

University of Massachusetts Boston

ScholarWorks at UMass Boston

Graduate Doctoral Dissertations

Doctoral Dissertations and Masters Theses

5-2020

How Does Grading Schools Impact Florida's Teachers and Students? The Need for a New Approach to School Accountability

Luke Aubry Kupscznk

University of Massachusetts Boston

Follow this and additional works at: https://scholarworks.umb.edu/doctoral_dissertations



Part of the [Education Policy Commons](#), and the [Public Policy Commons](#)

Recommended Citation

Kupscznk, Luke Aubry, "How Does Grading Schools Impact Florida's Teachers and Students? The Need for a New Approach to School Accountability" (2020). *Graduate Doctoral Dissertations*. 535.

https://scholarworks.umb.edu/doctoral_dissertations/535

This Open Access Dissertation is brought to you for free and open access by the Doctoral Dissertations and Masters Theses at ScholarWorks at UMass Boston. It has been accepted for inclusion in Graduate Doctoral Dissertations by an authorized administrator of ScholarWorks at UMass Boston. For more information, please contact scholarworks@umb.edu.

HOW DOES GRADING SCHOOLS IMPACT FLORIDA’S TEACHERS AND
STUDENTS?
THE NEED FOR A NEW APPROACH TO SCHOOL ACCOUNTABILITY

A Dissertation Presented
by
LUKE AUBRY KUPSCZNIK

Submitted to the Office of Graduate Studies,
University of Massachusetts, Boston,
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY
May 2020

Public Policy Program

© 2020 by Luke Aubry Kupscznk

All rights reserved

HOW DOES GRADING SCHOOLS IMPACT FLORIDA'S TEACHERS AND
STUDENTS?

THE NEED FOR A NEW APPROACH TO SCHOOL ACCOUNTABILITY

A Dissertation Presented

by

Luke Aubry Kupscznk

Approved as to style and content by:

Mark Warren, Professor
Chairperson of Committee

Frank Porell, Professor Emeritus
Member

Kristin Murphy, Professor
Member

Heather MacIndoe, Graduate Program Director
Public Policy Program

Michael Johnson, Chair
Department of Public Policy and Public Affairs

ABSTRACT

HOW DOES GRADING SCHOOLS IMPACT FLORIDA'S TEACHERS AND STUDENTS?

THE NEED FOR A NEW APPROACH TO SCHOOL ACCOUNTABILITY

May 2020

Luke Aubry Kupscznk, B.A., Rollins College
M.S., University of Massachusetts Boston
Ph.D., University of Massachusetts Boston

Directed by Professor Mark Warren

In 1999, Florida began grading schools on an A to F scale. These grades constituted part of the A+ package of policies advanced by Governor Bush's administration. Schools then earned grades based on student standardized test scores. These changes followed a decade of increasing dismay over the trajectory of American education and preceded national moves towards test-based accountability for students and schools. While many researchers have investigated the effects of high-stakes testing on students, few have looked at the impacts of school-level accountability on non-test outcomes. This study considers the impacts of receiving a failing-grade on variables other than test scores for students and teachers.

This study utilizes qualitative and quantitative methods, including in-depth interviews with 12 teachers, two parents and a non-profit representative involved in Florida's scholarship programs. Additionally, using 2010-2015 data from the Florida Department of

Education, it employs fixed-effects regression models to investigate the association between receiving a failing school grade and expected changes in attendance, graduation, drop-out and teacher-separation rates.

The regression results do not suggest strong relationships between failing grades and attendance, graduation, or drop-out rates; they show modest evidence that teachers are more likely to leave a school after it fails. Descriptive statistics illustrate alarming gaps between failing and non-failing schools. Interviews explain some of the reasons for this and illuminate the views of teachers on the practice of grading schools. Most teachers support accountability, but think school grades are simultaneously simplistic and confusing, reducing complex school environments to a single grade. According to teachers, school grades ignore student growth as well as unique challenges and strengths of schools, including socioeconomic and racial divisions. A failing-grade seems to make attracting highly-qualified teachers difficult, while shaming those who work with struggling students.

States like Florida have options for reform. This study suggests that states should incorporate within-year student growth into their evaluation metrics and integrate teachers into the evaluation and reform processes. Some school districts in other states have adopted school quality frameworks (SQFs). SQFs provide a model for how Florida might make grades more holistic and diverse, while better informing parents of school quality.

TABLE OF CONTENTS

ABSTRACT.....	iv
LIST OF FIGURES	viii
LIST OF TABLES	ix
CHAPTER	Page
1. INTRODUCTION: EXPLORING THE NEED TO RECONSIDER SCHOOL GRADES.....	1
2. LITERATURE REVIEW AND THEORETICAL FRAMEWORKS: IMPACTS OF GRADES ON MOTIVATION AND COMMUNITY OUTCOMES.....	40
3. POSITIONALITY, RESEARCH DESIGN AND METHODOLOGY	68
4. PROBLEMS WITH FLORIDA’S SCHOOL GRADING POILCY AS REPORTED BY TEACHERS.....	99
5. TEACHER RECOMMENDATIONS: IDEAS FOR IMPROVING SCHOOL ACCOUNTABILITY	135
6. THE ROLE OF ECONOMIC AND RACIAL SEGREGATION IN SCHOOL PERFORMANCE	180
7. TEACHER THOUGHTS ON THE PUBLIC STIGMA OF FAILING SCHOOL GRADES.....	208
8. TESTING THE IMPACTS OF RECEIVING A FAILING SCHOOL GRADE ON NON-TEST OUTCOMES	243
9. CONCLUSIONS AND POLICY IMPLICATIONS: ADVOCATING HOLISTIC MEASURES OF SCHOOL QUALITY	285
APPENDIX	
A(2): INTERVIEW GUIDE FOR TEACHERS AND PRINCIPALS	315

APPENDIX	Page
B(2): CONSENT LETTER FOR INTERVIEW PARTICIPANTS.....	316
C(2): CODE LIST	318
A(8): DISPLAYS OF SKEW FOR PRE- AND POST-TRANSFORMED VARIABLES	321
B(8): RESULTS OF SCHEFFE MULTIPLE COMPARISONS TEST, TEACHER SEPARATION RATE BY SCHOOL GRADE.....	332
C(8): RESULTS OF SCHEFFE MULTIPLE COMPARISONS TEST, SEPARATED TEACHER EXPERIENCE BY SCHOOL GRADE	333
D(8): RESULTS OF SCHEFFE MULTIPLE COMPARISONS TEST, DROP-OUT RATES BY SCHOOL GRADE	334
E(8): RESULTS OF SCHEFFE MULTIPLE COMPARISONS TEST, GRADUATE RATES OF FOUR RACIAL SUBGROUPS BY SCHOOL GRADE.....	335
F(8): TABLE OF SAMPLE MEANS, STANDARD DEVIATIONS, MINIMUM AND MAXIMUM MEASURES	339
REFERENCE LIST	340

LIST OF FIGURES

FIGURE	Page
1. Percentage of Low Socio-Economic Students by School Grade	220
2. Mean Teacher Separation Rate by Year	259
3. Mean Teacher Separation Rate by School Grade	261
4. Mean Combined Years Experience for Separated Teachers by School Grade..	262
5. Relative Mean Drop-Out Rates.....	266
6. Mean Drop-Out Rates for All Schools by Year (2010-2014).....	267
7. Mean Drop-Out Rates (2010-2014) by Racial Subgroup	268
8. Mean Drop-Out Rates by Racial Subgroup, by Year	269
9. Mean Average Daily Attendance Rates by School Grade	272
10. Graduation Rates by Race in Failing and Non-Failing Schools	279
11. Mean Graduation Rates by Race, by Year.....	280

LIST OF TABLES

TABLE	Page
1. St. Petersburg/Tampa Bay Times “School+Grade” Search Results, Proquest Central Archives by Year.....	75
2. Participation Data.....	82
3. Variables in Initial Dataset	95
4. Breakdown of School Grades by Year	245
5. Variables of Interest – Means, Standard Deviations of Schools (N=3255), in Percent, by Year.....	246
6. Drop-Out Rates – Means, Standard Deviations, in Percent, by Year, by Race	247
7. Chi-Squared Test of the Association between a Failing Grade in Years t-1 and Year t.....	250
8. Chi-Squared Test of the Association Between a Failing Grade in Years t-2 and Year t.....	251
9. Correlation Matrix (Coefficient, R-Value, Number of Observations).....	252
10. Correlation Matrix Including Alternative, Logged Variables.....	254
11. Results of Fixed-Effects Regression Models with Dependent Variable - Log Teacher Separation.....	257
12. Results of Fixed-Effects Regression Models with Dependent Variable - Log Drop-Out Rate.....	264
13. Results of Fixed-Effects Regression Models with Dependent Variable - Log Absence Rate	271
14. Results of Fixed-Effects Regression Models with Dependent Variable - Log White Failure Rate	274
15. Results of Fixed-Effects Regression Models with Dependent Variable - Log Black Failure Rate.....	275
16. Results of Fixed-Effects Regression Models with Dependent Variable - Log Hispanic Failure Rate.....	276
17. Results of Fixed-Effects Regression Models with Dependent Variable - Log Asian Failure Rate.....	277

CHAPTER 1

INTRODUCTION: EXPLORING THE NEED TO RECONSIDER SCHOOL GRADES

In 1999, Florida began grading schools on an A to F scale. These grades constituted part of the A+ package of education policies advanced by Governor Jeb Bush's administration. At the time, schools earned grades based on student standardized test scores – an A for success, an F for failure. These changes followed a decade of increasing national dismay over the trajectory of American education and preceded national moves towards test-based accountability for students and schools.

School grades and other labels used by states contribute to painting a dire picture of a nation in decline. A RealClearPolitics opinion poll in 2019 asked Americans how they would rate the overall performance of the K-12 public school system. Only nine percent of Democrats and 11 percent of Republicans ranked them “excellent.” 55 percent of Democrats and 50 percent of Republicans ranked schools “only fair” or “poor” (Cannon, 2019). Schneider (2017) reports that much of this depression has grown since the passage of No Child Left Behind (2002). In 2002, “24 percent of respondents in the annual Phi Delta Kappa/Gallup poll gave the nation's public schools an A or a B grade. Sixty percent gave

them a C or D.” In 2012, “only 18 percent gave the nation’s schools an A or a B grade. Seventy-two percent gave them a C or a D” (p. 222).

Despite these pessimistic polls, Schneider (2017) argues that “on almost every measure, including the National Assessment of Educational Progress...American schools are doing better today, on average, than at any previous time” (p. 223). Why is there such discrepancy between public opinion and academic outcomes? He concludes that:

Instead of capturing a real decline in quality, then, it may be that the data available to us since 2002 have merely exacerbated the rhetoric of crisis, which has become increasingly pervasive over the past several decades. Such rhetoric is politically useful insofar as it generates momentum for specific candidates or particular proposals; crisis, after all, calls for action. Because this rhetoric does not appear to be closely tethered to reality, though, it has fostered a gap between actual performance and perceived performance (p. 223).

If Schneider is right, and schools are being mischaracterized as in crisis, then students of color and socio-economically challenged students face the brunt of the panic. In Florida, schools serving such students are more likely to be labeled as failing and consequently investigated, restructured, or closed in favor of charter schools. Students within those schools may face exaggerated discipline responses, fueling the school-to-prison pipeline

Questions about the impacts of school labels remain mostly unasked and unanswered in academic research. Researchers have focused on the reliability of tests and the validity of grades to parents, but few have questioned teachers. Some academics have developed more holistic means of evaluating schools and have made suggestions about reporting the data in nuanced, constructive ways, but away from a few of the nation’s most progressive locales, such suggestions have been ignored by policymakers. In places that have adopted more

holistic metrics, like New York City, non-test outcomes sometimes remain disconnected from state accountability measures (Schneider, 2017). I hope to accomplish three things in this dissertation: first, as they are key stakeholders in our schools and implementers of any policy, I intend to demonstrate the value of consulting teachers in evaluating education policies. Second, while it is important to ask whether school rankings or grades are accurate, it is equally important to ask how they impact teachers, students, and communities. In short, we should analyze both the inputs and the outputs of school labels. Third, I argue that academic research can serve an important role in connecting educators, students, and parents with those crafting policy.

This study stems from the lessons and questions of my first years teaching, as well as my studies in the doctoral program of Public Policy at the University of Massachusetts, Boston. In my dissertation, I explore how teachers at various Florida middle and high schools feel about the school grading system, standardized testing, and general state education policy. Teachers tell me how grades affect them emotionally, how they shape school policies, and how they impact students. Their first-hand observations of failing school environments inform their opinions on the system and their suggestions for how to fix it.

Teachers in interviews come across as informed, thoughtful, and passionate about school grades and other policies affecting them and their students. I feel honored to incorporate their words into my findings. Originally, I intended to also include an entire content analysis section, because I suspected that the media might play a role in stigmatizing failing schools. For reasons discussed in my methods section, I ultimately rejected this idea and instead use the quotes and reporting of journalists to substantiate, contextualize, and

complement the statements of teachers. In my quantitative research, I explore the effects of receiving a failing grade on possibly overlooked educational outcomes: attendance rates, drop-out rates, graduation rates, and teacher separation.

Policy Background

Centralized school accountability in Florida dates back to the 1920s. Michael and Dorn (2007) identify three reasons for the state's transition to a more business-like education model. First, elites pushed for school boards to be occupied by experts and professionals, as opposed to elected community members. Second, school systems rapidly expanded in the 1920s and 1930s and taxpayers demanded efficiency. Third, intelligence testing's rise to prominence allowed for the ability-based placement of students in classes. Despite these advocate tenets, testing was controversial even then and a national achievement gap already existed between white students and their black peers (Michael & Dorn, 2007; Miggins, 2014).

In the 1970s, the Reubin Askew Administration began a series of reforms that would eventually lead to A+. At the time, Florida school districts – which had long resisted desegregation using non-racial variables such as “socio-economic” factors as alternatives to race (Cobb-Roberts & Shircliffe, 2007; Lee et al., 2007) – were increasing efforts at integration following the passage of the Elementary and Secondary Education Act (ESEA; 1965) and the 1964 Civil Rights Act. By 1972, the pressure to equalize access to education for students of color prompted passage of the Florida Accountability Act. The act required

that teachers pass a state examination, that schools report their budgets, and that students take the first statewide standardized test – the State Student Assessment Test (SSAT). Michael and Dorn (2007) argue that, at the time, the Citizens Committee on Education made a conscious choice to “interpret accountability [and equal access] as focusing on the outcomes of education rather than what was spent on the different programs” (p. 92) – a sentiment shared by the Jeb Bush Administration decades later.

As in the 1930s, the 1970s saw a significant achievement gap between white students and their peers who had previously attended all-black schools – often with different resources, textbooks, and curricula. Early in the 20th century, black schools often operated in “churches, lodge halls, and sawmill camps” and often lacked “drinking water, sanitary toilets, desks, and adequate supplies” (Cobb-Roberts & Shircliffe, 2007, p. 23). Furthermore, black children often attended school for only three months during the year, due to planter demands for labor in the winter. Some Florida counties spent nearly three times more on white children than on their black classmates (Cobb-Roberts & Shircliffe, 2007).

The state government did not recognize or take responsibility for past or then-present inequalities and instead tasked counties and districts with integration and reform. Michael and Dorn (2007) describe “this shifting of responsibility” as “an important aspect of decision making that, to this day, remains a dominant theme in Florida’s educational politics” (p. 94, see also Lee et al., 2007). The 1973 and 1976 Education Accountability Acts also failed to address inequalities, instead requiring that all students take a standardized state exam (Hamilton, 2003). In *Debra P. v. Turlington* (1979) a federal court judge ruled these conditions unfair and delayed implementation of the tests until 1982, after the graduation of

students who had experienced segregation. Michael and Dorn (2007) argue that this caused the Askew accountability system to “fizzle out” and left “no requirement for programmatic changes or acceptance of responsibility by the state, the legislature, or the district” (p. 95).

The next significant evolutions in state education policy occurred under the Lawton Chiles Administration (1991-1998). It was during Chiles’s tenure that the Florida Comprehensive Assessment Test (FCAT) was developed. The FCAT was designed not only as an assessment, but also as a guide for classroom instruction. Nascent A+ policies accompanied the FCAT, including financial rewards for high-scoring schools and the Florida School Recognition Program, which publicized school rankings. At the time, failing schools were supposed to receive additional resources, expert staff instruction, student transfer options, new staff, and a new principal – none of which was ever used between 1992 and 1995, when the policy changed (Michael & Dorn, 2007).

In the second half of the Chiles Administration, the Commissioner of Education, Frank Brogan exerted more control over education policy. It was Brogan who began using public school rankings and annual published lists of the lowest-performing schools to increase motivation. When Jeb Bush was elected in 1999, voters approved a constitutional amendment that shifted more control over education policy to the governor. Together with Brogan, now Lieutenant Governor, Bush formulated the A+ Plan which connected a large bundle of policy reforms to student performance on the FCAT (Michael & Dorn, 2007).

I include this pre-A+ history in order to stress three things: first, while not as segregated as some other southern states such as Alabama and Mississippi (Lee et al., 2007),

Florida consistently struggled with – or actively resisted – integration throughout the 20th Century. The Federal government has done little to press the issue since 1980 and white students have consistently out-performed students of color (Goldstein, 2014). 11 counties in Florida are still under court-ordered de-segregation requirements (Mack, 2014) and much of the integration achieved in the 1970s and 1980s has been reversed (Goldstein, 2014; Orfield & Ee, 2017). In 1980, 50.6 percent of white children attended racially diverse schools. By 1998, only 37.2 percent did (Cobb-Roberts & Shircliffe, 2007). At the national level, polls often show that whites believe students of color receive equal education, yet few white families would voluntarily send their children to predominantly black or Hispanic schools (Booher-Jennings, 2005).

These achievement and opportunity gaps are reflected in the rest of the nation. Nationally, the average white eighth grader scores higher in both reading and math than the average black or Hispanic twelfth grader; additionally, the dropout rate for Hispanic and black students is more than twice the national average. Similar gaps exist based on family income, parental education, and native language (Linn, 2008). My second point is that, in Florida, these gaps as well as the state's history of racial segregation have been consistently ignored in the crafting of state education policy.

Finally, reforms gradually shifted responsibility for educational equity to counties and districts – which then passed it on to teachers and administrators – while centralizing power in Tallahassee. As Welner and Carter (2013) explain: “schools, principals, and teachers are told that they have ‘no excuses’...this accountability, however, is rarely extended to those making these demands” (p. 3). Ultimately, those in Tallahassee and other state capitals have

not been “required to provide supports necessary for equitable learning opportunities.” (p. 3). While Governor Bush is often cited as the formative figure in Florida education reform (Miller, 2011), his policies actually built upon a history of slowly-evolving centralization and accountability mechanisms (Hamilton, 2003), mechanisms which never “took the next crucial step: holding policy makers accountable for ensuring the conditions and resources necessary to create and maintain a system of excellence that offers universal opportunity” (Linn, 2008, p. 4).

The A+ Plan

Governor Bush intended the A+ Plan to increase achievement among all student sub-groups. While NCLB required schools to measure and report data by student sub-groups, including historically low-performing racial and socioeconomic groups; A+ only designates one – the lowest scoring 25 percent of students. These students have to make certain learning gains or a school cannot achieve an A grade. This was intended to be race neutral – capturing all at-risk students regardless of their race or other demographic factors. In fact, Jeb Bush and others have consistently touted the race-neutral nature of the entire plan, arguing that poor or minority students should not be written off, but held to the same high standards of success as their white, affluent peers (Borman & Dorn, 2007). This approach ultimately delayed approval of Florida’s Every Student Succeeds Act (ESSA, 2015) plan longer than that of any other state.

