth, S. C., & Little, K. S. (2012). Responding to the collective and individual “best interests of the students”: Revisiting the tension between administrative practice


Horn, C. (2003) High-stakes testing and students: Stopping or perpetuating a cycle of failure?, *Theory Into Practice, 42*(1), 30-41. doi: 10.1207/s15430421tip4201_5


Muzaliwa, A. I. & Gardiner, M. E. (2014). Narrative inquiry (ni) as an exemplary method of social justice leadership. In I. Bogotch & C. M. Shields (Eds.), *International handbook of educational leadership and social (in)justice (pp. 183-198)*. Dordrecht, NL: Springer


Archives/Archivos Analíticos de Políticas Educativas, 20(20), 1-30. doi:


Oakes, J. (2008). Keeping track: Structuring equality and inequality in an era of
accountability. Teachers College Record, 110(3), 700-712.

9-44. doi: 10.3102/002831208320573

inequality. Cambridge, MA: Civil Rights Project at Harvard University.

Ospal, T., Woglemuth, J., Cross, J., Kaatna, T., Dickmann, E., Colomer, S., & Erdil-Moody, Z.
(2015). “There are no known benefits…”: Considering the risk/benefit ratio of
qualitative research, Qualitative Health Research, 26(8), 1137-1150. doi:
10.1177/1049732315580109


https://www.huffingtonpost.com/alan-singer/people-are-pissed-at-pear_b_6921702.html

https://bits.blogs.nytimes.com/2015/03/17/pearson-under-fire-for-monitoring-students-twitter-posts/

Shapiro, J. P. & Gross, S. J. (2008). Ethical educational leadership in turbulent times:

Shapiro, J. P. & Stefkovich, J. A. (2001). Ethical leadership and decision making in education:


275


http://digitalcommons.unl.edu/geographyfacpub/2


Theoharis, G. (2009). *The school leaders our children deserve: Seven keys to equity*, 278


APPENDIX A: INTERVIEW QUESTIONS

1. What aspects of implementing or administering a high-stakes assessment have you agreed with and/or disagreed with?
   a. What was this experience like for you?
   b. Can you describe a particular time or situation that exemplifies this?

2. Can you describe a situation where you knew students were taking tests for reasons you personally disagreed with?
   a. What was this experience like for you?
   b. Can you describe a particular time or situation that exemplifies this?

3. Tell me in your own words how you feel about high-stakes assessments? Do you have any personal ethical concerns? Do you have any reasons you might support them?

4. How does being the implementer of someone else’s educational policies make you feel?
   a. What is this experience like for you?
   b. Can you describe a particular time or situation that exemplifies this?

5. Do you ever feel conflicted about implementing high-stakes assessments?
   a. Explain the basis for these feelings? What is the internal conflict like to experience?

6. Do you ever feel like you struggle to reconcile your own personal beliefs with the requirements of your job?
   a. Can you give me an example of this?
7. What is the purpose of high-stakes assessment? Do you agree with the purpose? Can you give me an example why you agree or disagree?
APPENDIX B: UNIVERSITY IRB APPROVAL LETTER

February 12, 2018

Jennifer Canady
L-CACHE - Leadership, Counseling, Adult, Career & Higher Education
Tampa, FL 33612

RE:  Expedited Approval for Initial Review
IRB #: Pro00031697
Title: Tales of Ethical Dilemmas: Educational Leadership, Ethics, and High-Stakes Assessment

Study Approval Period: 2/9/2018 to 2/9/2019

Dear Ms. Canady:

On 2/9/2018, the Institutional Review Board (IRB) reviewed and APPROVED the above application and all documents contained within, including those outlined below.

Approved Item(s):
Protocol Document(s):
Study Protocol

Consent/Assent Document(s)*:
Informed Consent.pdf

*Please use only the official IRB stamped informed consent/assent document(s) found under the "Attachments" tab. Please note, these consent/assent documents are valid until the consent document is amended and approved.

It was the determination of the IRB that your study qualified for expedited review which includes activities that (1) present no more than minimal risk to human subjects, and (2) involve only procedures listed in one or more of the categories outlined below. The IRB may review research through the expedited review procedure authorized by 45 CFR 46.110. The research proposed in this study is categorized under the following expedited review category:
(6) Collection of data from voice, video, digital, or image recordings made for research purposes.

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

As the principal investigator of this study, it is your responsibility to conduct this study in accordance with IRB policies and procedures and as approved by the IRB. Any changes to the approved research must be submitted to the IRB for review and approval via an amendment. Additionally, all unanticipated problems must be reported to the USF IRB within five (5) calendar days.

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

Sincerely,

Kristen Salomon, Ph.D., Vice Chairperson
USF Institutional Review Board
2/18/2019

Jennifer Canady,
L-CACHE - Leadership, Counseling, Adult, Career & Higher Education
17808 Sunrise Dr.
Lutz, FL 33549

RE: Expedited Study Determined Exempt at Continuing Review
IRB#: CR1_Pro00031697
Title: Tales of Ethical Dilemmas: Educational Leadership, Ethics, and High-Stakes Assessment

Study Approval Period:

Dear Dr. Canady:

On 2/8/2019 5:02 PM, the Institutional Review Board (IRB) reviewed and APPROVED the above application and all documents contained within including those outlined below. Please note that this protocol is now approved under the 2018 Common Rule (45 CFR 46) and is now exempt. Thus, Continuing Review is no longer required and your application will be closed per USF HRPP policy.