A+ was built on a market-competition model: through the use of high standards, it would incentivize stronger effort on the part of students and teachers and thus raise achievement. Schools that could not raise achievement might lose students through vouchers or a public choice option to schools that could (Goldhaber & Hannaway, 2004). The Sunshine State Standards had been in development since the early 1990s and had been formally adopted during the Chiles Administration. A+ built upon them through three other components: assessment, public reporting, and consequences (Hamilton, 2003).

Assessment

Initially, A+ relied exclusively on the FCAT as an assessment tool, but expanded the test by adding a science section along with the existing math, writing, and reading sections. Additionally, A+ required students in grades three through 10 to take the test, so that learning gains from year-to-year could be measured (Hamilton, 2003). On the test, students were rated at level one (failing) through level five (highest-scoring) and in 2002, learning gains were added to the equation, such that not only were students expected to pass the threshold level (three), but those at each level of the scale had to show improvement over their scores the previous year. Students in the third grade had to pass the test in order to advance and students in the tenth grade had to pass the test in order to graduate high school (Florida Department of Education, 2014).

Public Reporting

In 1995, prior to A+ and under the Chiles Administration, the state began to publish a list of the lowest-performing or “critically low”- designated schools. In 1996, the state moved away from this label and instead rated schools 1 through 5, just as students were rated on the FCAT. The A+ Plan replaced this numerical ranking with the current A through F grading system. At the time, these grades were based solely on FCAT scores and learning gains of the lowest-performing 25 percent of students (Hamilton, 2003). Over time, the grading rubric became much more complex (Florida Department of Education, 2014).

Consequences

The A+ Plan attempted to combine incentives and punishments to encourage school improvement. Schools that received an A received additional bonus funding based on the number of students while F schools were subject to sanctions and intervention by the state. Students who attended schools that failed two out of four consecutive years would be offered vouchers to attend private or charter schools, or a “public choice” to attend a different public school. In 1999 and 2000, only two schools received two failing grades in a row; about 50 of their students took advantage of vouchers to attend private schools (Greene, 2001; Herrington, 2005).

No additional schools failed twice in a row under initial grading and the voucher program was ruled unconstitutional by the Florida Supreme Court in 2006 (Bush v. Holmes, 2006). Schools failing a fourth year would have been susceptible to state restructuring,

although this was also never triggered (Greene, 2001; Herrington, 2005). The logic of vouchers and public school choice was simple market theory. If schools were forced to compete for students, and neighborhood monopolies were eliminated, then the increased competition would lead to improvements (Figlio & Rouse, 2005).

Additional Changes

Throughout the years since 1999, Florida's school and student accountability system has changed several times. The school grading formula gradually shifted from solely FCAT-inspired to a point system that combined student learning gains, graduation rates, and acceleration rates – which refers to the number of students in advanced placement, dual enrollment, and other college-level courses (Foundation for Excellence in Education, 2012).

As the grading rubric has become more esoteric, parents and educators have expressed frustration and sometimes anger (Maxwell, 2012). In 2010, educators led a state-wide mobilization effort that pushed Governor Charlie Crist to veto legislation that would have overhauled testing requirements, teacher tenure, and tied teacher pay to FCAT scores (Sampson et al., 2010). However, many of these features were signed into law the following year by the subsequent governor, Rick Scott (Harrison & Cohen-Vogel, 2012). For the 2014-2015 school year, an entirely new test – the Florida Standards Assessment (FSA) – along with end of course exams, took its place. Alberto M. Carvalho, Superintendent of Miami-Dade County Schools told the New York Times that these changes, and particularly new requirements for 2015, have had an “emotional effect on students, teachers and parents [that]

has been damaging; the manifestation of sadness and frustration is real” and he called for “the state to pay attention to it” (Alvarez, 2014).

This frustration and emotional turmoil worsened when school grades were published late in the winter of 2016. Teachers and administrators argued that this release – late in the school year – made little sense and failed to give teachers information that could have impacted instruction for the year. Furthermore, the state called grades “partial grades” since the state had to make changes to the rubric to ensure that record numbers of schools did not fail. Grades ultimately did not count towards school evaluations (Akkaraju, Atamturktur, Broughton, & Frazer, 2019).

Significance

There is a significant gap in existing literature regarding outcomes of public school grading other than standardized test scores. Outcome variables such as drop-out, attendance, graduation and teacher retention rates, as well as more qualitatively measured things like discipline infractions – are important both as educational outcomes valued on their own and as mitigators of academic performance.

Attendance problems – both truancy and frequent moves between schools – have been shown to negatively affect assessment scores as well as other outcomes such as delinquency rates (De Witte & Csillag, 2014; Parke & Kanyongo, 2012). Discipline infractions often lead to student suspensions, ensuring absence from class. Additionally, suspensions have been found to correlate with decreases in academic performance for the

suspended students, as well as with “collateral damage” to other students at the school (Perry & Morris, 2014). Worse still, discipline infractions occur more often in schools with higher percentages of students of color (Anyon et al., 2014) and students with disabilities (Sullivan, Van Norman, & Klingbeil, 2014). These disciplinary measures, as well as student drop-out rates, contribute to lower-income students of color disproportionately ending up in the criminal justice system, a phenomenon known as the school to prison pipeline (Skiba, Arredondo, & Williams, 2014).

Too often, policymakers ignore these contributing factors or consider them in isolation from accountability measures. However, if the publicized failure of schools and the subsequent impacts on school cultures and policies adversely affect any of these measures, then there may be important links between the school-grading policy and the academic struggles of students. For example, many attribute high rates of teacher turnover directly to student demographics, under the assumption that particular student sub-groups are more difficult to educate. Challenging this attribution, Johnson, Kraft, and Papay (2012) find that teachers are more likely to leave schools with poor work environments or hostile school cultures, which are more common in schools that serve low-income students of color. Others find that teachers leave schools with lower or failing ratings for those more highly graded (Burns, Martin, & Collie, 2019). Although such research does not show causation, it demands further investigation into the effects of grades on teacher perceptions.

These possible associations suggest that this dissertation should be a valuable contribution to our understandings of accountability policies and student achievement. Additionally, it may prove useful to the state of Florida, which despite a consistent reluctance

to change course on school grades, is constantly massaging education policy. For example, in 2012 the state – in a move that angered many parents – published racially-based academic goals. For reading, the goals were 90 percent proficiency for Asian students, 88 percent for white students, 81 percent for Hispanic students, and 74 percent for black students. While the state insisted that these goals were an attempt to recognize the different starting points of different student sub-groups, they came across as racist to many community members (Toppo, 2012).

As another example: in early 2015, the Scott Administration, in response to what Orange County School Board Chairman Bill Sublette called a “parent revolt,” moved to eliminate the Florida Standards Assessment (FSA), a reading test taken by students in the 11th grade. This did little to assuage parents angry that high school students will take a new version of the FCAT test as well as end of course exams in many subjects (Davis, 2015). Additionally, communities are upset over a 2012 policy to begin publically ranking school districts and schools by grades. One teacher complained that “schools just can’t be rated like shampoos” and that “worst of all is using these scores not just to rate but to rank, so that the emphasis is on who’s beating whom.” Another teacher said of the rankings, “they’re very accurate measures of the size of houses near a given school,” referring to the link between student poverty and school failure (Postal, 2012).

The most recent – 2015 – changes have not proved any more popular. The new FSA test, meant to replace the FCAT, confused and angered parents, students, and teachers who were not even told what score was required to pass (Veiga, 2015). After many schools struggled with implementing the test and the state delayed score reporting, many districts

sought waivers from accountability. Barbara Jenkins, the Superintendent of Orange County Schools in Orlando sent a letter to teachers and parents that read:

I believe the negative consequences, *both immediate and subliminal*, are without merits. OCPS found a means to hold teachers harmless for student performance portion of their evaluation, and the state suspended mandatory retention for students, *but that does not negate the impact of labeling schools with limited and suspect data.*¹

Following such communications from schools and districts, school grades for the 2014–2015 remained pending into the following school year. When eventually released, they stemmed from a temporary grading formula given the lack of comparable scores from which to calculate learning gains.

These policy reforms and community protests illustrate recognition of the need for change. Yet, while the state has changed the metric for grading schools and even the test used to establish it, possible associations between school rankings and student challenges have not been explored. Such an exploration should bear significance for other states as well. While not every state uses the A to F scale, NCLB does require that states have some kind of evaluation system, what Politico’s Maggie Severns calls “prodding and shaming schools into improvement.”

Some other states have challenged the terms used to label schools. Policymakers and politicians in New York, for example, have spent recent years debating the value of both the policy and terminology responses to “failing schools.” Many superintendents overseeing the state’s “failing” schools argue that their districts are equal to any district in terms of student

¹ Quoted from an e-mail sent to teachers on October 2, 2015; emphasis added

abilities, but are underfunded. They argue that the existing response – to label a school “failing” and enlist a charter staff to take it over – will not solve the underlying problems. They are in favor of other responses, such as a more equitable funding formula (Gagnon & Schneider, 2019).

In late 2015, New York legislators attempted to alter the language of the original reform package – the Education Transformation Act (2015). Assemblyman Angelo Santabarbara and others wanted to replace the term “failing” with “struggling.” Santabarbara said that he attended the Hamilton School – now labeled “failing” – and that “we [at Hamilton] have some of the best teachers in the state...and some of the brightest students...that [failing] terminology is just not fair” (van de Grift, 2009). The amendment he supported stated:

The stigma of being labeled as a failing or persistently failing school only further perpetuates the issues and difficulties these schools face. By changing the designation status from failing and persistently failing to struggling and persistently struggling, this indicates that this is a temporary situation for these schools and that additional management powers and funding will be associated with such schools in order to improve their overall academic performance in order to be removed from these more negative accountability statuses (van de Grift, 2009).

Governor Andrew Cuomo vetoed the amendment legislation, arguing that the original bill’s rhetoric was inseparable from the policy reforms, which he supports (van de Grift, 2009). These concerns in Florida, New York, and other states, are not only relevant to schools and education policy makers. Ladson-Billings (2013) compares the larger education status-quo to the national debt. The federal government often engages in deficit spending and over time deficits form our national debt. She explains her metaphor:

We “budget” – that is, plan – for a certain level of student performance, but students regularly perform at lower levels. The long-term failure to produce equitable conditions to address these deficits creates the education debt. The idea of an education debt is not simply metaphorical. Economists calculate the loss of productivity, the need for remediation, the drain on social services, and the increased costs of law enforcement and imprisonment that result as a lack of educational attainment (p. 13).

The “year to year progress score,” as Ladson-Billings calls it, obsesses the nation and leads to attempts by spectators and participants to address the needs of the “losers” – failing students and schools. There is no mention of how the categorization itself – “losers” – might worsen the situation.

In the coming years, these questions – and this research – will become even more relevant to state education departments. The Every Student Succeeds Act (ESSA 2015), which replaces NCLB, returns much of the power to states in determining how to evaluate teachers, what goals to set for schools, and how to intervene when schools do not meet those goals (Brevetti, 2014; Elmore, 2004; Stretesky & Lynch, 2002; Wright, 2018). Specifically, the law requires states to intervene in the 5 percent of schools that rank lowest, but does not require that they be designated “failing.” Additionally, while test scores do have to be incorporated in school grading metrics, so do other standards such as graduation rates and English proficiency (Elmore, 2004; Wright, 2018). States also have to include at least one non-academic, “opportunity to learn” factor – such as student or educator engagement. These factors have to be independently considered for each student subgroup (Elmore, 2004).

These changes suggest, amongst other things, that Federal legislators and the Obama Administration, recognized many of the arguments posited here: that school evaluation must

be a complex process that recognizes the unique struggles of different students in different schools and that failure to succeed must be met with targeted intervention. As states continue to gain increased freedom to respond to Federal initiatives, they must investigate the impacts of existing school accountability models on both individuals and institutions.

Other academics and state education agencies have explored alternatives to designating failing schools, in some cases relying on teacher and community input. Jack Schneider (2017), for example, together with his research team, conducted focus groups with students, teachers, principals and administrators, and parents and community members from Somerville, Massachusetts. Based on these focus groups, Schneider argues for a framework for evaluating schools composed of five categories. Three of these categories – teachers and the teaching environment, school culture, and resources – are “inputs.” The remaining two – academic learning and character and well-being – are outputs. Some of these categories could be measured using traditional, standardized data, while others would rely on surveys of students and teachers.

In addition to more accurately measuring school achievement, such a system would give way to more nuanced measurements with several advantages over school grades. First, as Schneider (2017) contends, grades or rankings naturally pit schools against one another and suggest that not all schools can be “good.” Massachusetts, for example, divides schools into achievement quintiles; however, Schneider believes this approach “is flawed insofar as schools may not naturally sort into quartiles or quintiles” (p. 193). He points out that rank-ordering schools “makes the lowest-scoring of the ‘good’ schools appear to be far worse than they actually are” and “the top-scoring ‘bad’ school is going to look far better than it should”

(p. 193). The same could be said of the D schools in Florida that miss the threshold for a C by one point.

This illusion of success or failure muddies the supposedly crystal-clear message of grades. Furthermore, different states grade schools in different ways, calling into question whether grades have objective meaning at all. Schneider (2017) explains that in Florida an A school earns at least 62 percent of possible points on the grading rubric; however, in Oklahoma, “the cutoff for an ‘A’ school is 93 – the same cutoff point frequently used by teachers in grading student work – and schools can earn ‘bonus points’ for achieving established criteria in attendance, advanced coursework, graduation, and college entrance exams” (p. 194). If parent-understanding is the argument for grades, then Oklahoma’s 93 percent threshold seems superior to Florida’s 62 (which would earn a student a D or an F in most classes), but the different metrics for the same label undermine the argument that grades serve as objective, universal signals.

As these points suggest, my findings are not only relevant in Florida, but across the nation. While every state’s ESSA plan has been approved, the bi-partisan push for reform to NCLB demonstrates a willingness on the part of lawmakers to correct mistakes. States should take that willingness as an opportunity to craft more nuanced systems of holding schools accountable and publishing their success. Without such action, and without investigations like this one, we might expect to continue our exclusive reliance on measurable outcomes dating back to the early 20th century. While branding schools as failures has succeeded in inspiring some school turnarounds, we must question the educational validity of those turnarounds for student learning. Furthermore, we must examine why some students and

schools remain consigned to failure. Other researchers have illustrated compelling links between demographic and social conditions and academic performance; while policymakers have increasingly focused on how to measure and label that performance. Educators and scholars frequently question the appropriateness of those measures, but we must also question the appropriateness of the labels. This study does that by exploring the effects of school grades on both measurable and qualitative outcomes other than standardized test scores.

Moreover, research in this area must include the voices of teachers, who are most directly responsible for educating children. As those in our classrooms every day, teachers have valuable insight into and information about their students and abilities. Policymakers ought to value such insight. Beyond this seemingly pragmatic fact, there is a systemic need to include teachers in the policy process. As Paige explains – and as I explore more fully in Chapter Five – when teachers oppose or do not understand policies, they are unlikely to implement them effectively and may even undermine them. Jones et al. (2007) cite “lack of respect and recognition” and “lack of influence over school policy” as two of the top five reasons that teachers leave the profession (pp. 137-138). The Orlando Sentinel columnist Scott Maxwell writes that Florida drives 40 percent of its teachers out of the profession because politicians “turn deaf ears to public school teachers and often show disdain” (Maxwell, 2017). I hope that my dissertation serves to amplify the voices of some teachers, but policymakers should not consider my writing sufficient, rather consider it a first step in opening the door and listening to teachers in Florida, and throughout the country.

To this end, the purpose of my study is to explore potentially unintended or unaccounted for consequences of Florida's school grading policy. Primary research questions for the qualitative phase of my research include: How do teachers react to the reception of school grades? How do teachers view the school grading policy? How do school grades impact school culture and policies? The primary research question for the quantitative phase is: What effects do the reception of failing grades have on academic outcomes other than standardized test scores?

Chapter Summaries; Brief Statement of Findings

Governor Bush's A+ policies marked the culmination of decades of increasing school accountability. In Chapter Two – which includes my literature review – I discuss the precursors to Bush's policies, which gradually concentrated power over education in the state's capital of Tallahassee while shifting responsibility onto schools and teachers. The accountability movement, including standardized testing and publicized school rankings, grew alongside efforts to resist desegregation. Progress on the latter front has unfortunately stalled and reversed, a phenomenon discussed at length in my Chapters One and Six. In my literature review, I also look at literature on psychology, organizational culture, and sociology to explore how shame, or stigma, might work to motivate or de-motivate students and teachers. Researchers like Steele (2010) question how labels and stereotypes affect students on an individual level. Steele has found that students stereotyped as low-performing feel such anxiety not to confirm the stereotypes that the opposite result occurs: they confirm

the stereotype. Throughout my dissertation, I wonder what happens when students and teachers experience the stereotype of attending a “failed” school. Does that stigma spur renewed effort and success or does it create whole new levels of performance-crippling anxiety, apathy, or hostility?

In Chapter Four – the first of my findings chapters – I focus on what teachers find misguided about the current grading policy. I begin by establishing some context for their views. Teachers describe their relationship with state policymakers in oppositional terms. Through a combination of a host of lawsuits between the Florida Education Association and the state, as well as the individual statements of teachers, it is clear that educators believe that policymakers do not prioritize the best interests of public schools. Like, Welner and Carter (2013) teachers see increasing top-down, but not bottom-up, accountability. In other words, the state places all of the responsibility for student success on teachers and schools without accepting the reciprocal responsibility of recognizing challenges schools face and providing needed funding, amongst other things.

Educators also feel frustrated by shifting expectations. At the state level, the grading formula for schools has changed nearly every year since grading schools began in 1999; while at the federal level, expectations have shifted from little formal accountability, to No Child Left Behind (2002), to Race to the Top, to the Every Student Succeeds Act (2015). These shifts have created a feeling among teachers that the “goalposts” of success will always be moved. Some teachers and parents who initially supported school grades have lost faith in them after repeatedly feeling on the brink of success, only to have the rubrics changed. Politicians exacerbate this cynicism by publicly praising charter and private

schools, many of which are not subject to the same requirements and limitations as public schools. President Trump and Education Secretary Betsy DeVos, for example, have made publicized visits to private schools while spurning public schools. Some teachers suspect that the hidden purpose of accountability is not the improvement of public schools, but their closure.

It is in the context of this oppositional relationship that I present the issues teachers take with school grades. In Chapter Four, I divide these issues into four themes. First, while grades are supposed to be straightforward and transparent, they are actually misunderstood and mistrusted. Grades have been shown to influence families' real-estate and enrollment choices (Figlio & Lucas, 2004), but teachers suspect that most parents no longer care about grades. Shifting standards have led to apathy. Others fear that grades are – in the words of one principal – “baked-in,” that “bad” schools will always be deemed bad by the community, regardless of grade improvements. Educators fear that parents who do care about grades do not understand how they are formulated. This is problematic insofar as parents assume that schools are monolithic, when in reality, even F schools sometimes have magnet or gifted programs, special clubs, or passionate teachers who might benefit some students.

Second, teachers believe that grades ignore student growth. A quarter of the grading formula does stem from “learning gains” which measure whether students make benchmark gains, or move up a level on standardized tests. However, such benchmarks do not account for where students begin the year. Teachers stress that they can raise their students entire grade levels and still not make “learning gains” if students start at enough of a disadvantage. The frustration of teaching at a struggling school, for some, is that years and years of

significant achievements in student learning still result in an F grade. Florida has resisted federal calls to include English-Language-Learners (ELLs) in school scoring because they start the year at a disadvantage; some argue that the same logic applies to students with poorly educated parents, home or food insecurity, and little exposure to technology, among other challenges external to schools.

Third, teachers feel frustrated that grades do not attempt to measure the talents and efforts of staff and students. A teacher's education and background, experience, hours worked, technical certifications, and club sponsorships all have an impact on students, but are not measured. When students win science championships, debate tournaments, or football games, it showcases their educational experiences, but is not captured in an A to F grade.

Fourth, teachers believe that both the challenges and missions of schools fundamentally differ. Often, the same schools that fail, serve students who lack technology at home, who have not travelled beyond their neighborhoods, who lack a family background in education, or who face violence on a daily basis. Teachers feel driven to respond to these challenges by installing new technology, taking students on field trips, exposing them to the college admission process, and providing social services. These are things that some teachers at A schools admit to taking for granted. Teachers at F schools, by contrast, see a host of socioeconomic challenges that their schools may or may not be equipped to respond to, but they want credit for trying. One teacher describes his job at an A school as “totally different” from the job of teachers at failing schools. This view of education – a mission to respond to the needs of communities – fundamentally differs from a standardized, one-grading-system-

for-all view. I stress that teachers do not want to avoid accountability, but they want a more authentic, holistic accountability.

In the Chapter Five, I highlight the things that teachers want recognized by a school evaluation system. Teachers feel profoundly disrespected by the state and districts in which they work, and in some cases by their own administrators. This disrespect – a feeling of not being considered “professionals” – has given them the belief that no policymaker will listen to their ideas. Most of my participants express surprise that I am interested in their views on school grades. Some tell me that they do not have the “qualifications” – such as a PhD or principal position – to have their ideas listened to. In my experience, this feeling of disrespect stems partly from the reliance of policymakers on researchers and “experts” who have no teaching experience, something that frustrates and angers teachers. It also stems from what teachers see as low pay, unfair accountability, and unrealistic expectations.

These feelings have led to a disconnect between teachers – who view themselves as experts on their field and their students – and the state of Florida, confident that it knows what criteria belong in a school grading policy. Yet, whether they believe anyone will listen or not, teachers do have ideas. First, teachers want school grades to account for or recognize environmental factors. Every day, they witness the ways in which neighborhoods, housing, nutrition, safety, family structures, and other realities of students’ lives outside of school impact their educations. To have these factors ignored and all of the blame for failure placed on their shoulders, is offensive. As one teacher writes in a letter to the editor to the *Tampa Bay Times*:

To suggest that my colleagues and I are less [committed] is a slap in the face and lacking common sense. Does the governor think our jobs are easier teaching students whose families are poorer and less-educated in schools that are older and less-equipped than those “high-performing” schools where our colleagues get bonuses each year? Who would have to be more dedicated to go to work each day? I know for a fact that most of our staff worked hard all year despite huge obstacles only to be rewarded with a low grade based on a test we are not even allowed to see (Robinson, 2002).

In addition to the offense taken, teachers and parents tell me that – by creating the illusion that all schools “compete” on an even playing field – the state obscures the real needs of schools. One parent, for example, tells me that what is most needed in her local public school is psychological services or counseling. Some teachers tell me that they need resources. One school – as of 2016 – is getting a new building and upgraded technology, but teachers lament that for years they have taught in a disrespected, deteriorating school with air-conditioning problems, broken furniture, and outdated text-books. On some of my visits with teachers, I see these things for myself. A simple letter grade, according to teachers, does not communicate these factors, instead it sends a message that teachers, or students, are failures. As another teacher writes the *Tampa Bay Times*, “more affluent schools are rewarded while the poorer children who need more resources are told they are failures...these are children and their families, not defective products going by on an assembly line” (Shubart, 2002).

Teachers suspect that policymakers do not understand the links between demographic, environmental factors and educational outcomes. The state, for example, pursues funding equity by dividing state-wide property tax revenue equally among districts and schools based on student enrollment. However, teachers emphasize that some schools’ budgets are supplemented by wealthy parent groups, fund-raising efforts, and donors. They

speak of schools that have restaurants and swimming pools, versus those that struggle to provide basic accommodations. These inequities contribute to failing schools having difficulty attracting and maintaining a highly-qualified teaching staff with the experience to work with challenged students. Meanwhile, schools that receive A grades receive still more bonus money.

Second, teachers want the state to differentiate grades to account for student needs, the same way they are tasked with differentiating instruction in their classrooms. All of the teachers I speak with mention “multiple intelligences” – either as part of their training or their teaching philosophy. Their idea is that students learn in different ways: some verbally, some kinesthetically, others in reading or writing. Additionally, students have different areas of expertise and skill. Some who perform poorly on math tests succeed on the debate team, or the science club. While only one teacher stresses support for standardized testing, all of the teachers I speak with at least acknowledge that testing has a place. What they ask for is a more “authentic testing” that recognizes the multiple intelligences of students.

Testing is the foundation of the school grades metric, so teacher requests for more authentic testing would significantly impact school grades, but they also want the grades themselves to be a more nuanced measurement. The grade might, for example, measure opportunities for students such as a debate or science team, and the success of those programs. It might consider students at different levels – standard, honors, or advanced – differently. One teacher tells me that the scripted lesson plans she receives from the county outline different learning strategies and mastery goals. She feels incredulous that her school administrators told her “we’re not about mastery, we’re about exposure.” She sees this as a

betrayal of the school's stated educational philosophy and aim to help kids. The betrayal, as she sees it, is in pursuit of a higher school grade by any means necessary.

Third, teachers want grades to be understandable by and useful for teachers, students, and families. Most take offense to the idea that the state or county knows better than they do how their students are learning. Teachers do not believe that a test can say more than their daily activities with students. Likewise, they do not believe that a letter grade can say more to the public than interactions with parents. One principal in particular stresses her communications with parents and open-house nights as the most important tools for building trust with the community. She praises the school's "familial" atmosphere. Others mourn the lost intimacy of a "community school" that kept parents informed organically through school and community events.

Teachers feel suspicious that a school grade could ever be useful. As one community member writes of school grades, in a letter to the editor to the *Tampa Bay Times*:

The teachers don't want it, the parents don't want it, the students certainly don't want it, yet we have let our state government cause incredible pain to the public school system in Florida...we have given one test, FCAT, the power to destroy an entire school with a failing grade. Testing should be diagnostic in nature, not punitive. And certainly not humiliating to the citizens of Florida (Smith, 2002).

Something "punitive" and "humiliating" with the power to "destroy an entire school" seems beyond redemption, which is how teachers see grades. However, several would like more detailed data from the state that is diagnostic. One principal emphasizes that the VAM – or value-added-model – data that teachers receive is a "secret science" instead of a useful tool. VAM data factors into teacher evaluations based on the same standardized test results that

inform school grades, yet teachers do not know how the number is calculated or what exactly it means, and they receive it into the fall of the following school year, after they have already begun to work with a new group of students.

In Chapter Six, I explore segregation as an increasing problem in Florida's schools. While Florida long resisted desegregation orders throughout the second half of the Twentieth Century, what little progress had been made is being reversed (Borman & Dorn, 2007; Cobb-Roberts & Shircliffe, 2007; Orfield & Ee, 2017). Some drivers of this reversal include the growth of charter schools, which often draw primarily students of color away from, or replace, more integrated public schools, and the letter grade system itself, which teachers accuse of driving white and affluent families from schools with lower grades. Teachers believe that race and class are deeply entwined, what Orfield and Ee (2017) call "double-segregation." Florida's failing schools serve majority black and Hispanic students, with higher rates of free-and-reduced lunch status.

The teachers I speak with tell me that the governor and legislature's solution is school choice – scholarship opportunities for students to attend private schools and the competition of charters. This philosophy inspired legislation such as the 2017 "Schools of Hope" bill, which redistributes money to charter schools that open near failing traditional schools. Teachers accuse programs like this of robbing their schools of much needed resources. They point out that while it is logical for money to follow students, that logic does nothing to fix their buildings, upgrade technology, or improve transportation, among other needs.

Some highly-graded schools are accused by teachers and the media of intentionally pushing students at-risk of not graduating towards charters and adult-education programs to prevent them from having a deleterious effect on the school grade. Unfortunately, these charters often lack sufficient funding and learning conditions, and do little to track student progress. They, along with private schools, lack the oversight and requirements that public school teachers say burden their schools.

Teachers want even accountability for these schools, but they also want accountability that is reflective of the unique needs and challenges. Teachers insist that the state should not ignore demographic factors such as race and poverty. If a school's needs are unique, then so is its mission; thus, it ought to be graded differently. By insisting that the only problem plaguing disadvantaged students is low expectations, as former governor Jeb Bush has argued, policymakers fail to consider the type of interventions that might actually help students.

In Chapter Seven I discuss the feelings evoked by school grades, possibly the section of my findings least explored by other researchers. Part of the reason feelings about school grades go relatively undiscussed is that they are uncommonly used. While every state has some kind of evaluation label for schools, only 15 states assign letter grades that mirror student grades. According to the Florida Education Association, via Frontline: “Part of the reason that 70 percent of states don’t grade schools is because ranking schools on an A-F system has few, if any, tangible benefits for students” (Frontline, 2019). While it may seem like a trivial issue to some, I argue that the ways we measure schools have profound impacts – good, bad, or mixed – on teachers and students.

The media as well as other researchers have referred to failing school grades as a “stigma” (Bowen & Trivitt, 2014; Chakrabarti, 2013). For the most part, teachers confirm this association. One principal describes receiving a failing grade as “a punch in the gut” that seems to ignore all of his staff’s efforts and successes. While some teachers report that an F grade makes them work harder, other teachers leave or refuse to work at failing schools in the first place. This is especially true for teachers who teach “core” subjects like math and language arts; these teachers receive a disproportionate pressure from administrations.

While the media does not work as a source of shaming, as I initially suspected it might, teachers at failing schools do report seeing their schools represented in the media often, usually in negative terms such as crime coverage. They worry that such coverage scares away talented teachers and students. They want the media to paint a more three-dimensional, realistic portrait of their schools that includes acknowledging the heterogeneity of study bodies. Covering programs, such as advanced placement classes and successful extracurricular activities, could provide a more nuanced picture of a campus. Educators also think that the media could attract attention to the causes of failure, to provide the public with context and highlight possible solutions.

In Chapter Seven, I also question the position of the Florida Education Association that school grades do not benefit students. One of the schools at which I interview teachers and principals recently rose from a decade of F and D grades to a C grade. The former principal tells me that after receiving yet another F grade, he and his staff collaboratively came up with a plan of action. Defenders might use such a case to say that the stigma of the F

grade served its purpose by inspiring teachers to work harder, implicitly suggesting that they were not working hard before.

In this case, it is unclear whether the turnaround stemmed from the F grade inspiring harder work, or to significant interventions by the county. The county offered bonuses to attract qualified teachers, built the school a new campus, and upgraded technology. All things that teachers suggest to me might help failing schools. My qualitative data collectively suggests that the letter grade was irrelevant. There are startling correlations between school grades and income brackets. I argue that rather than ignore socioeconomic status, counties and the state might intervene in lower-income or highly segregated schools before they undergo a decade of failure. Such intervention would arguably provide a decade of children with better educations. It is also questionable whether interventions – even if they raise the grade – are genuinely beneficial to teachers and students. Teachers – even those who support some standardized benchmarks – make clear to me that they do not believe state tests correlate with classroom success. In short, the state and counties of Florida ought to reconsider the timing, causes, and methods of intervention in schools.

Teachers have different beliefs about student reactions to grades. Some think that students are not aware of the school's grade. Others think that students are highly aware. This difference may correspond – at least in part – with a student's age (Jones, Jones, & Hargrove, 2007). For students who are aware, it remains unclear how they feel affected – if at all. According to their teachers, students who perform poorly may be affected by testing in four ways. As pressure “trickles down” – from the state, to the superintendent, to administrations, to teachers, and finally to students – it may increase already-existing anxiety, for both

teachers and students. As teachers feel more anxious, they may teach less effectively, or even say inappropriate things. Teachers disagree about the value of pressure: while some think that students should not feel burdened by their school's grade or teacher evaluations, others think that students need to understand those connections more.

Some believe that students feel exhausted from an over-abundance of tests. As they take more and more, they might lose stamina, or might come to view tests as unimportant. Some students grow fatigued quickly, especially when tests require long reading passages. These students tend to "Christmas-tree" answers in order to get out of the test as quickly as possible. Some do this without foresight of the consequences for their schools, teachers, or their enrollment in remedial classes. Teachers believe that this is true of a majority of students – that they do not try. A couple of teachers blame laziness, but most tell me that apathy is a way for students to maintain agency. If they try and fail – they feel dumb. If they refuse to try – they feel rebellious, and in control.

Finally, teachers believe that low expectations – especially at the school level – form a kind of self-fulfilling prophecy. Teachers at failing schools can come to expect less of their students. Inappropriate behavior during tests is accepted, as are low scores. Likewise, students can impose negative stereotypes on themselves or others. These beliefs call for a different kind of intervention, clearly teachers believe that more testing practice, scripted lesson plans, or writing learning targets on the board do not work to inspire struggling students.

In Chapter Eight, I employ quantitative methods – namely a series of fixed-effects

regression models – to analyze possible impacts of receiving a failing grade on schools’ teacher separation, attendance, drop-out, and graduation rates. Fixed-effects models control for the relatively static variables of schools, such as demographic make-up, and use within-school variance to test differences in outcomes after an F or D grade (Mehmetoglu & Jakobsen, 2017).

Overall, the results of these regressions are more ambivalent than the strong statements made by some teachers in interviews. For the most part, teachers do not suspect that the school grade impacts the attendance rates of students. This is born out by the data. There is not sufficient evidence to reject the null hypothesis that attendance rates are similar before and after the reception of a failing grade. This is also supported by between-school comparisons. The mean attendance rates at both failing and non-failing schools are similar and desirably high, between 94 and 95 percent.

Results do not suggest a significant correlation between the reception of a failing grade and teacher separation rates. While there is some modest evidence suggesting that a failing grade may contribute to a slight increase in teacher separation, findings seem at odds with the statements of some teachers and both principals I interview – that failing schools have a hard time attracting and retaining staff. I suspect that unavailable explanatory variables may have more to do with retention than the grade itself. Perhaps the stigma of an F that I feared warded off teachers is less of a driver than factors of location, income, race, discipline, and administrative support, among others. It is possible, as some of my teacher participants attest, that while a failing grade repels some teachers, it attracts others who feel a

calling to work with struggling students.

There are concerning trends in teacher separation rates across all schools. Between 2010 and 2014, teacher separation rates increased from a mean of 5.6 to 8 percent. The data does not reveal why this occurred, but teachers suggest in interviews that policy upheavals, increased accountability, and decreased respect for teachers play a role. Whatever the reasons, failing schools clearly lose more teachers. The mean teacher separation rate of F schools is nearly double that of A schools (roughly 11 versus six percent). Also worrying, the teachers who leave failing schools have less cumulative experience than their peers at higher-achieving schools. Interviews suggest that failing schools either hire less experienced teachers or cannot maintain staff long enough for them to develop more experience. Either case is problematic. Principals lament devoting energy and resources to train new staff, only to see them depart after a year or two. Experienced staff could potentially be more effective on their own, as well as help to assist and train new teachers. I argue that for these reasons and others, it is impossible to evaluate the quality of schools without considering staff experience, training, retention and morale.

In regards to failure rates, the inverse of graduation rates, there is evidence suggesting that school grades accurately capture decreases in graduation rates. Disturbingly, however, the evidence also suggests that two years after a failing grade, the failure rates of black and Hispanic students continue to increase. In Chapter Six, I explore possible reasons for this. It is possible that the interventions of counties in failing schools may initially help students of color to stay in school and graduate, but that these efforts lack sustainable impacts and so –

two years after a failing grade – graduation rates decline once more.

There are other worrying trends regarding graduation and failure rates. While graduation rates are relatively high, there is a notable achievement gap between white and Asian students and their black and Hispanic peers. The failure rates of the latter groups appear to be either negatively impacted by a failing school grade, or to be unresponsive to county and state interventions. As expected, for all subgroups there is also a gap between students at failing schools and non-failing schools. If one of the aims of school grades is to drive failing schools to improve in this and other measures, then grades seem to have been ineffective between 2010 and 2014. In fact, graduation rates actually fell in 2011 and otherwise remained stagnant in the five-year sample.

Finally, I look at drop-out rates. Oddly, drop-out rates do not appear to be associated with the reception of a failing grade. School grades partly rely on graduation rates, yet drop-out rates seem to have no significant effect. This implies that drop-out rates are not an inverse of graduation rates, which the state Department of Education confirms. I argue that this is somewhat misleading, especially since teachers and the media suggest that some schools hide drop-outs and failures by encouraging students at risk of not graduating to attend adult GED programs or charter schools. Policymakers should be conscious of these potentially “hidden” drop-outs.

Between school data does show that failing schools have average drop-out rates nearly double those of non-failing schools (2.3 versus 1.2 percent). While this divide is concerning, it is positive that both rates are – assuming their accuracy – quite low. Moreover,

drop-out rates increasingly fell between 2010 and 2014. Mean rates decreased from 1.4 percent in 2010 to 1.15 percent in 2014 – a small change but in the desirable direction.

Collectively, these chapters illustrate striking disconnects between educators, parents, and policymakers at the state and county levels. Based on prior research, media accounts, and their own statements, the politicians and policymakers behind Florida's school accountability system see the primary challenges of education as ineffective teachers, a stagnating education association, low expectations for poor and minority students, and a lack of choice for parents. Consequently, they see the solutions as stricter accountability measures for teachers, including merit pay, weakening the Florida Education Association, holding all schools to the same expectations, and promoting school choice via funding for charter schools and scholarships for students to attend private schools.

Educators view these positions as parts of a war on public schools. They consistently tell me, in interviews and in conversations during my years as a Florida teacher, that Florida politicians want to undermine and hopefully close public schools in order to privatize education. They feel cut out of the policymaking process, which only increases their antipathy for school grades. Teachers and principals see the primary challenges of education as insufficient funding for public schools, particularly those serving students of color and socioeconomically challenged students, a lack of respect and trust for the expertise of teachers combined with unrealistic expectations, standardized testing that limits rather than enhances classroom teaching, and a failure to address systemic issues beyond the control of schools, such as housing and family security. All of these problems are exacerbated by

growing racial and class segregation.

Much of the energy of parents, academics and education reformers has targeted the presumed harms of standardized testing and top-down school accountability. In Florida, these efforts have had some success. Former governor Rick Scott signed several bills into law that initiated studies of the validity of standardized tests. He supported the views of some parents and teachers that students spent too much time testing and he followed through on eliminating a couple of required end of course exams (Sherman, 2017). Despite progress like this, as an educator and researcher, I worry that political expediency will ensure that governments continue to value simple metrics, easily publicizable results, and a veneer of accountability. I endorse the views of Fullan and Quinn (2016), who rhetorically ask, “why politicians endorse solutions that don’t work.” They answer:

The answer is not complicated: because they can legislate them; because they are in a hurry; because the remedies can be made to appeal superficially to the public; because (and unkindly on our part) some of them really don’t care about the public education system, preferring that education be taken over by the private sector; and (more kindly) because they do not know what else to do (p. 3).

If true, we should not expect the wholesale reversal of the standardization trajectory, no matter how opposed to it some teachers stand. So, I often ask, how can we mitigate some of the harms of education “progress” in America? If we must have one scale by which all school performance is measured, how can we ensure that said scale helps more than it hurts? Can we make it useful and meaningful for educators and parents alike? Can it be as easily understood as an A to F grade, yet nuanced enough to communicate the real strengths, challenges, and progress of a school? If part of the problem is, as Fullan and Quinn (2016) conject, that

politicians “do not know what else to do,” then one answer is to offer ideas and command the state’s attention. I attempt to do so in the final section of my dissertation: Conclusions and Policy Ideas.