The IRB determined that your study qualified for exempt review based on federal exempt criteria for exemption in the federal regulations as outlined by 45 CFR 46.104(d):

(2) Research that only includes interactions involving educational tests (cognitive, diagnostic, aptitude, achievement), survey procedures, interview procedures, or observation of public behavior (including visual or auditory recording) if at least one of the following criteria is met: (i) The information obtained is recorded by the investigator in such a manner that the identity of the human subjects cannot readily be ascertained, directly or through identifiers linked to the subjects; (ii) Any disclosure of the human subjects’ responses outside the research would not reasonably place the subjects at risk of criminal or civil liability or be damaging to the subjects’ financial standing, employability, educational advancement, or reputation; or (iii) The information obtained is recorded by the investigator in such a manner that the identity of the human subjects can readily be ascertained, directly or through identifiers linked to the subjects, and an IRB conducts a limited IRB review to make the determination required by 45 CFR 46.111(a)(7).
As the principal investigator for this study, it is your responsibility to ensure that this research is conducted as outlined in your application and consistent with the ethical principles outlined in the Belmont Report and with USF HRPP policies and procedures.

Please note, as per USF HRPP policy, once the exempt determination is made, the application is closed in ARC. This does not limit your ability to conduct the research. Any proposed or anticipated changes to the study design that was previously declared exempt from IRB oversight must be submitted to the IRB as a new study prior to initiation of the change. However, administrative changes, including changes in research personnel, do not warrant an Amendment or new application.

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

Sincerely,

Melissa Sloan, PhD, Vice Chairperson
USF Institutional Review Board
Informed Consent to Participate in Research Involving Minimal Risk
Pro # 00031697

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

We are asking you to take part in a research study called: Tales of Ethical Dilemmas: Educational Leadership, Ethics, and High-Stakes Assessment.

The person who is in charge of this research study is Jennifer Canady. This person is called the Principal Investigator. However, other research staff may be involved and can act on behalf of the person in charge. She is being guided in this research by Dr. William Black.

The research will be conducted at a location of your choice.

Purpose of the study
The purpose of this study is to:

• Explore the stories and experiences of educational leaders who implement high-stakes assessment policies.
• Create spaces for discussion regarding how to educational leaders work within the tensions between their own personal beliefs and organizational requirements.
• Share stories and experiences of educational leaders who have ethical dilemmas between their own personal beliefs about high-stakes assessment and requirements of their day-to-day jobs.

Why are you being asked to take part?
We are asking you to take part in this research study because of your experience as a school administrator implementing high-stakes assessment policies that you may not agree with have may led you to feel ethical dilemmas.

Study Procedures:
If you take part in this study, you will be asked to:
Participate in three one-hour semi-structured interviews at a location convenient for all participants; maintain an electronic diary weekly using a digital voice recorder provided to you by the principal investigator to record at least one story, and approximately one hour of verifying transcripts and themes. The first focus group interview will occur early Winter 2018 (late January or early February), the second focus group interview will take place in late Winter 2018 (late February or early March), and the third focus group interview will take place in Spring 2018 (late April or early May). You will return your electronic diary to Ms. Canady at the third and final focus group interview.

Transcripts for the first focus group will be made available for participant review before the second focus group and the second focus group transcript will be made available for each participant to review before the third focus group. Transcripts from the electronic diaries and the third focus group will be made available by the end of May 2018.

With your permission the focus group interviews will be recorded and transcribed. To maintain confidentiality, you will be given a pseudonym in all transcriptions and you will not be identified by name on the recording. Transcription software and/or professional transcriptionist may be used to transcribe audio files.

The audio files will be locked in Jennifer Canady’s house. Each participant will be offered a copy of their own audio files and a copy of their own transcriptions. The participants and principle investigator will be the only ones with access to the audio files. The master audio file will remain in Jennifer Canady’s possession and will be destroyed five years after the publication of the dissertation.

Total Number of Participants
About four individuals will take part in this study at USF.
Alternatives / Voluntary Participation / Withdrawal
You do not have to participate in this research study.

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

Benefits
We are unsure if you will receive any benefits by taking part in this research study. A potential benefit of participating in the study could be an increased understanding of dynamics around ethical tension related to high-stakes assessment polices, which could lead to further informing your personal understanding and practice.

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

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

Costs
It will not cost you anything to take part in the study.

Privacy and Confidentiality
We will keep your study records private and confidential. Certain people may need to see your study records. Anyone who looks at your records must keep them confidential. These individuals include:

- The research team, including the Principal Investigator, or study coordinator.
- Certain government and university people who need to know more about the study, and individuals who provide oversight to ensure that we are doing the study in the right way.
- The USF Institutional Review Board (IRB) and related staff who have oversight responsibilities for this study, including staff in USF Research Integrity and Compliance.

We may publish what we learn from this study. If we do, we will not include your name. We will not publish anything that would let people know who you are.
You can get the answers to your questions, concerns, or complaints

If you have any questions, concerns or complaints about this study, or experience an unanticipated problem, call Jennifer Canady at 813-789-6696.

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

Consent to Take Part in this Research Study

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

______________________________________________
Signature of Person Taking Part in Study

Date

______________________________________________
Printed Name of Person Taking Part in Study

Statement of Person Obtaining Informed Consent

I have carefully explained to the person taking part in the study what he or she can expect from their participation. I confirm that this research subject speaks the language that was used to explain this research and is receiving an informed consent form in their primary language. This research subject has provided legally effective informed consent.

______________________________________________
Signature of Person obtaining Informed Consent

Date

______________________________________________
Printed Name of Person Obtaining Informed Consent