While the labels and metrics used to grade schools may seem like a relatively small matter in the scope of school reform, my findings suggest that the labels matter. My quantitative findings suggest that school grades’ importance is not as simple as cause-and-effect, but I argue that makes them no less significant. Instead, we should see them – in contrast to the words of Former Governor Bush and others – as intricately linked to socioeconomic, racial, and other demographic factors. My qualitative findings confirm that grades matter to parents and teachers, and as I discuss in the conclusion of Chapter Two, grades have created problems and confusion in other states that use them. Moreover, letter grades affect perceptions in ways different from other kinds of school ratings (Jacobsen, Snyder & Saultz, 2014). In this context, we should consider grades a minute but important matter, one realistic to tackle, and a way to make the road of education reform in all states as helpful – and harmless – as possible for students, teachers, and families.

CHAPTER 2

LITERATURE REVIEW AND THEORETICAL FRAMEWORKS: IMPACTS OF GRADES ON MOTIVATION AND COMMUNITY OUTCOMES

Since the 1983 publication of *A Nation at Risk: the Imperative for Educational Reform* by the National Commission on Excellence in Education, state and federal education policy in the United States has gradually shifted towards centralization, standardization, and accountability (W. Au, 2010; Goldstein, 2014; Hamilton, 2003). The 1999 passage of a bundle of education reforms in Florida – the “A+ Plan” – marks an important moment in this shift. At the time, it set a model for the Bush Administration’s No Child Left Behind (NCLB) reforms and went on to influence the policy reforms of other states (Borman & Dorn, 2007; Herrington, 2005; Jones, Jones, & Hargrove, 2007; Ujifusa, 2012). Central to the A+ Plan is an evaluation system that ranks both students and schools based primarily on standardized test scores (Gershenson, Holt, & Papageorge, 2016). Schools, like students, are assigned letter grades between A and F, the logic being that the stigma attached to failing grades should motivate schools to improve (Bowen & Trivitt, 2014; Chakrabarti, 2013; Goldhaber & Hannaway, 2004; Jones et al., 2007; Michael & Dorn, 2007).

The A+ Plan and subsequent Florida reforms have sparked controversy with teachers and parents (Harrison & Cohen-Vogel, 2012; Maxwell, 2012) as well as researchers who fear that testing narrows curriculum, increases student anxiety, widens achievement gaps, and fails to accurately measure student ability (Agee, 2004; Amrein & Berliner, 2002; Jones et al., 2007; Linn, 2008). Despite controversy, some evidence suggests that the A+ Plan has worked. When Governor Jeb Bush took office in 1999, Florida schools were frequently rated among the nation's worst. A decade later that had changed dramatically. In 2010, the conservative American Legislative Exchange Council ranked Florida schools third best in the nation; in 2011, Education Week ranked them fifth (Foundation for Excellence in Education, 2012; Miller, 2011). Greene (2001) finds that failing schools in particular have showed improvements. Between 1999 and 2000, those schools whose students would have been eligible for vouchers after a second failure improved their scores by more than twice the margin of higher-scoring schools. Likewise, Rouse et al. (2007) find that schools with an F grade in 2002 adopted new instructional policies that had a lasting positive impact on student performance.

Some detracting evidence, from across the country, suggests that these improvements are superficial and not due to educational quality, but instead to types of curriculum perversion – such as a near-exclusive focus on a particular subject or subset of students – or to inaccurate score-inflation (Chakrabarti, 2013; Jones et al., 2007; Linn, 2008). For example, Goldhaber and Hannaway (2004) observe one school that focused exclusively on writing, since at the time passing at least one of the primary subjects – math, reading, and writing – assured the school a passing grade. As the school's principal explains, “with writing there's a

script” to follow, meaning writing was the easiest subject to coach (p. 602). The observed school did increase its pass rate from 17 percent to 71 percent – but still failed in reading and in math. This exclusive focus on certain subjects can, critics argue, poorly prepare some students for future academics and alienate others who see their favorite subjects eliminated from the schedule (Agee, 2004). Others have found evidence that schools increase disciplinary suspensions during testing periods to prevent characteristically low-scoring students from dragging down aggregate scores (Greifner, 2013). Additionally, some teachers practice types of “unethical test preparation” such as giving students answers or changing answers for them (Jones et al., 2007).

Disregarding these debates over the merit of individual school reforms, supporters tout the A+ Plan as a policy success. However, it fell short of the No Child Left Behind (NCLB, 2001) mandate that all students be proficient in reading and math by 2014, and some schools in particular have remained stubbornly low-achieving. Between 1999 and 2011, of the 552 high schools in Florida, 318 received a D and 73 received an F at least once (Florida Department of Education, 2011). In 2013-2014, the number of failing schools rose 68 percent over the previous year to include 178 schools statewide (Kurtz, 2018). In 2014-2015 the failure rate grew even more, out of 3219 schools, 365 received a D and 184 received an F (Ballou, Sanders, & Wright, 2004). These numbers suggest that, while grades for most schools have improved, some schools remain consigned to failure. Struggling schools are often concentrated in particular areas, including the most urban or rural where incomes, populations, or both are lowest (Lee, Borman & Tyson, 2007). If A+ truly is a blanket success, why is it not working for all schools?

In the past decade there has been a healthy national debate about whether the struggles of certain schools and students are attributable to differences in resource allocation (Grodsky, Warren, & Felts, 2008; Jimenez-Castellanos, 2010; Mehmetoglu & Jakobsen, 2017), a history of racial discrimination (Cobb-Roberts & Shircliffe, 2007), family supports and home environments (Jacob, 2004; Schwartz, McCabe, Ellen, & Chellman, 2010; Yoder & Lopez, 2013), or – as Governor Bush argued – poor standards that underestimate the abilities of disadvantaged students (Borman & Dorn, 2007; Cobb-Roberts & Shircliffe, 2007). Researchers have also investigated, with mixed findings, whether the stigma of a failing grade and the vouchers that originally accompanied failure in Florida, spur school reform (Bowen & Trivitt, 2014; Figlio & Rouse, 2005; Greene, 2001). Figlio and Rouse (2005) concluded at the time of their research that – in Florida – the threat of stigma associated with an F was more clearly associated with school reform than voucher threats.

Most of the research in this area has focused on the impacts of voucher threats or stigma on individual and aggregate student test scores (Chakrabarti, 2013; Gershenson et al., 2016); additionally, some has focused on school-level policy changes (Goldhaber & Hannaway, 2004; Rouse et al., 2007). We have little or no understanding of how the stigma of a failing school grade affects students and teachers as individuals and members of their schools – beyond test scores. Further, while some studies have controlled for student-level factors, we lack sufficient measures of a failing school grade's impacts on student- and teacher-level non-test outcomes such as teacher turnover, student discipline infractions, and drop-out rates (Jones et al., 2007). We have measurements of these variables; yet, most state

policymakers do not seem to consider these in connection with the assignment of school grades.

Literature Review

Much has been written on the reliability of standardized tests as measures of student and teacher performance, as well as their impacts on students, curriculum, and teachers. A substantial amount of work has also been done linking the assignment of school grades to various school and community level outcomes. However, we know little about the impact of school level grades on students themselves and even less about the impacts – excluding instructional practices – on their teachers. In this section, I will briefly summarize some of the research on the effectiveness of A+ in turning schools around, the perceived reliability of standardized tests, the feelings of parents and teachers, and the effects of school grades on community and school outcomes.

A+ and School Reforms

In recent years, state policymakers have become increasingly interested in different models of school and student accountability. Typically, status models are compared to a range of growth models. In a status model, student scores are compared to a target score and potentially to another cohort of students. For example, one might ask: did tenth graders this year meet our target score – and – how did tenth grade scores this year compare to tenth

grade scores last year? Growth models, which are required post-NCLB, measure changes in performance from year to year by the same cohort of students, usually with both a target score and a target level of improvement (Goldschmidt et al., 2005). Florida employs both, with a portion of school grade points going to student proficiency and a portion going to student growth, both for students at large and the lowest 25 percent.

Prior to the 2014-2015 school year, student performance was measured by the FCAT, which comprised about half of a high school's grade. Additionally, schools earned points from their graduation rates, SAT results, student enrollment in advanced classes, and student earnings of college credit (State Impact, 2014). Most consider the diversification of this formula to have been a positive step – although it has played its own role in entrenching inequalities. For example, rural schools with smaller enrollments cannot afford to offer as many advanced placement or college-prep classes, which means they struggle to earn A's (Gonzalez, 2012).

Despite these challenges, the majority of schools seem to be working hard to increase their school grades. Rouse et al. (2007) posit that the stigma of receiving an F grade inspires school-level policy changes that are associated with increases in student achievement. To reach this conclusion, they conduct a cross-sectional regression analysis of fifth grade students in 2002-2003. They find that not only are the policy responses at F schools significantly different from those of other schools, but also that students attending F schools earned higher gains than students at all non-F schools. In order to control for the possibility that score-increases were the result of curriculum perversion or “teaching to the test,” Rouse et al. (2007) compare student scores on the relatively high-stakes FCAT with the relatively

low-stakes Stanford-10 Assessment. While these increases were associated with narrowed curriculums, F schools also adopted policies such as block scheduling, increased class time, and the provision of additional resources for teachers. Rouse et al. (2007) thus argue that a significant portion of score increases can be attributed to these policy-changes.

While he generally seems to agree, Betebenner (2008), in a review of the Rouse et al. (2007) study, points to some worrying flaws in its argument. First, he argues that because the sample size of persistent F schools was so small, it is difficult to attribute gains in scores to the ongoing stigma and pressure to reform associated with a failing grade, or to policy changes other than test-prep. Second, Betebenner (2008) explains that if vouchers and other accountability measures do promote changes in policy, policymakers still have no way to single out what policy change(s) prompted the effect. I might additionally argue that the comparison between FCAT and the Stanford-10 Assessment is hardly a reliable control, since “teaching to the test” curriculum focus would likely improve performance on both standardized tests.

This uncertainty gives way to Bowen and Trivitt (2014), who seek to determine whether the private school vouchers once associated with A+ or the stigma associated with a failing grade had the greater impact on school turnarounds. Their study compares changes in student performance pre- and post- the 2006 Bush v. Holmes decision to overturn the school voucher component of A+. They find that elementary students attending schools that had recently been graded F made statistically significant improvements in math and reading. However, the interaction between a failing grade and the threat of a school voucher showed no impact on math scores and a statistically significant decrease in reading scores. One of

their hypotheses to explain this surprising decrease is that parents give up on failing schools when they know vouchers offer a way out.

Aside from the tenuous link drawn by these studies between sanctions and increases in test scores, existing research tells us little about how individual students and teachers are affected by school grades, nor does it attempt to measure non-test outcomes. It is assumed that accountability mechanisms only affect the standardized test scores and associated policies implicit in the school grading formula, without accounting for the possibilities of teachers leaving, students dropping or acting out, and increased suspensions and expulsions that remove problem students. Furthermore, while some studies (Figlio & Rouse, 2005; Rouse et al., 2007) only allude to it, other research strongly suggests a reality that turnarounds result not from true educational improvements but from curriculum perversion and narrowed learning (Chakrabarti, 2013; Goldhaber & Hannaway, 2004).

Community Reactions

Much of the existing research in this area is concerned with the accuracy of standardized tests at measuring student learning and teacher effectiveness. While I share these concerns, the primary focus of this research is not on the reliability of tests, but rather on their acceptance as academic indicators by students, their families, and teachers who may react in unexpected ways to the assignment of a failing school grade. In other words, whether a student's score is consistent with his academic abilities or not, he will still be attending a

school designated a failure and he may still know that his individual score makes him, to some degree, complicit in that failure.

My research also considers whether students and teachers in Florida's failing schools experience feelings of shame or stigma associated with their schools. The best indicator for such qualitative measures may come in Goldhaber and Hannway (2004)'s case studies of five schools in Florida, three of them graded A and the other two F. They find that both A and F schools faced substantial pressure to perform due to the "easily understood" grading system that "carries a certain amount of symbolic value" (p. 601) and significantly narrowed their curriculums and student activities as a result.

At A schools, principals report that parents cared about school grades and apply substantial pressure on principals and teachers to "do whatever it might take to perform well on the test" (Goldhaber & Hannaway, 2004, p. 601). This does not necessarily coincide with a belief in the accuracy of the FCAT. One principal tells researchers that "even though they [parents] know better than to believe that any one test can show how good we are, they still want us to get an A" (p. 601). Nor do parents want excess time committed to test preparation, as in the case of one elementary school that cancelled all field trips and projects scheduled prior to the FCAT. Yet, "while parents lamented this and knew their own child's life chances would not be affected by his or her test results, they nonetheless claimed to be 'type A personalities' who 'still wanted to see their school with the highest grade'" (p. 601). One school dropped from an A to a C and its principal had to "expend considerable effort" (p. 601) convincing parents that it was still a good school.

Schools were not only pressured by parents, but by other schools as well. Goldhaber and Hannaway (2004) report that schools felt pressured to perform well by their district administrators, as well as via competition with other schools that were often “envious” and “wanted to see the [A] school stumble” (p. 601). This pressure translates to individual teachers and students. Parents cry in interviews with researchers and one parent says of their daughter, “[she] gets so nervous at testing time that she pulls her hair out” (p. 601). A teacher calls the pressure placed on testing “the most criminal form of child abuse that I have seen” (p. 601). At these – already successful – schools, policy reforms included daily countdowns to the FCAT and, at one elementary school, a “lock-in” before the test so that children could study through the night.

At F schools, Goldhaber and Hannaway observe that teachers and principals feel more positive about the grading system, but do not face as much pressure from parents. One principal says that the parent reaction to the school’s failing grade was “almost nonexistent” (p. 602). This is consistent with the suggestion by Bowen and Trivitt (2014) that parents at such schools give up. One reason for apathy may be doubt as to the accuracy of the FCAT and school grades, or outright hostility to them as assessment standards. Also likely, parents at these historically socioeconomically-challenged schools may face outside pressures – such as work and transportation – that limit their involvement in school.

Community Outcomes

While research is limited on the effects of school grades on non-academic outcomes (Jones et al., 2007), there is evidence to suggest that grades do have reverberations in the larger community, most notably on housing values and neighborhood schools. Evidence further suggests that these outcomes disproportionately affect neighborhoods of color and have the cyclical effect of diminishing funding for already struggling schools (Booher-Jennings, 2005). If such outcomes can be observed, then we have an impetus to investigate other non-score outcomes.

Figlio and Lucas (2004) observe a strong correlation between the assignment of school grades and the housing market. After the first year of assigned grades, controlling for other property attributes like neighborhood and geographic region, they find that houses zoned for A schools were appraised 19.5 percent higher than those zoned for B schools. Likewise, houses zoned for B schools were valued 15.6 percent higher than those zoned for C schools. This also corresponded to real estate prices. Houses zoned for A schools averaged nine percent higher sale prices following grade assignments than did houses zoned for B schools.

These findings suggest that, regardless of whether or not parents recognize school letter grades as academic indicators, some do value them enough to factor into housing decisions. This presents its own problems, such as obstacles for lower class students to attend high quality schools. While Florida funds schools equally by population, wealthy schools can supplement their budgets via parent fundraising or private donors, making funding gaps

persist (Loubert, 2005; Ready, 2013). Across the country, such funding gaps continue to increase (Barshay, 2015). Furthermore, due to connections between race and class, schools primarily serving students of color are the most likely to suffer. This plays into a long history of punishing majority-black schools in Florida, including the closing of historically black schools to satisfy white parents during de-segregation orders (Cobb-Roberts & Shircliffe, 2007).

Defenders of Florida's policies argue that parents can determine where to move based on school grades. Yet, as Orfield (2013) explains: "Although this is generally accurate for white and many Asian families, black and Latino families often face discrimination at many stages of housing choice" (p. 45). This discrimination leads to "functioning middle-class schools for white and Asian families and segregated schools with substantial concentrations of severely disadvantaged children for black and Latino students" (p. 45). Although well-intended, school choice options – dating back to the 1960s – have had the counter-effect of further segregating schools by encouraging the flight of white and otherwise affluent families from primarily black and Latino schools.

Desegregation Versus Accountability

While school choice has been shown to exacerbate segregation (Orfield & Ee, 2017), school accountability and integration do not sound like mutually exclusive goals. Unfortunately, the early school choice plans, in the 1960s, were often designed by

segregationists who wanted to help white families leave integrated neighborhoods. Today, white and affluent families not only have access to greater resources, but have relatively more knowledge and connections in choosing neighborhoods and schools (Orfield, 2013). This disparity exists partly because over the past two decades, policymakers have worked – intentionally or not – to erode the progress of desegregation. Goldstein (2014) persuasively argues that the United States never really committed to integration, and instead – particularly in the 1980s – turned towards a teacher accountability agenda. By focusing explicitly on the test-score achievement gap between white students and students of color, policymakers gravitated towards a “no excuses” reform agenda that paved the way for leaders like Jeb Bush to argue that the biggest roadblock to struggling students of color was low expectations.

When accountability laws gained prominence in the 1970s – which Goldstein (2014) points out was a “time of antitax sentiment” – poor funding meant that “competency laws were often unfunded mandates with little or no money provided to actually implement new ways of training and evaluating teachers” (p. 168). The commission that published *A Nation at Risk* identified multiple priorities including rigorous high school curriculum, a longer school day, an increased school year, and “encouraging the federal government to play a larger role in setting the national education agenda and funding it” (p. 170). Goldstein continues:

But the report was released into a political climate of budget-cutting fervor—federal aid for poor children’s education was cut by 6 percent in Reagan’s first term—and a culture war. There was little enthusiasm in Congress for providing the massive influx of funding needed to extend learning time, the costliest proposal in *A Nation at Risk*. The idea of more rigorous, universal curriculum standards was also a nonstarter; the American Right had long demonstrated an overactive paranoia about supposedly

liberal national attempts to influence the curriculum of local schools—a paranoia Reagan fully shared as a former anticommunist activist. Consequently, despite *A Nation at Risk's* broad set of recommendations, policy makers focused increasingly on teachers alone: their training, demographic traits, and how they were evaluated and paid” (pp. 170-171).

So, opposition to desegregation and public spending on education helped give birth to an accountability movement that increasingly restricted accountability to a narrow set of actors: teachers and schools.

Kantor and Lowe (2013) argue that No Child Left Behind continued this trend. They write:

It [NCLB] was attractive, for example, to many suburban Whites who favored expanding educational opportunities for the least advantaged while preserving their own access to good schools. Yet, because it effectively shifted the blame for educational failure from the child to the school, it was also popular with some civil rights organizations, Black and Latino parents, and liberal advocacy groups. These organizations and groups believed that accountability offered a more robust sense of opportunity than Title I, which had been framed in a language of cultural deprivation that blamed poor children for their own educational failures (p. 37).

These quotations summarize a disturbing progression that shifted responsibility away from governments that had sanctioned segregation and consigned students of color to the poorest neighborhoods and schools, to the culture of the students and families, and finally to teachers and schools. By framing accountability and desegregation as mutually exclusive goals, policymakers have done little to address the disparities in home-educational resources, environmental concerns, housing security, medical problems, poor nutrition, unemployment, and the availability of successful role-models (Orfield, 2013). Orfield (2013) concludes that:

While test scores are strongly related to family resources and peer-group influence, low-performing minority schools are disproportionately branded as failures. Generally, teachers are blamed for their schools' failure, and many view such evaluation policies as deeply unfair. Consequently, frustrated teachers tend to move away, leaving schools where they are unfairly blamed for schools where they are credited for the success of more-privileged students. Schools with high concentrations of students needing strong academic support are often staffed largely by inexperienced teachers who are not yet effective educators, and some of them do not want to be there" (p. 41).

Teachers who do stay at segregated and struggling schools often face constant, unrealistic pressure which they report internalizing and passing on to students (Jones et al., 2007).

Summary

Research has found that the system of school grades – primarily through the stigma associated with a failing grade – has yielded some success in turning schools around (Figlio & Rouse, 2005; Goldhaber & Hannaway, 2004). However, there is some conflicting evidence as to whether turnarounds are due to curriculum perversion or to more holistic reforms (Goldhaber & Hannaway, 2004). The FCAT, on which school grades have been primarily based, has not been accepted as a genuine measure of student learning by many parents and teachers (Borman & Dorn, 2007; Jones et al., 2007). However, both groups still place immense pressure on schools and students to perform at the highest level (Rouse et al., 2007).

Some students and parents report high levels of anxiety associated with this pressure. Other parents, primarily at failing schools, appear to be apathetic and withdrawn from school (Goldhaber & Hannaway, 2004). One hypothesis is that they lose faith in the school system

altogether, especially given the – at least theoretical – possibility of changing schools. This can cause resentment among teachers who feel parental involvement is vital to improving grades. For this and other reasons, high quality teachers may be reluctant to work at the schools needing the most help (Johnson, Kraft, & Papay, 2012; Schwartz et al., 2010). In addition to such school-level impacts, school grades show evidence of impacting certain community outcomes such as home values and the closing of schools and subsequent rezoning of students, especially students of color (Figlio & Lucas, 2004). These relocations can affect the funding and success of neighborhood schools (Booher-Jennings, 2005). This calls for an exploration of other possible outcomes including student motivation, drop-outs, attendance, and teacher turnover, especially in light of the disturbing degree to which students of color are segregated in the most challenged conditions.

While this study has national implications, much of the research I have found and focused on here is Florida-based. As explained elsewhere, Florida set a model for the national requirements of No Child Left Behind and for the reforms of other states. I think that for this reason, other researchers have also focused on it. Furthermore, having instituted grades far earlier than most (in 1999), Florida has a longer history of data and policy changes from which to draw.

Theoretical Perspective

My hypotheses rest on three theoretical foundations: motivational theory; Creed et al.'s (2014) organizational “shame nexus”; and Steele's (1995) theory of stereotype threat.

Together, these three foundations suggest the following: first, a school accountability system that relies on an external motivational mechanism – school grades – may contribute to student apathy or anxiety. Second, the assignment of failing grades to schools and individuals may function as a dual-shaming of already stigmatized students while encouraging quality teachers to leave. Third, the application of shame may increase the likelihood of student failure. Further, a flawed theory of motivation has led policymakers to try and foster competition between teachers and schools. Instead, teachers may feel the same anxiety, apathy and shame as their students.

Motivational Theory

Psychologists and researchers have identified some of the roles that motivation can play in stimulating action and achievement. The literature divides motivation into two types: extrinsic motivation and intrinsic motivation (Alispahić, 2013; Harlen & Crick, 2003; Jones et al., 2007; Schunk, Meece, & Pintrich, 2013). Generally, motivation is described as the instigation and maintenance of effort and can have internal (intrinsic) or environmental (extrinsic) sources. For students, an example of intrinsic motivation is the desire to master a skill, such as reading a book for personal enjoyment. An example of extrinsic motivation is the desire to achieve a particular score on a test.

Research has shown intrinsic motivation to be a powerful stimulant of student learning and achievement, while extrinsic motivation has been found to have deleterious associations such as anxiety and stress, reduced creativity, inflexible thinking, and an

increased likelihood to drop out of school (Alispahić, 2013; Jones et al., 2007; Schunk et al., 2013). Studies have found that upwards of 60 percent of surveyed teachers report that students show signs of increased stress related to standardized tests (Jones et al., 2007). Conversely, students with learning as opposed to achievement goals “show more evidence of superior learning strategies, have a higher sense of competence as learners, show greater interest in school work and have more positive attitudes to school (Harlen & Crick, 2003, p. 12).

Moreover, some research has shown that “intrinsic motivation decreases in the presence of other external events that are experienced as controlling” (Jones et al., 2007, p. 80). These can include punishments, goals, and deadlines. For example, Jones et al. (2007) quote one parent whose second grader still enjoys “music, art, and PE” but explains “the kids know that the test is the real thing and they all think something will happen to them if they don’t pass it” (p. 81). Most research has focused on this “something [that] will happen” or the external consequence of testing for students. Yet, research shows that teachers experience similar pressures and can subsequently “become less autonomy-supportive and more controlling” (p. 85).

The relevant question for this research is: what effect might a more macro-level outcome – such as a school grade – have on an individual’s motivation? An answer to this question will in part require knowledge about the extent to which educators draw explicit connections for students between their individual performance and the school’s overall grade. In my own time as a Florida public school teacher, I was strongly encouraged by the school administration to make that connection by informing students of the FCAT’s importance for

them as individuals as well as for the school as a whole. Even without explicit connections, studies show that when teachers feel pressured and become more controlling, students are less likely to experience the intrinsic motivators most likely to ensure high quality learning (Jones et al., 2007).

Harlen and Crick (2003) posit that high-school students are much more likely, than younger students, to be aware of connections like those between their individual scores and school grades. Older students are more likely to perceive grades as unfair, yet place more importance on them, and are more likely than younger students to focus on performance outcomes – or extrinsic motivators – instead of learning processes. Additionally, older students who perform poorly are more likely to grow apathetic and answer tests with random guesses, while feeling resentment, anxiety, and mistrust of the assessment. Assuming high school students in Florida are aware of the importance their individual scores play in their schools' grades, or are affected by changes in teacher behavior, these findings suggest that students at failing schools may be more likely to internalize school failure while either losing motivation or suffering the anxiety and other ill-effects of solely extrinsic motivation.

The Shame Nexus

Creed et al. (2014)'s discussion of shame is grounded in institutional theory. Institutions, they argue, are inhabited by individuals whose shared values make certain behaviors or choices desirable. These values are encoded in the formal structures and symbols of institutions – such as, I argue, grade and discipline systems in schools.

Individuals are not machines, and they therefore experience emotions, including emotions that connect people and lead to social bonds. According to Creed et al. (2014), social emotions “emerge as reactions to our perceptions of our social standing and that of others relative to norms and standards within social structures” (p. 10). They may be “other-directed” such as anger or disgust, or “self-directed” such as guilt or shame. The latter occurs when one feels that it is they who have acted against the mores of the institution. In schools, values and mores are represented by grades and students have long been pitted in competition with one another. Similarly, merit pay and other punitive and reward mechanisms have similarly encouraged competition among teachers (Goldstein, 2014).

In an institution, shame works to shape the behavior of individuals and to reproduce the institution’s values. Creed et al. (2014) argue for the existence of a “shame nexus” composed of four concepts: felt shame, systemic shame, sense of shame, and episodic shaming. Felt shame is “a person’s experience of negative self-evaluations based on anticipated or actual depreciation by others due to a failure to meet standards of behavior” (p. 3). People feel bound to these standards in order to maintain social bonds and their standing in the social group. Felt shame signals to a person how and when to act in such a way as to maintain status.

Sense of shame is defined as “an internal mechanism of intersubjective surveillance and self-regulation” (Creed et al., 2014, pp. 3-4). This internal mechanism works together with systemic shame – a “form of disciplinary power comprising shared understandings of the conditions that give rise to felt shame” (p. 4). Creed et al. explain this interplay:

The high existential stakes of life and meaning derived from important social bonds set the stage for disciplinary power...if persons lose their ties to valued social relations, they risk losing their connections to the symbolic systems that make their lives meaningful (p. 14).

So, shame is in part a fear of losing social bonds and consequently one's place in institutions of power. Institutions use symbolic responses as "episodic shaming" – essentially a way to punish transgressions and induce felt shame.

I believe that these concepts can be appropriately applied to schools. Schools have commonly agreed upon values and norms including academic and disciplinary standards. Under a standardized accountability model, test scores and school grades work as systemic shame, while the application of those scores and grades to individuals or schools, and the publication of those scores and grades, serve as episodic shaming. On an individual level, class placement may also serve this role. While a teacher, my high school consigned the lowest performing students on the FCAT to special "study skills" classes, an embarrassing schedule assignment among their peers. Additionally, some research shows that poorly performing students are suspended or expelled to prevent their dragging down of test scores (Greifner, 2006). These public brandings, like categorizations based on race, class, gender, sexual orientation, and athletic ability, can affect the psychological development of children and their own self-images as they age (Bates & Glick, 2013). As teachers are held more accountable for the performances of their students on standardized tests, they are caught within the same net of episodic shaming.

On a more macro, school-wide level, Goldhaber and Hannaway (2004) find through case studies that "there is a very real and tangible social stigma attached to being judged a

low-performing or failing school, at least for teachers” (p. 605). In fact, studies consistently attribute changes in school performance to the stigma associated with teaching at or attending a failing school (Bowen & Trivitt, 2014; Figlio & Rouse, 2005; Goldhaber & Hannaway, 2004). One teacher tells Goldhaber and Hannaway that:

With that stigma [of an F] it was one of the hardest years I’ve ever taught. All your buddies are at other schools with C or D or A or B. No matter what they say, they look at you and their look asks, “What do you have over there? Why is your school an F?” (p. 604).

Theoretically, the blame implied by that question – why is your school an F – is internalized by teachers and students. Welner and Carter (2013) explain that, in education policy, “we are told that poor children – who are less likely to possess the family, neighborhood, and material resources that we know improve test scores and other measures of achievement – have no excuses for not performing as well as middle-class and affluent children.” If these disparities in opportunity are “no excuse” then the blame must logically fall on the teachers or students as individuals.

Furthermore, some researchers have speculated that teachers feel reluctant to teach at schools that may be associated with certain students – be they poor academic performers or simply poor (e.g. residents of public housing) (Reingold, Ryzin, & Ronda, 2001; Schwartz et al., 2010). Assuming this “shame-nexus” – a combination of individual and institutional shame – turns around failing schools, is it desirable? Creed et al. (2014) argue that shaming can “trigger renewed efforts to preserve those [social] bonds” (p. 14) and research has found that stigma is associated with improved school performance (Bowen & Trivitt, 2014).

Yet, all schools do not turn around. Creed et al. (2014) explain that there is a possibility that for some, shaming prompts rejection instead of acceptance. This acceptance or rejection depends on factors such as “the value of the social bond, the effort and social, emotional, or material resources required...and expectations of success” (p. 24). Complying with expectations is difficult, the authors argue, if compliance “demand[s] too much or may not even be within the transgressor’s power” (p. 25) or if personal characteristics are stigmatized such as “styles of social behaviors associated with poverty or lower social-class origins” (p. 26). While some may respond by attempting to transform themselves or the institution, others will leave or actively resist its norms. In the case of schools, some students may act out, drop out, or grow apathetic; while some teachers may change schools or leave the profession altogether.

Some researchers have noted the rebellion of teachers against policies that they do not understand or are not based on their input (Paige, 2013). Harris and Herrington (1994) describe the “one common theme of policy implementation literature” as: “policies rarely affect practice as intended” (p. 72). They continue:

Educators try to make sense of reforms and will often follow through when they have sufficient understanding, support, and capacity. But when these conditions do not hold, educators will tend toward minimal compliance, and school leaders will buffer their teachers from unwanted intrusions (p. 72).

Furthermore, the imposition of accountability with which they disagree may be “undermining the morale of teachers” (p. 73). When I first left teaching to pursue my graduate studies, many teachers that I knew confirmed these theories by saying that I “was getting out at the

right time” – coinciding with the ramping up of accountability policies. Jiang et al. (2015) find that these feelings do not stop with individual teachers, but affect their perception of their principals, schools, and colleagues as well as increasing their stress and dissatisfaction with the profession.

We also have evidence to suggest that certain students and schools face steeper challenges. Grodsky et al. (2008) find that students of different races, ethnicities, and income-levels have significantly different “opportunities to learn.” – “resources available to students, most often in the classroom setting, that facilitate their acquisition of knowledge or skills...[including] teacher pedagogy, ability and effort, instructional pacing, and curriculum” (p. 388). Likewise, Jimenez-Castellanos (2010) argues that educational resources positively correlated to academic achievement – such as higher teacher salaries, newer schools, and more space per pupil – vary significantly, with Latino, low-income, and English Language Learner (ELL) students notably disadvantaged relative to their white peers.

Some schools are further disadvantaged by student zoning. F schools studied by Goldhaber and Hannaway (2004) had significantly more students who were eligible for free or reduced lunch (due to low income) than did A schools, while Porfeli et al. (2009) find that schools in the southeast with low community capital housed the majority of students with disabilities and the least numbers of gifted students. Based on these kinds of imbalances, Borman and Dorn (2007) conclude that although Governor Bush “deserves credit for arguing that schools should hold all students to high standards...his policies have pushed for equality in a narrow way” (p. 3). They continue: “we are skeptical about seeing significant and sweeping improvements in how well schools work for all students” (p. 3).

Stereotype Threat

Shaming may be particularly dangerous if it encourages existing prejudices or insecurities, particularly if those prejudices and insecurities interrupt academic performance. On the one hand, stereotypes may affect students of color insofar as they want to avoid “acting white” by performing too well in school (Ferguson, 2008). Herbert Kohl (2016), in his now famous essay “I Won’t Learn From You” categorizes this as active “not-learning” – the conscious decision and labored process of refusing to learn something one otherwise could. In a study of U.S. high schools, Signithia Fordham and John Ogbu observe that black students aware of systemic oppression develop oppositional identities to protect themselves from racism (Bates & Glick, 2013). On the other hand, fear of adhering to these stereotypes about underperformance may produce increased anxiety and lower scores among students of color or other stereotyped groups who do not develop oppositional identities, for whom school achievement remains important.

Claude Steele (1995) calls this later phenomenon “stereotype threat.” Stereotype threat occurs whenever a person fears confirming a negative stereotype about their identity. According to Steele (2010):

Whenever we’re in a situation where a bad stereotype about one of our own identities could be applied to us – such as those about being old, poor, rich, or female – we know it. We know what “people could think.” We know that anything we do that fits the stereotype could be taken as confirming it. And we know that, for that reason, we could be judged and treated accordingly (p. 5).

Steele proposes that this fear of confirming stereotypes and being judged accordingly is strong enough, and produces anxiety to the degree at which it can interrupt performance – particularly academic performance.

Steele and others have found evidence in multiple studies since 1995 that seems to confirm this theory (Carr & Steele, 2009; Lawrence, Marks, & Jackson, 2010; Osborne, 2001; Rossman & Rallis, 2003; Shih & Pittinsky, 1999; Steele, 2010; Steele & Aronson, 1995; Van Loo, Boucher, Rydell, & Rydell, 2013; Wout, Danso, Jackson, & Spencer, 2008). Researchers in these cases have taken various stereotyped groups – including women (stereotyped as bad at math), Asians (stereotyped as good at math), African Americans (stereotyped as bad at academics), and white males (stereotyped as good at math and academics generally) and given them standardized test samples. It has consistently been the case that, when the stereotyped group and non-stereotyped group are each given a test described as a measure of their abilities (an intelligence or math test), the stereotyped group underperforms dramatically. Likewise, when a usually non-stereotyped group (white males) is given a test and told that Asian students typically outperform white students on that particular test – the scores of white males go down (Steele, 2010). However, when a third group composed of stereotyped students is told that the test will not be a measure of their abilities, but instead a measure of tests in general, the performance gap disappears.

These results suggest that certain phrases or contexts serve as symbols that activate stereotype threat. Something about a test that measures individual ability provokes a different response than does a test that does not measure the test-taker's abilities, but the quality of something external to them – such as test-takers in general, or the test itself. Psychologist

Sian Beilock (2010) theorizes that this is due to a distracted left pre-frontal cortex, which has difficulty processing both identity anxiety and problem solving at once. This is not only true of typically under-performing students. In fact, research has found that the more highly a student values education and their academic performance, the more their score drops when stereotype threat is activated (Steele, 2010). So, a framing that divorces the test from high stakes associated with individual evaluation may decrease gaps in performance. However, our evaluation systems increasingly do the opposite.

Standardized tests in the United States are increasingly presented as measures of individual ability and as associated with high-stakes (W. W. Au, 2009). Furthermore, states in the south with a higher percentage of black and Hispanic students who statistically under-perform are more likely to associate higher stakes with their tests, particularly in high-poverty areas. Interestingly, following the adoption of the FCAT in Florida and high stakes tests in a host of other states, graduation rates went down (Amrein & Berliner, 2002).

I suspect that school grades increase the risk of stereotype threat in two ways: first, if students know that their individual scores influence their school grade, which affects their teachers and classmates, then their performances are even more high-stakes than they would be if scores only affected them as individuals. Increasing high-stakes have been linked to higher levels of anxiety and increased score disparities (Osborne, 2001), explaining why students who care more about school may be more highly affected (Steele, 2010). Second, the branding of a student's school with a D or an F grade introduces a new stereotype – that of the failing school. Students at repeatedly failing schools are expected to fail. Just as some white students in Steele's experiments were told that they are typically out-performed by

Asians (which prompted under-performance), so students at failing schools are told – in a highly public way – that they are typically out-performed by the students at A and B schools. Furthermore, teachers at failing schools may expect their students to fail or may themselves feel stereotyped as poor educators. Research has shown that something as simple as having a more feminine or masculine sounding name can influence a student's academic outcome; how much more might a school-wide label affect them (Rossman & Rallis, 2003)?

Considered together with motivation theory and Creed et al. (2014)'s shame nexus, stereotype threat provides a compelling model for how school grades may produce decreases in certain school outcomes. Perhaps, school grades serve as a mechanism for shaming students, some of whom respond by rejecting school norms. Perhaps, the emphasis on standardized outcomes – individual and school grades – externalizes motivation such as to decrease engagement; and perhaps the knowledge that one attends a failing school and the shame associated with that label activates stereotype threat, interrupting academic performance. This may happen even for students who personally pass exams, but fall under the umbrella of their school's grade. Finally, affluent families might pull their students out of schools perceived as failures, just as white families often switch out of predominantly black or Hispanic schools (Booher-Jennings, 2005). If these processes really occur, then even if aggregate test scores increase, school grades may prompt negative outcomes from some students and their teachers.

CHAPTER 3

POSITIONALITY, RESEARCH DESIGN AND METHODOLOGY

Having established the need for an investigation into the non-test outcomes of school grades in Florida, this chapter explores my own positionality, as well as outlines the research design and methodology for a mixed-methods – qualitative and quantitative – study of school grades, primarily from the perspective of teachers.

Positionality Statement

In 1999, the same year the state began grading schools, my family moved to central Florida. I was 13 and about to enter high school. As a student, I never considered school grades. I had the privilege not to. My parents had always valued my education and had the financial means to either pay private school tuition or ensure we lived near a public school with a positive reputation. I took it for granted that, when I started my ninth grade year in 2000, my new high school had an A.

Ten years later, in 2009, I was hired for my first full-time job, as a high school teacher. Once again, I took it for granted that the school had an A grade. Lower-scoring schools seemed to exist in a different world than the one I inhabited. So, I felt shocked when – in the summer before I began teaching – the state released school grades and my new workplace earned a C. The grade contradicted the school’s reputation. People in the community told me that the grade must be wrong, that the state had changed the grading formula and the new metrics were inaccurate. A C seemed suspect.

The first faculty meeting felt depressed. A teacher I met for this study calls a similar meeting at her school “funeral” – which I think captures the tone. The principal acknowledged that the grading formula had changed and that the school had been rezoned to capture slightly more lower-income students and students of color. He did not offer these as excuses, but as explanatory variables to be addressed. He spent the bulk of the meeting on a presentation aimed at improving the school’s performance. Teachers shouted; some cried. They felt unfairly judged. They called the expectations unrealistic. Their dignity and reputation as A-school educators had been stolen, as had the roughly 800 dollar reward bonus they normally received.

For the first time, I thought about school grades: what they meant, how they were scored, what it felt like to not have an A. As a former straight-A student, it felt embarrassing. I felt the need to defend my school to friends and neighbors who wondered why it had declined. I thought about the low-income students and students of color who some blamed for the low score. I had never taken classes with students like them. Why might they

struggle? I thought – if we feel like this over a C grade, how do teachers at persistently failing schools feel?

That year I taught standard-level language arts classes. For the first time, I shared a classroom – albeit as a teacher – with students who lived in their cars, who struggled to afford lunch, who were addicted to drugs, who got arrested, or who went home to primarily black and brown neighborhoods. I also taught classes with mostly white, affluent and honor roll students, but I was already familiar with them. I was not familiar with the links between what some researchers call “opportunities to learn” – differences in factors of nutrition, housing, and family – and academic success. I was unfamiliar with these things because students are segregated – both between and within schools – by race and class. The students who took honors and advanced placement courses, as I had, were fairly homogeneous, while students in the standard classes I now taught were visibly and invisibly diverse.

I taught at that school for three years, during which time it returned to earning an A. The pride of that grade felt nice, as did the annual bonus money. Despite those nice feelings, I began to see the whole system as absurd and unequal. So did other teachers. To paraphrase a teacher who I quote later in this dissertation: school grades – like class levels – seemed to say more about neighborhood than they did about school quality. If grades said the wrong things, did they also cause unintended consequences? I learned that things important to teachers – attendance, discipline, drop-out rates, job security – went unmeasured by grades. Did they relate in unmeasured ways? I also learned that teachers felt adversarial towards state policymakers who they accused of ignoring student factors that influenced academics the

most – neighborhood, family, income, race – none of which factored into standardized test scores.

I include this personal history for the sake of transparency about my own interests and experiences with school grades. There is always the risk of researcher bias, particularly in qualitative work (Maxwell, 1996; Schutt, 2012; Weiss, 1994). As a Florida teacher, I have my own professional feelings about school grades. Maxwell (1996) stresses that the proper response to bias is “integrity” rather than “indifference” (p. 91). I, like Maxwell, believe that no researcher can entirely divorce his research from his person. Thus, I have tried my best to be open about my own feelings and experiences regarding school grades and education policy. In this sense, I align with a more constructionist paradigm, defined by Rubin and Rubin (2012) as those who “emphasize that all meaning is sifted through people’s prior experiences and biases” (p. 16). My academic speculations on school grades are encapsulated in my literature review. Personally, I have only limited exposure insofar as I have worked at three Florida schools – all relatively high-performing. I can attest to an incredible amount of pressure to achieve an A grade being placed upon teachers at each school by the administration and upon students by teachers.

My experience at a C-school, albeit embarrassing, does not mirror the “stigma” teachers associate with F grades. Moreover, my school earned an A grade the following year and shortly thereafter I left the county. So anecdotally I can speak to feelings of shame among the staff; however, I did not experience a lasting impact on student performance or school culture. Albeit limited, this experience impacts my thinking about drops in school grades, especially drops to D or F grades that are theoretically even more impactful than my

school's C grade. At all of my career schools, I have sat through many lunch and staff meeting conversations about how unreliable, inaccurate or unfair school grades are. Admittedly, these conversations helped inspire my research questions in this dissertation.

While my perspective is somewhat critical, insofar as I hope my research to help address a social problem (Rubin & Rubin, 2012), I have attempted to maintain some distance from my participants. My goal is, with some influence of my own teaching experiences, to report the feelings and ideas of teachers in an effort to communicate information to policymakers. Thus, my own advocacy is tempered by what I see as politically feasible and, when appropriate, I have checked the reports of teachers against available data (e.g. on the size of a school or the effect of a prior policy). In so doing I hope to have struck a balance between advancing the best interests of teachers, reporting “truth” insofar as it can be measured and suggesting realistic policy reforms.

Research Design

I initially envisioned this dissertation research as a three-phase project, starting with content analysis of major Florida newspapers, followed by qualitative and finally quantitative research and analysis. I chose a mixed-methods design in order to: first, explore the issue of school grades from multiple perspectives; second, understand the distinct roles played by different actors; and third, verify findings. Schutt (2012) calls this “triangulation” which “suggests that a researcher can get a clearer picture of the social reality being studied by viewing it from several different perspectives” because “each [method] will have some

liabilities in a specific research application and all can benefit from a combination of one or more other methods” (p. 17, see also: Rossman and Rallis, 2003). As an example of such a benefit, I suspected that teachers hold intense emotions regarding their evaluations and so the use of media and quantitative measures might verify qualitative findings. For example, I worried that teachers might exaggerate or personalize the extent to which high quality teachers left their schools each year; however, my dataset confirmed their testimony, showing that teachers do leave failing schools at higher rates than at non-failing schools.

As some researchers (e.g. Paige, 2013) argue, teachers play a fundamental role in the implementation of education policy. Moreover, they are the actors who primarily interact with students and parents – arguably the patrons of the education system – on a daily basis. In this way, their communication, their attitudes, their non-quantifiable fears, hopes and ideas, all have serious policy implications that might only be captured through long-form interviews that allow them to feel comfortable and speak openly. My research might be called “exploratory” insofar as interviews introduced me to problems and ideas that I, despite my teaching background, had not anticipated (Schutt, 2012). For example, some teachers emphasized the lack of a culture of success at failing schools, something I had never heard described and had not experienced. Finally, I chose to pursue qualitative research because one of my goals is to further public understanding of teachers and students and to shape education policy debates. As I hope the text of my dissertation demonstrates, the qualitative words of teachers have the potential to be uniquely powerful in doing so.

In addition to verifying the statements of teachers, the quantitative research methods I pursued have hopefully amplified the scope of my findings and increased their

generalizability. Due to time and resource constraints, I could only meet with fifteen participants at a total of four schools; however, my dataset spans every consistently-graded school in the state. As Schutt (2012) explains, quantitative data is able to encompass “few data on many cases” as opposed to “many data on a few cases” in qualitative research; further, qualitative data is “sensitive to context” rather than quantitative data which seeks “universal generalizations” (p. 325). The “many cases” and “universal generalizations” of quantitative research are necessary in the world of education, which, as I argue here, is consistently driven by standardized goals and findings.

Due to some preliminary findings and research obstacles, the project’s design has not developed as I envisioned. In this section, I aim to explain the research design, reforms, and implementation. As I planned the project in three phases, I organize this section accordingly.

Phase I – Content Analysis

Bowen and Trivitt (2013) compare the performances of failing schools before and after the Florida Supreme Court ruled vouchers unconstitutional in 2006. They find no evidence that vouchers uniquely drove school improvement and conclude that avoiding the “stigma” of a failing grade may motivate schools. Based partly on these findings and my anecdotal experiences, I initially suspected that the media may serve as a stigmatizing force; that by calling attention to failing schools, newspapers may invoke additional shame. The primary goal of this phase was to analyze how the media portrays school failure. My early media analysis surprised me.

I selected three newspapers to review: the *St. Petersburg/Tampa Bay Times*, the *Orlando Sentinel*, and the *Sun Sentinel*. These papers not only hold the top three largest distributions in the state, but represent different geographic regions: the west coast, central Florida, and south Florida. I began with the *St. Petersburg/Tampa Bay Times*. Its archives are available via ProQuest Central. Using ProQuest, I conducted a keyword, in-text, and in-title search of the phrase “school+grade” for the period of 1999, when schools were graded for the first time, to 2014, just before the one-year waiving of school grade policy consequences. I eliminated the resulting articles that incidentally included the words “school” and “grade” but did not actually relate directly to “school grades.” The number of remaining articles by year is displayed in Table 1.

Table 1

St. Petersburg/Tampa Bay Times “School+Grade” Search Results, ProQuest Central Archives By Year

Year	Number of Search Results
1999	27
2000	21
2001	27
2002	20
2003	33
2004	37
2005	8
2006	15
2007	25
2008	38
2009	30
2010	16
2011	19
2012	23
2013	27
2014	26

In addition to the search counts, I catalogued the article titles, authors, quoted individuals, and web addresses in Excel. Additionally, I recorded the overall position of the article as against, in support of, or neutral on the issue of school grades. I listed pertinent quotations, the speakers of said quotes, and their positions. Finally, I made special reference to mentions of particular schools, especially if mentions implied shame or stigma (e.g. flunked, bombed, ashamed).

This process did not prove relevant in the ways I had expected. I had expected journalists to accept grades as an appropriate standard and consequently, to implicitly condemn failing schools. For the most part, they did not do so. Instead, the only cases in which journalists at *the Times* reported grades without acknowledging their controversy was to report on the overall achievement of the district. For example, articles discussed the number of A-grades in the district rising or falling, without profiling particular schools with any depth. The majority of articles fell into one of three other categories: (1) politicians campaigning against or in support of school grades; (2) news of state policy changes; and (3) overt discussions of grades as controversial, citing teachers or parents.

Of these three categories, the last two played significant roles in my research. Updates on state policy changes provided a significant complement to the research literature and Florida's own publications. These articles helped me to outline policy changes between 1999 and 2015, including the views of policy-makers, parents, and teachers on each change. Articles on the controversy of grades validated my own experiences and conjectures that

many teachers, parents, and local politicians oppose grades as a measurement. Many see them as unreliable, stigmatizing, and political.

In 2004, for example, Democratic State Senator, Frederica Wilson, attacks grades as arbitrary and political. She projects fewer school failures because “she doesn’t think the students are any smarter, [but] the state made the test easier because this is an election year and the governor’s brother is running for president.” She adds that, failures would lead to “protest after protest every day leading up to the election, and they don’t want that” (Waite, 204). In a letter to the editor in 1999, Kip Mitchell, “an educator and parent” critiques the premise of the state’s reward system. He writes that “It doesn’t take a brain surgeon to figure out that when a ship is sinking, you go to it and do whatever it takes to save it and its’ occupants. However, the Legislature and governor believe that instead of saving it, you abandon it and instead praise and reward ships that are watertight and sailing into the sunset” (Mitchell, 1999).

Teachers also express shame and confusion. Marilyn Oberle, another teacher, writes in a 1999 Letter to the Editor that:

I felt relief as I focused on our B grade; an A grade would have been better, but a B shows we try. Or does it? As some schools scored better and some worse, I felt confusion. Just what are we telling the public? I know teachers, students and parents at many schools. We all try! So does this mean I should feel better about what I do at Mildred Helms than the teachers at the C schools and worse than the teachers at the A schools? Maybe I should transfer to an A-grade school and take my friends (teachers, students and staff) with me. I've been at Mildred Helms for 21 years and I can tell you, no matter what the grade, we are a top-notch elementary school. We work hard every day to do and be the best (Oberle, 1999).

Another educator, Faye Adler, asks, “How do children feel when they discover their school got a rating of C, D, or F? They feel inferior and definitely not worthy” (Adler, 1999). In the same year, Rob McMahon, the president of Pinellas Classroom Teachers, calls school grades a “deliberate subjection of our children, our teachers, our support employees and our schools to humiliate and ridicule with an arbitrary system of grading” (McMahon, 1999). Similarly, Sandi Jacobs, a fifth-grade teacher, compares grades to “the Scarlet Letter” and says they are “clear enough” but “kind of demoralizing” (Hegarty, 1999).

These statements foreshadow my interview data and confirm my own experiences, but do not represent the third-party influence that I expected in this phase of my research. While I believe there is importance to a study of the media’s role in amplifying the voices of educators and parents, I scaled back the content analysis piece of my dissertation. I felt my research questions could be best answered by speaking directly with educators and parents, rather than reading their words through a press filter. I went on to complete a cursory review of articles in *the Orlando Sentinel* and *Sun Sentinel*, but do not include a content analysis chapter in my dissertation.

It is possible that the role I expected the media to play – that of implicating schools, or at least discussing their individual performances – does happen at the local level, in community or town newspapers as opposed to big-city papers with large distributions. However, my review of large papers still proved rewarding. My primary purpose was to build context for phases II and III, not to quantify school grade discussions at a comprehensive level. Reading the words of teachers and parents, the positions of politicians, and learning of policy changes between 1999 and 2015 helped determine the questions I

included in later interviews and the variables I pursued in my dataset, vital components of pre-structuring a qualitative or mixed-methods study (Maxwell, 1996).

Phase II – Qualitative

The primary goal of the qualitative phase of my research is to explore impacts of school grades such as feelings of shame and anxiety, changes in school culture, and accounts of teacher-parent and teacher-student interactions. Research questions include: How do teachers react to the reception of school grades? How do teachers view the school grading policy? How do school grades impact school culture and policies?

I also hoped to develop a relatively holistic view of school environments and cultures and the interplay between those contexts and the receptions of school grades. Because I wanted participants to speak openly and frankly, and possibly introduce concepts I had not anticipated, I chose an in-depth, semi-structured interview approach. I asked each participant the same questions – as informed by my interview guide – but also followed tangents and conversations, allowing for the introduction of new ideas (Rubin & Rubin, 2012; Weiss, 1994). This had the beneficial effect of making participants feel more relaxed and speak more openly. For example, one parent spoke in very short, reluctant responses until we began to tangentially discuss her children, at which point she admittedly “opened up.”

I sought to meet with teachers at two Florida schools. In order to better understand how failing grades cause certain reactions, I hoped to select one D or F-graded school and one A or B-graded school, to make comparisons. In qualitative research it is not possible to

isolate variables and prove causation, but as Maxwell (1996) explains, researchers “tend to ask how x plays a role in causing y, what the process is that connects x and y” (p. 20). I thought these answers could be better answered if I talked with teachers at both ends of the grade spectrum. I planned to interview ten teachers or administrators at each. I considered only middle and high schools since teenaged students are most likely to understand and internalize the stakes of testing (Jones, Jones, & Hargrove, 2007).

While Florida is a large and demographically diverse state, and evidence shows that schools in different regions have access to different resources (Gonzalez, 2012), I sought two schools located in the same county. I made this choice partly due to feasibility and travel concerns, but also to focus the scope of the study on the impacts of school grades as opposed to geographic differences between sites. Theoretically, while these differences have been shown to influence student performance (Grodsky, Warren, & Felts, 2008), I suspected that the reception of a failing grade would have negative consequences for all demographic groups, across regions (Steele, 2010).

Finding two such schools proved more difficult than I had anticipated. I initially chose three high schools in different areas of the state that had received failing grades in the 2013-2014 school year, in the hopes of identifying a county that I could work in. I contacted department chairs in English, Mathematics, Social Studies, and Science departments via e-mail. I received no responses. Follow-up e-mails were also unsuccessful. I then chose three additional schools and contacted their department chairs, who also did not respond. While I cannot know for certain what caused teachers to not respond, I suspect that they may have had confidentiality concerns. Speaking critically about school grades and their school

administrations could be seen as politically sensitive and professionally risky. Additionally, I reached out during a school year, when teachers are often overwhelmed with work and e-mails.

I turned to my content analysis work. I began to contact principals and other educators who I had seen quoted in articles. I reasoned that educators who had granted interviews to journalists had demonstrated a willingness to speak openly. I finally received a response from Alex, who had just been transferred from his position as principal at a failing, urban middle school. Alex graciously agreed to speak with me and from there I followed a snowball method for identifying additional participants. In other words, I ended my interview with Alex – and each subsequent interview – by asking if he knew any teachers or administrators who might be willing to speak with me. Table 2 details each participant in my study, their position, school information, and who referred them to me. I ultimately interviewed five teachers at Alex’s former school, as well as his replacement as principal. I contacted the other individuals with references as “newspaper” or “website” based on finding them quoted on a school site or newspaper article, with all other participants contacted via snowballing.

Table 2

Participant Data

Name (Pseudonym)	Position	School Site	School Level	School Grade	Reference
Alex	Principal	School 1 (Formerly)	Middle	F	Newspaper
Joy	Teacher - Science	School 1	Middle	F	Alex
Sandra	Teacher - Science	School 1	Middle	F	Joy
Rhonda	Teacher - Science	School 1	Middle	F	Joy
Jennifer	Principal	School 1	Middle	F	Rhonda
Lakisha	Teacher - Reading	School 1	Middle	F	Jennifer
Martha	Teacher - Social Studies	School 1	Middle	F	Jennifer
Theresa	Teacher - English	School 2	High	B	Website
Jane	Teacher - English	School 2	High	B	Theresa
Sarah	Teacher - English	School 3	High	C	Jane
David	Teacher – English/Social Studies	School 4	High	A	Website
John	Teacher - Math	School 4	High	A	David
Jason	Nonprofit	N/A	N/A	N/A	Newspaper
Angela	Parent	N/A	N/A	N/A	Jason
Carla	Parent	N/A	N/A	N/A	Jason

I do not include race and gender information in this table, although both could be relevant. I did not ask participants to identify their race and gender in interviews and while some self-identified, others did not.

This method of contacting participants carries both benefits and costs. By having personal references, I increased my response rate and potentially built trust more easily with my participants. I also reached participants who I may have never found otherwise and who had proved difficult to connect with via unheralded e-mails or phone calls. On the other hand, snowball sampling may bias my sampled population. By relying on teachers who know and recommend one another, or who spoke to reporters, I may have limited the range of views held by my participants (Schutt, 2012). This concern does not worry me greatly, as several teachers recommended colleagues who they admitted to strongly disagreeing with, not knowing that well, or working from different departments. For better or worse, this sampling method also led me outside of my planned two school sites in the same county. I ultimately spoke with teachers and principals from four different schools, in three different counties, as well as a non-profit representative and two parents – a total of 15 participants.

I developed an interview guide for teacher interviews both inductively – based on my secondary research and content analysis phase – and deductively, along a naturalist-constructionist paradigm. That is, interview questions focused on how participants interpret and perceive their experiences and interactions, as well as policies at their schools (Rubin & Rubin, 2012). Interviews were semi-structured such that each participant was asked prepared questions as well as follow-up and conversational questions (Rubin & Rubin, 2012; Weiss, 1994). I kept interviews as conversational as possible while still exploring the issues in a comprehensive way. I believe this increased the levels of trust and openness, as did my own experiences as a Florida teacher. Prepared questions for teachers and administrators explored issues such as: school policies before and after the reception of failing grades; student and

teacher anxiety; curriculum and testing practices; teacher relationships with students, parents, and administrators; professional decisions; school morale; and policy recommendations. I did not expect the opportunity to speak with parents and my interviews with them took on an even more conversational, free-flowing form. Interviews lasted approximately one hour.

Trustworthiness and Rigor

All participants signed a letter of consent previously approved by the University of Massachusetts, Boston Institutional Review Board. I kept these letters in a confidential and secure location. Each interview was audio-recorded and transcribed by the service www.rev.com. In order to allay any fears of confidentiality, I guaranteed participants anonymity. I assigned each a pseudonym that I used in all written products, including on transcripts during coding. I tried to choose pseudonyms that protected the participants' identities while still reflecting the character of their real names. This was an admittedly subjective practice that I followed in order to try and make them relatable and true for readers. Surprisingly, several teachers told me that I could quote them by name, although I ultimately did not do so. Rossman and Rallis (2003) explain that the use of written products may extend beyond what the individual or researcher foresee or control; they argue that anonymity should be maintained "unless a compelling reason is put forward not to" (p. 74). I see no such compelling reason.

I stored their audio recordings and transcripts on a hard-drive accessible only by me. In order to further protect confidentiality, I censored all mentions of particular school names

or locations. While many Florida school districts have their own Institutional Review Boards, I did not seek their approval. Some require names of participants and school sites, as well as written findings, the provision of which would violate my assurances of confidentiality.

Regarding participants, I did not perform post-interview member checks, or “systematically soliciting feedback about one’s data and conclusions from the people you are studying” (Maxwell, 1996, p. 94). I had several reservations about member checks as a validity tool. First, contacting teachers had proven difficult and may have been more so given their propensity to change schools or neglect e-mail during summers. Second, as Maxwell explains, feedback cannot necessarily be taken as accurate. I wanted to include the authentic thoughts of teachers, not those edited after consideration of what peers or supervisors might think. I did ask several early participants if they would be interested in participating in member checks; all declined. I did begin interviews by stressing that interviewees could withdraw participation at any time or specify parts of the interview as confidential. Nobody requested to withdraw from or censor our discussions.

In all of these steps and conditions, I have attempted to conduct a trustworthy study. Rossman and Ralis (2003) outline two conditions for a trustworthy study: that it conforms to “standards for acceptable and competent practice” and “standards for ethical conduct with sensitivity to the topic and setting” (p. 63). In short, they write, “an unethical study is not a trustworthy study” (p. 63). It is my understanding that no harm befell any participant as a result of their participation. Their identities have been protected in every stage, from their ability to choose the location of our interviews, to their anonymity in written products, to

their assignment of pseudonyms. Finally, the purpose of the study is to contribute to understand and improvement of teacher circumstances (Rossman & Rallis, 2003).

In addition to recording and transcribing interviews, I took extensive field notes before and after interviews in order to capture the process, details, and my impressions of participants. I allowed each participant to choose the location of our interview. Chosen locations ranged from the participants' homes, to coffee shops, to their classrooms. My field notes also describe the settings of our interviews. Notes and reflections, especially my descriptions of classrooms, helped to interpret transcripts and in later analysis (Maxwell, 1996).

I believe that the use of memos after interviews, prior to coding and – in some cases – after coding, has contributed to the rigor of my study. In addition to comparing memos and interviews to one another, in some cases I engaged in follow-up research. For example, one participant suggested I read a news article that had informed her knowledge. I did so and incorporated elements of it into my dissertation. In cases where participants cited numbers or statistics, I attempted to verify them. In this way, I hope to present a clear, accurate, and nuanced vision of reality from multiple perspectives.

I ended the interview phase of my study after fifteen participants mainly due to time and resource constraints. While my latter interviewees suggested other teachers to interview, I began to have trouble getting responses from their suggestions. This may be because a school year had ended. Ironically, teachers are difficult to contact during the school year, but during the summer vacation they often leave town on trips or take care of their own children.

I experienced some degree of saturation in the interviews that I conducted, with teachers repeating similar themes; however, I still hoped to interview more. I think my study could be strengthened by reaching a greater diversity of teachers, although my participants do vary by race, school, and years of experience. I only speak with two parents and I think future research should focus on parent views of grades with a more extensive number of parent participants.

My first step in qualitative analysis was an exploration of transcripts by reading and writing memos that summarized the content of interviews as well as my personal and academic reactions (Maxwell, 1996). For each interview, I wrote – as Maxwell (1996) advises – a total of two memos: one immediately after conducting the interview and another after reading – or in some cases listening to – the transcript. In some cases, these memos provided both verification and clarification to the content of an interview. For example, at one school, my memo included notes on the condition of the school: vandalized bathroom walls, an out-of-order sink, damaged textbooks and stained floors. In a subsequent interview at the same school, a teacher discussed the poor condition of the school and its effects on students. After reading that transcript, I was able to connect my observations and the teacher's to an emerging theme in several interviews about the need for material support at some schools.

Second, I coded interviews using the computer software MAXqda. I coded each transcript using a combination of inductive and deductive codes. I developed deductive codes in advance based on the research questions and my literature review, while my interview summaries and content analysis helped me develop inductive codes (Maxwell, 1996; Weiss,

1994). Maxwell (1996) warns against “context stripping” or losing the original context of coded text. In order to avoid doing so, I used memos extensively during the coding process. While I did not take a case-study approach per-se, I did – for example – keep track of which teachers worked at which schools and connected common themes across their interviews. The aforementioned struggle for resources, for example, arose as a theme at two of the four schools from which I met teachers.

Third, these codes were compared and aggregated to connect and form more general themes, a process that Weiss (1994) calls “inclusive integration.” In some cases, this led to surprising connections. For example, after collecting portions of each transcript related to the aforementioned struggle for resources, I asked what other things struggling schools lacked access to. Not all of these resources were tangible. At one failing school, several teachers discussed students lacking access to a “culture of success” that would teach them what it means to achieve and instill in them a desire to do so. This required me to re-read interviews with a more expansive definition of the term “resource” in order to capture what each participant meant by it (Schutt, 2012).

During this phase of analysis, the general structure of my qualitative chapters was formed. Rather than construct a linear narrative about school failure, I organized concepts around a logical policy evaluation process, similar to that proposed by Bardach (2005). I attempted to: diagnose aspects of the problem, assemble evidence, construct alternatives and criteria, and project outcomes (see also: Creswell & Miller, 2002; also see: King, Keohane, & Verba, 1994; Rossman & Rallis, 2003).

Ultimately, codes clustered in seven groups related to: community, internal to schools, policy, school choice and charters, school grades, student performance, and teacher emotions and job satisfaction. (A complete list of codes, organized by topic, is featured in Appendix C). Rather than organize my qualitative chapters using these same topics, I chose to weave all seven topics into a narrative. As my research questions and many of my interviews stemmed from issues teachers took with the school grading policy, I started there. I began my writing by focusing on what teachers see wrong with current policies then moved to the solutions they oppose and favor, the demographic factors standing in the way, and finally their feelings about these topics.

In my writing, I have done my best to preserve the context and intent of my participants' quotations. In some cases, this has led to atypically long quotes. I resisted breaking up some of these quotes in order to illustrate the interconnectedness of issues stressed by teachers. With the exception of changing names and removing identifying details about schools and locations, I have attempted to capture the voices of my participants. I have removed excessive verbal tics (e.g. like, umm, I mean) but only in cases where they obscured the message. On a stylistic note, I capitalize adjectives derived from proper nouns (e.g. Asian, Hispanic), but leave other sometimes-capitalized ethnic identifiers (e.g. white, black) lowercase. In referring to students of Latin and Central-American descent, I use either the descriptor Hispanic or Latino, rather than the increasingly accepted Latinx. I do so to mirror the language of either the state of Florida or my interviewees, depending on the context. I realize that writers and readers have different preferences in these regards but have tried to

maintain consistency. Finally, while I sometimes place my research actions in the past tense, I quote the words and actions of my interview participants in the present tense.

Limitations and Validity Concerns

I initially considered interviewing students as well. For a host of reasons, both practical and ethical, I chose not to pursue this option. Without student involvement, my interviews are limited to adults, which risks what Maxwell (1996) calls a “description threat.” In other words, the sample of participants may be too limited or homogeneous to illustrate a comprehensive view of school actors and environments. I have tried to address this by asking extensively about teacher and parent interactions with their students. A study focused on adults is not without strengths. Teachers receive the most direct pressure from administrators as well as the state to raise test scores and school grades; they also receive the most direct pressure from anxious students. As the key “middle-person” between administrators, students, and parents – teachers offer valuable insight.

The second major issue of concern is my limited sample size of fifteen participants, including 12 teachers across four schools. Only one school is represented by a significant plurality. Due to the nature of qualitative research, I am unable to achieve generalizability. However, Maxwell (1996) suggests that “the generalizability of qualitative studies usually is based, not on explicit sampling of some defined population to which the results can be extended, but on the development of a theory that can be extended to other cases” (p. 97). It is this kind of theory development that I believe can inform our views of schools and

accountability policies. Even with a larger variety of sites, it can be difficult to accurately describe the larger population (Maxwell, 1996). It is partly for this reason that I have attempted to speak with a variant sample of teachers, across different departments.

A potentially serious problem for authenticity is what Maxwell (1996) calls “reactivity,” or “the influence of the researcher on the setting or individual studied” (p. 91). This may be especially problematic if any participants feel a political obligation to protect school policy. During interviews, I only sensed this once, in an interview with a principal who at times sounded scripted. Additionally, Weiss (1994) warns of the possibility of participants not being forthright or honest, or attempting to improve the researcher’s impression of them. I hope that confidentiality helped to allay any fears of honesty. I believe that my experiences as a teacher helped me to achieve an open and frank dialogue with participants, many of whom seemed to open up once I commiserated with them about teaching experiences. Also, by allowing teachers to choose the locations of our interviews, I hope to have maximized their comfort.

I have worked, throughout all of these stages, to analyze data fairly, despite my views formed as a teacher. I attempted to consistently ask my participants how they might respond to contrary views, in order to tease out multiple perspectives. I have tried to stay abreast of relevant research across multiple disciplines. I also intentionally employed multiple research methods – qualitative, quantitative, and content analysis – in order to explore whether teachers perspectives echo in the press or in the statistical data. I hope that, despite a small sample size, this dissertation gives some voice to the diverse views of Florida’s educators.

Phase III – Quantitative

The primary research question for the quantitative phase of my study is: What effects do the reception of failing grades have on academic outcomes other than standardized test scores? I include both student-level outcomes such as attendance rates, drop-out rates, and graduation rates as well as the educator-level outcome of teacher separation. I planned a quantitative phase for a variety of reasons: first, to ensure that I explored these non-test outcomes even if they were not raised during the qualitative phase; second, to balance and “triangulate” the qualitative phase; and third, to align with the quantitative measures considered and published by the state of Florida.

The Dataset

My dataset includes the years 2010 through 2014 ($t=5$). While the first grades were assigned in 1999 and data are available as recently as 2018, I chose the years 2010 through 2014 for several reasons. First, the same school variables are mostly consistently available for these five years. Second, these years follow a 2009 change in grading formula and precede the 2015 introduction of a new state standardized test and the one-time waiving of policy consequences.² For analytic purposes, this is a period of relative stability.

² The 2009 formula change is explained in depth in the policy section of my literature review. Most importantly, certain triggers for rises and drops in grades were eliminated and the formula for high schools was diversified to include scores on various end-of-course exams, graduation rates, and accumulation of college credit.

Data on Florida's schools is both highly accessible and frustratingly obscured. A wealth of variables is available on Florida's website in Excel format.³ Unfortunately, the formatting of these spreadsheets differs from year to year, as do the number of schools represented and the school names.⁴ Some measurements are available for some years, but not for others. For example, the Department of Education provides a "total graduation rate" for each high school before 2010, but after 2010 only provides the graduation rate of each racial subgroup, without a total. Likewise, the racial makeup of each student body is only available post-2014. Florida only produced the "Retention of First-Year Instructional Staff" – a variable of high value to this study – from 2008 until 2012. Another valuable report – on discipline data – is unavailable prior to 2013.

Other variables are reported in inconsistent ways. For example, student free-and-reduced lunch status – the primary means of gauging financial need – is available as a percentage in some years, but as a student count in others. Data on English-Language-Learners is available by school, race, and gender for some years, but only by district and state for others.

These inconsistencies made compiling a dataset extraordinarily difficult, frustrating, and time-consuming. Ultimately, I constructed a dataset via consolidated spreadsheets from the Florida Department of Education website together with some more organized spreadsheets purchased directly from the state. I point out the laborious nature of this process

³ See the archives page at: www.fldoe.org/accountability/

⁴ In some cases, discrepancies exist because schools closed or new schools opened. In other cases, schools are missing for unexplained reasons. Some counties or schools changed names from year to year, others did not change but have been recorded in different ways (e.g. "Elementary School 123" versus "Elem. Sch. 123").

in order to emphasize that not every school, or variable, is included. I intentionally excluded schools without a published grade for each of my sample years. I made this decision partly because a failing grade is my primary independent variable of interest and partly because I do not know the reasons for the missing grades. I did include some schools officially listed by the state as UG (ungraded).

I also excluded online/virtual schools and juvenile-detention facilities. I made this decision primarily because – in a variety of ways – these schools are not comparable to traditional public schools. This incomparability is partly reflected by differences in or absence of key variables such as attendance rates and drop-out rates. Determining which schools to include drastically changes the picture of Florida’ schools. For example, NPR reported in 2013 that 53 schools received F grades in 2012, while I report 33. For 2013, NPR reported 107 failures, I report 101 (Borges, 2013).

The schools included (N=3,255) comprise an unbalanced panel, a type of panel data in which “we have an unequal number of time periods per individual (e.g. due to missing observations)” (Mehmetoglu & Jakobsen, 2017, p. 229). This is due partly to the occasional unexplained missing variable, but primarily to the lack of graduation and drop-out rates for elementary and middle schools, to which the measurements do not apply. For a list and description of the variables included (all for five years), see Table 3.

Table 3

Variables in Initial Dataset

Variable	Number of Observations	Description
School	16,275	the school name
County	16,275	the county/school district (out of 71 different counties)
Year	16,275	the academic year (2009-2010 to 2013-2014)
High School	2,425 (yes) + 13,850 (no)	indicates if school is a high school
Letter Grade	16,275	the school's letter grade (A-F)
ADA Rate	15,196	the average daily attendance rate
White Drop Out Rate	2,252	the percentage of white students who dropped out
Black Drop Out Rate	2,221	the percentage of black students who dropped out
Hispanic Drop Out Rate	2,204	the percentage of Hispanic students who dropped out
Asian Drop Out Rate	1,660	the percentage of Asian students who dropped out
Total Drop Out Rate ⁵	2,320	The total percentage of students who dropped out
White Graduation Rate	1,916	the percentage of white students to graduate
Black Graduation Rate	1,750	the percentage of black students to graduate
Hispanic Graduation Rate	1,695	the percentage of Hispanic students to graduate
Asian Graduation Rate	791	the percentage of Asian students to graduate
Total Teachers	15,183	the total number of teachers
Teachers Separated ⁶	15,183	the number of teachers to leave the school
Teacher Separation Rate	15,183	the percentage of teachers to leave the school

⁵ Florida also publishes Indian and Multiracial drop-out rates. I do not include these rates because, in most cases, the schools either have no reported students, or list a # or * to indicate a number of students fewer than 10. These students are presumably captured in the total drop-out rate. For the same reasons, I exclude similar categories of graduation rates.

⁶ Florida breaks separated teachers up depending on the reason for their departure. This may be valuable information but in most cases the numbers do not add up, suggesting that schools did not identify the reason for many departures. Consequently, I believe the break-downs could be misleading.

As I detail throughout this dissertation, qualitative evidence suggests that the reception of a failing grade represents a shock to the school. Either the state's arguments are correct and this shock leads to improved outcomes, or the state is wrong and outcomes remain the same or decline. In order to try and estimate the effects of receiving a failing grade, I added a dummy variable (Fail Grade) to indicate whether or not a school failed in a given year.

The reception of a failing grade represents a shock in at least two ways: first, it triggers policy responses; second, it may affect the decisions and attitudes of teachers, students, parents, and administrators. For example, it may prompt a teacher to leave the school. I suspect that the second type of effect may be delayed, as grades are often released during the summer, or sometimes even the fall of the following year. Rather than use conventional lagged effects, I created a second dummy variable (Fail Prior) to indicate whether or not a school failed in the previous year, as well as a third dummy variable (Fail 2-Years Prior) to indicate whether or not a school failed two years prior.⁷ I took this route because if a school rebounded from a failing grade and earned a C or even an A or B grade in a subsequent year, any effects of the failing grade would likely dissipate. The purpose of these three dummy variable is to try to – controlling for other factors – estimate what impacts a failing grade has on schools, and when those impacts take effect.

⁷ Although not included in my dataset, I used school grades for the years 2008 and 2009 in order to create accurate Fail-Prior and Fail 2-Years Prior variables for 2010 and 2011).

Hypothesis and Research Design

It is impossible to isolate any one factor as causal in a social setting, perhaps especially in a somewhat chaotic school environment. With this said, my literature review and qualitative data both suggest that failing grades discourage teachers and students. Between 2010 and 2014 this discouragement is not reflected in the data on mean drop-outs and attendance, since both changed in a desirable direction despite an increase in failing grades. However, mean graduation rates decreased and mean teacher separation increased – both undesirable trends.

My goal is to determine whether a fall in a school's grade to a D or an F is associated with expected changes in indicators – or dependent variables – other than test scores. Impacts on these variables are estimated using a series of fixed-effects models. Ideally, I would design an experiment with a control group of schools not assigned school grades, however no such group of (public) schools exists. Using fixed effects models, it is possible to fully control for un-measured characteristics of schools that do not change over time along with observed time-varying school characteristics, including school grades (Allison, 2005). With multiple years of school data, a fixed-effects model allows me to isolate a school's failing grade as a possible cause for changes in dependent variables measuring student and teacher outcomes over time.

I aim to test the null hypothesis that the outcomes of interest are the same before a drop to a D or an F grade, as they are after such a drop, against the alternative hypothesis that the outcomes of interest will change following the drop. A fixed-effects model has both

advantages and disadvantages in this context. To my benefit, fixed-effects control for all time-invariant variables, which may eliminate spurious relationships. This is necessary since my dataset is missing many possibly explanatory variables (e.g. free and reduced lunch status). Unfortunately, this means that I cannot estimate these variables' effects (Allison, 2005; Mehmetoglu & Jakobsen, 2017). As I am primarily interested in the effects of a failing grade – regardless of between-school variations – a fixed-effects model should be appropriate. While I do not have some demographic information – such as racial and economic make-up of schools – I assume that these remain relatively unchanged within each school between years. This is bolstered by the zoning of schools by neighborhood – often defined by housing values and thus incomes.

CHAPTER 4

PROBLEMS WITH FLORIDA'S SCHOOL GRADING POLICY AS REPORTED BY TEACHERS

In this chapter, I introduce context for teacher positions on school grades and outline their primary complaints. My study participants, as well as actions by the Florida Teacher Association, suggest that there is antipathy between many teachers and state policymakers. Based on this dynamic, teachers are doubtful that any policymakers will listen to their concerns or ideas and some believe politicians secretly want to close public schools. Despite these doubts, teachers are not opposed to accountability. They do see school grades as confusing for parents and teachers and as insufficient measures of student learning. They also think that grades ignore the strengths of students and teachers. Finally, I discuss teacher beliefs that schools have different missions depending on their student populations, something they want grades to reflect.

“A Really Strange Benchmark”: The Transition to School Grades

Theresa teaches advanced placement English courses at a top-rated high school in an urban area of Florida. This is her 23rd year teaching. She is white, as are most of her current students. Before her current job, she taught drop-out prevention and various English courses, as well as “whatever they threw at me” at a school she describes as “primarily students of color” and “even more urban.” Of all the teachers I interview, Theresa has the most experience, including National Board certification. She also has insight from having taught at both high and low-rated schools. Additionally, she can recount her experience of the policy transition to modern-day school grading. In her recount, she is unabashedly critical of school grades – “a sea change in the way we addressed children” – and contrasts them with “a much more inclusive approach to educating students in the early 1990s.” She explains:

We really couldn’t believe it was actually happening, about rating schools, kind of like you’d grade deli meat. It seemed really, really weird. The school where I was teaching was a very, very poor [high] school, with approximately 2,400 students. The average age of a student at that school, when I was there, was 16.7 years, which is higher than the median. Most of the students who were there needed social services. Most of the students who were there struggled with food insecurity, issues of homelessness or intermittent homelessness. So, we’re trying to graduate them. We’re trying to graduate them and help them find career paths.

We had a strong vo[cational]-tech program. When Mr. [Jeb] Bush inculcated the grading system along with FCAT, vo-tech took a hit. A lot of the students who could be successful as an auto tech, or as an air conditioning repair person, suddenly found themselves in an endless loop of failure on FCAT tests, that tested academic types of benchmarks, as opposed to some of what we now call “soft skills.” The ability to transition to work, career took a back seat to college prep. That started immediately. The school [where I worked] immediately received an F.

Theresa goes on to talk about school grades at a more macro level:

The letter grades sold well to the public. The letter grade was easy, because everyone's been in school in some capacity. Even if you're homeschooled, you know if you're an A student or a B student. Five-year-olds know what an A means. Five-year-olds also know that when you have an A group, a B group, and a C group, just by ability, they figure out very quickly who is who. When Jeb Bush and his foundation – which is still very active, and I think, very poisonous – when they were putting out press releases, they were easy to read. It was nice and graphic. Parents were making real estate decisions based on those grades. That ends up having a decade-long – and more – effect on school populations.

I ask if, regardless of the grades' impacts, her school's F grade was accurate. Theresa says “no.”

It was based strictly on one benchmark. It was based strictly on tenth grade FCAT scores, which is a snapshot of one day of one slice of the school. Nothing else was taken into account at the time, it was purely FCAT scores. Now it's a complex algorithm that changes at the whim of politicians. Now it involves AP pass rates, which is ridiculous, AP participation rates, which is even more ridiculous, graduation rates, which are largely inflated.

It was one benchmark, not a clear reflection of what the school was capable of delivering. It also seemed to be a really strange benchmark. We had used something called the HSCT test for years. It tested basic, what-you-need-to-function-in-society skills: computational mathematics, communication skills. It was a doable test, and if a kid didn't pass, we were allowed to sit one-on-one and remediate them through the skill they missed and then give them a certificate to pass. With FCAT, it's one and done, and with FCAT, it's reading passages about HMOs versus PPOs, or doing math that's very abstruse, high-level math that they honestly are not going to use as general citizens. That sea change was the first shift, but then using that as a benchmark, as the only measurement...I don't know any business on the planet that only has one yardstick. None.

A Strained Relationship: Teachers and the State

I quote Theresa at length because she vividly recalls the 1999 policy transition, but also because she eloquently expresses many of the themes ubiquitous in my interviews. Most importantly, for this chapter, she makes emphatically clear that she does not support the school grading policy, nor does she think it accurately represents school quality. Of the twelve school employees that I meet with – ten teachers and two principals, ranging from F to A schools – two feel ambivalently about school grades, the other ten agree unequivocally with Theresa. In my six years of teaching in Florida's schools, I met very few teachers who disagree.

As evidenced in my introduction and literature review, standardized tests have a controversial history in Florida and still alarm many teachers. While the majority of teachers I speak with echo those concerns, a couple voice support for standardized testing as an accountability measure, others feel ambivalent about it. Despite that ambivalence, the educators I interview all believe that school grades are inaccurate. In other words, even the teachers who support standardized testing doubt the current school grading policy.

Some education researchers like Mark Paige, worry that a policy cannot succeed without the support of those charged with implementing it. Paige (2013) speaks to this problem in relation to teacher collective bargaining rights. As of 2011, the Florida Education Association cannot negotiate teacher contracts, which are now solely the state's jurisdiction.⁸

⁸ The Florida Education Association (FEA) is a statewide association that identifies itself as an education union but lacks collective bargaining rights as per Fla Stat. 1012.34(1)(a) (2011). FEA is voluntary and currently claims roughly 140,000 members. For more information: <https://feaweb.org>

Paige calls the lack of bargaining rights a democratic “paradox.” By making teachers accountable directly to the elected state government, voters theoretically have more control. However, if the union has no “skin in the game” (p. 29), then teachers are cut out of the policy making process. That exclusion leads to teachers resisting or even obstructing policy, instead of instituting it.

Clearly, a sizeable number of Florida teachers feel excluded from – or even persecuted by – state policy making. Since adoption of the A+ reforms in 1999, teachers or organizations working on their behalf, have sued the state over school vouchers (2000), funding adequacy (2009), tax-credit scholarships (2014), teacher bonuses (2015), teacher licensure exams (2017), and charter school funding (2017), among other issues.⁹ In 2010, the Florida Education Association (FEA), together with individual teachers, launched a state-wide social media and protest campaign designed to prevent the passage of Senate Bill 6, a comprehensive education reform bill. While it was ultimately vetoed by Governor Charlie Crist, it was succeeded by a similar bill signed by Crist’s successor, Rick Scott.¹⁰ These

⁹ In 2000, the Florida Supreme Court ruled vouchers unconstitutional. For more on this see Borman and Dorn (2007). For a contemporary (and national) reaction see Jodi Wilgoren, “School Vouchers Are Ruled Unconstitutional in Florida.” *New York Times*, March 15, 2000. Later, in 2001, the Florida legislature created a new tax-credit scholarship program that exchanges state tax credits for corporate donations, with which students can attend private schools. The FEA challenged the program in *McCall v. Scott* under the grounds that it helped the state fund religious schools, but in 2017 the Florida Supreme Court declined to hear the case. Most scholarships are administered by the non-profit Step Up for Students. For more info see the Step Up website at www.stepupforstudents.org.

¹⁰ Scott signed S.B. 736 (colloquially known as “the teacher tenure” bill because it – among other things – discontinued tenure which teachers had previously acquired after three years). In 2017, Scott signed another controversial law: House Bill 7069, which introduced “Schools of Hope” or charter schools given financial incentives to locate near failing public schools. At least five Florida counties sued the state over H.B. 7069 in 2017. The other lawsuits I mention refer to: (1) *Citizens for Strong Schools v. Florida State Board of Education* (which can be read in its entirety at <https://law.justia.com>); (2) an ongoing struggle between aggrieved teachers and the state, together with the Pearson testing company, over how teachers are tested for certification; and (3) The “Best and Brightest” Scholarship program, which awards money to teachers based on their evaluations as well as their scores on the SAT or ACT – which many teachers see as an unfair measurement. All of these lawsuits can be explored in depth in Florida’s major newspapers, including *the Orlando Sentinel*, *the Tampa Bay Times*, and *the Miami Herald*.

struggles demonstrate a large disconnect between what state leaders believe is best for students and what teachers think. It also – as Paige warns – shows a large diversion of time and resources toward legislative battles and away from the education of children.

Elsewhere in my dissertation I more fully explore conflicting beliefs on standardized testing, teacher accountability, charter schools, and other issues being fought about in the courts of law and public opinion. For now, I primarily want to place teacher mistrust of school grades in the larger context of policy disagreement. The actions of state leaders suggest that the state does not trust teachers, and teachers have been open – with me and with the press – about not trusting the state. The words of teacher participants in this study should be read with this animus in mind. This is a strained relationship encircling a simple question: what constitutes a “good” school?

We Just Can't Win: Teacher Frustration with Shifting Standards

The formula for grading Florida's schools has changed nearly every year since 2000, and while state averages have improved over time, some failing schools remain stubbornly static. Other schools have seen their grades fluctuate from year to year, sometimes corresponding with state policy changes. In 2015, Florida waved penalties for failing schools in recognition of testing flaws associated with the new Florida Standards Assessment (FSA) test. Additionally, federal requirements have changed through No Child Left Behind (NCLB, 2001), the Obama Administration's Race to the Top initiative, and the Every Student

Succeeds Act (ESSA, 2015). These state and federal policy changes have given some teachers and principals a suspicion that the goalposts will always be moved.

Alex, who has a doctorate in Education from one of Florida's largest universities and briefly worked as a principal at a failing urban middle school, has first-hand experience of this policy change. He is a young, Latino man who taught in Florida for five and a half years before becoming an administrator for three, and finally a principal. We meet in a temporary office void of any personal effects or professional merits, presumably they fill the cardboard boxes against the wall. He is in the process of moving after being transferred – for reasons unknown to him – from his former school after 16 months. He speaks clearly about the shifting expectations of the state and the “latest iteration” of the formula. While he gives the state credit for trying “to make a system that is equitable for all of the schools,” he sees the end results as unjust. He explains:

The way that [school grading] is done, I don't think is equitable. Just because we have so many schools at so many different starting points. And I can say that, because when you look at a breakdown of the schools, who are the A schools, B, C, D, F, I mean, it's almost like a trend. Like, it doesn't change that often and when it does, the state changes how they grade the schools. And that's what happened before, because they said there were too many schools that were getting As and Bs. So, [they say]: “Clearly we're doing something wrong. So, we have to change, we have to modify it again so we can bump some people back down.”

But your more affluent schools end up on the A side and your inner city, more struggling, lower sense of economics schools, they end up on the failing side. And I don't think that's the intent. I think it's an unintended consequence. But, there really is, in my opinion, no genuine way to fully grade the schools if we're going to use testing as an indicator.

I cannot find any cases of a Florida state or county official stating that too many A schools requires changing the formula, but five of the teachers I speak with say something similar. I explicitly ask Joy – a teacher at Alex’s school who started in 2013 and experienced the 2015 change in grade formula – how that change affected the school:

We’ve stayed down. We’ve stayed down. I think it just made it harder for us to go up. Because I was hearing [that] if it wasn’t changed that we would have gotten a C, apparently, because [the state] changed it, it kept us at a F.

I ask if the staff feels like they “cannot win”? I wonder if it is disheartening to think that the school might improve, only to have the state change the rubric. Joy nods her head enthusiastically and says, “Yeah, there is that, to some extent.” I point out that state legislators and governors have argued that raising expectations each year motivates teachers and students to work harder and improve more (see: Education Next, 2014). Joy agrees that some staff feel motivated, but she thinks others feel like their efforts are futile and ignored, that if the school is going to fail anyway, why try?

Alex and Joy suspect that the system is rigged against their school. Additionally, some teachers feel that it is rigged against them, personally. Just over half of participants suspect, or have heard rumors, that their counties instructed school administrators to rate fewer teachers “highly effective,” which saves money from teacher bonuses, since highly-effective ratings qualify teachers for a “Best and Brightest” state scholarship. Florida politicians and many county officials strongly deny these claims and tout the non-biased nature of grades as a strength. In April 2018, former governor Jeb Bush claimed “[Florida’s

test score improvement] required a willingness to be bluntly honest about the state of our schools” (Bush, 2018).

While politicians praise school grades, the participants in my study mostly voice agreement with Theresa’s critiques and the frustrations of Alex and Joy with changing expectations. Additional complaints coalesce into four themes related to grade accuracy: (1) the school grading formula is not well understood by teachers or parents; (2) school grades do not give adequate weight to student improvement; (3) grades do not represent the efforts of teachers and students; and (4) grades do not reflect the most important challenges faced by schools.

In my experience, some Florida residents and politicians characterize these complaints as “whining.” They feel that teachers want to avoid accountability and that the Florida Education Association – commonly but mistakenly called a “union” – seeks to protect bad teachers and the status quo in general. To the contrary, teachers I meet do not oppose accountability or policy changes. Rather, education may be unique among professional fields insofar as teachers disagree with policymakers about what constitutes a job well done. They see accountability – in the form of standardized tests for students and grades for schools – as not only misguided, but as a ruse, a ploy to make them look bad.

Teachers view the Republican dominated Florida state government as hostile to them and to public education, as supporting charter and private schools at their expense. As an analogy, imagine if postal workers defined success by delivering mail efficiently and on time, but the government defined it by post office profit and advocated encouraging

customers to switch to FedEx. If the remainder of this chapter accurately captures teacher perspectives, then the policy solution to this divide is not no accountability, but a more authentic accountability that measures schools in a holistic, contextual way.

These themes also dispel the notion that school grades represent a fringe, insignificant policy question. Some rank them relatively low on the list of concerning education issues. This dismissal does not make sense given the extreme importance of school quality to families. Researchers have found that some make some of the biggest financial decisions of their lives – where to buy a home, whether to pour savings into private school tuition – based on school reputations (Figlio & Lucas, 2004). If parents no longer care about grades in the same way, does that represent a loss of trust in the system or a change in values? Moreover, teachers argue that grades spur policy responses and affect their job stability, school satisfaction, and potentially their salaries. Unfortunately, grades are only part of an education policy package that bothers educators. I will explore each of the aforementioned themes at length, but first I think it is important to place them in the context of teachers’ oppositional relationship with state policymakers.

“They Made Them Fail”: Teacher Beliefs in a “War” Against Public Schools

The initial logic of school grades relied on their being an informative and reliable tool for parents to gauge school quality. In 2014, reflecting on his hallmark grading policy, Jeb Bush described it as: “a straightforward and transparent A-F grading system to give parents simple and accurate information about the effectiveness of their schools.” This accurate

information should, according to former Governor Bush, “inject competition into the system through school choice” (Education Next, 2014). As Theresa points out: “The letter grades sold well to the public. The letter grade was easy, because everyone’s been in school in some capacity.” Other researchers have also concluded that school grades are easily understood symbols that – for better or worse – have significant impacts on parental decisions, including where to live (Figlio & Lucas, 2004; Goldhaber & Hannaway, 2004; Rouse, Hannaway, Goldhaber, & Figlio, 2007). Teachers confirm these findings in our interviews but are troubled that parents do not understand the changing metrics. They question how the system can be “transparent” if the formula is esoteric. Most of all, they worry that the system is actually cover for a campaign against public education.

This suspicion arises, in part, from what teachers see as less critical assessment of private and charter schools. In order to explore this, I speak with some parents who made the switch for their students. Angela lives in a densely populated, urban area of Florida. She is a single mom with four kids enrolled at private schools. She cannot afford private school tuition on her own; her kids receive funding through Florida’s tax-credit scholarship program. They formerly attended a public elementary school, where Angela says they “excelled.” So, I ask her why she decided to switch to private schools.

The reason why I send them to private school is because the area we live in, the school that [the county] wanted them to go to middle school, it was a C school and I wanted my kids to be in an A school. I didn’t know anything else about it, just the fact that it was a C school, it was below my standards for their education.

Angela does not know how public schools are graded. She says that she “read about it a long time ago” but no longer understands. I ask how she ensures that the private school is doing a good job.

They do have something that’s equivalent [to the FSA]. I don’t remember the name of it, but they basically take a standardized test, I guess, and then they’ll let them know if they’re average, above average, below average, compared to the other kids in the United States.

I also speak with Carla – a friend of Angela’s whose children also attend private schools via the state scholarship program. She, like Angela, knows about school grades, considers low grades unacceptable, but does not actually know how schools earn them. Like Angela, she assures me that the private school is held accountable but when asked how – she speaks vaguely about her children’s grades. Neither mentions accountability for the school itself, although Carla says she trusts teachers who send home work that makes her think, “Whoa, okay. It’s advanced work, what they’re doing.”

Throughout my interviews with teachers, several mention parents’ different standards with frustration. They feel that private and charter schools are often perceived as better than public schools without facing the same public accountability; furthermore, they believe that ignorance and confusion about school grades is being exploited in order to close public schools (see also: Prothero, 2017).

Education Secretary Betsy DeVos as well as President Donald Trump have lauded Florida as a model for the nation. Ironically, and to the disdain of public education advocates, they joined Governor Scott as he signed controversial education reform H.B. 7069 – not at a

public school – but at St. Andrew’s: a private, Catholic school (Camera, 2017; Strauss, 2017). Trump has never visited any of Florida’s public schools, even though he has a residence in the state. Scott Maxwell – an editorialist with *the Orlando Sentinel* – often criticizes state education policy in his column, *Taking Names*. He points out that Trump, DeVos, and Scott lavished praise on St. Andrew’s, which participates in the state’s scholarship program. He writes: “*This* was an example of education done right, reformers said. *This* is how education should be done.” He continues:

Well, would you like to guess how many of Florida’s top-down, test-obsessed, bureaucracy-laded rules applied to this school? Virtually none. So Florida politicians champion non-stop testing and mandates for public schools – and then tout private schools that are exempt from those rules (and where they often send their own kids) as examples of success (Maxwell, 2017).

Some teachers I speak with believe this accountability imbalance is an intentional “war against public schools.”

Theresa not only believes that such a war is being waged, she believes it will entrench existing inequalities. She explains that if a parent can afford to leave a failing school for a “better” one, they will. Those same parents are “financially invested and have the time to donate to the school,” which will enrich their children’s new schools but sap the public schools of support. She explains:

I think that money and privilege is money and privilege in any capacity. The letter grade system has created another layer of privilege in some communities that’s deeply problematic. The Jeb Bush Foundation’s goal, if you actually go on their

webpage, is to create a choice system, which is a euphemism for charters and vouchers for kids. They say that public systems are failing. They made them fail.¹¹

Teachers see evidence of such ill-intent on a national scale. In 2004, Rod Paige, then Secretary of Education in the Bush Administration, called the National Education Association a “terrorist organization” (Pear, 2004). In 2017, *the Atlantic* published an article under the heading – “The War on Public Schools,” which quotes members of the Bush, Obama, and Trump Administrations disparaging public schools (Christakis, 2017). *The Washington Post*’s Valerie Strauss, writing about the passage of H.B. 7069 – the “Schools of Hope Bill” – calls the bill a “scam” that “sparked a tsunami of opposition from parents, school boards, district superintendents and unions.” She writes:

They [the parents, school boards, etc.] argued that it will harm traditional public school districts, threaten services for students who live in poverty and curb local control of education while promoting charter schools and state-funded voucher programs.

When the legislation becomes law, more than 100 traditional public schools given low grades by the state will be converted into charters – even though the charter section in Florida is deeply troubled (Strauss, 2017).

I know that many teachers and parents – including the parents I speak with – believe that charter and private schools offer valuable help to their children. With that said, it is apparent that many public school teachers feel like politicians are for alternative schools but against them. Voices in the media echo and amplify those feelings. The following themes should be considered with this oppositional relationship in mind. Teachers consider school grades to be

¹¹ The “Jeb Bush Foundation” that Theresa refers to is ExcellnEd

part of an antagonistic relationship with the state and they do not critique grades in neutral tones. In our interviews, they often sound oppositional, angry, and defensive. They fear that a public lack of understanding or even apathy about grades creates an opening for opponents of public schools to exploit.

Straightforward and Transparent: School Grades as Reliable Indicators of School Quality

Do school grades fulfill one of Bush's stated objectives – to adequately inform parents of school quality? My interviews provide no one answer. Grades clearly hold symbolic value for some parents and are used as indicators of quality; yet, those same parents admit to not understanding the grades. That lack of understanding upsets teachers who want their work noticed. Since teachers see grades as unreliable, they want parents to consider the underlying factors instead.

Most teachers tell me that they do not know whether parents understand grades, but feel certain that parents do not care. Others say that parents cared and understood at one time but have grown cynical as the grading formula has changed. Media accounts often feature or quote parents with open hostility toward grades. While this may seem like a separate issue from understanding the formula, teachers see it as connected. They argue that confusion breeds apathy and they link understanding grades with caring about them. For example, David – a fifth year teacher at an A school – tells me:

When they changed the formula in 2013 or 2014 – I can't remember – but when they changed it I was relatively new and every time it came up in a meeting I had to

Google the formula because I couldn't remember it. So, I can't imagine that parents really know it. So, maybe an F grade looks bad and an A grade looks good but if a school goes from an A to a B or an F to a D, do parents care about that? I think they probably don't really know the difference.

According to David, parents who do not understand the grading formula might recognize and care about the extremes, but feel indifferent about changes from one grade to another.

Sandra and Carla trusted grades sufficiently to remove their children from a C school. However, they do not know how grades are calculated. Booher-Jennings (2005) claims that in 2002, the grading system was "impenetrably complex for most educators and the general public" (p. 104) and that, while teachers understood the 2002 formula change, most parents remained unaware of it. Additionally, Florida publishes both state and federal statistics, each calculated differently. A majority of teachers in my study believe that, partly due to confusion and grade inscrutability, most parents actually rely more on anecdotal information than they do on the school's grade.

Theresa's high school dropped from an A to a B in 2016. I ask her – what was the parents' reactions? She says: "nothing."

The parents were like, "It's Bryan High." In fact, the parents said, "Oh my gosh, you're not getting a bonus, can we help?" Bear in mind, my parents are not normal parents. I heard not a peep from parents. The kids were like, "we got a B, ha ha ha, B for Bryan!" We were a C for three years in a row. Our principal was kind of mildly concerned about it, but said, "well, these are just the numbers." He was a different guy [than our current principal].¹²

¹² Bryan High is a pseudonym to obscure the identifying name of the high school, while retaining the students' alliteration.

Theresa believes that because her school is an older school, with a good reputation, involved parents, and a supportive principal – the parents did not care about the grade drop, at least not to a B. Moreover, she thinks that parents no longer trust the grading system.

I'm sensing a weariness [among parents]. I think that as time has gone on, they're getting very jaded and cynical about it. Tallahassee will put out a press release saying, "we've adjusted the scale." What? Okay. I think parents sometimes don't know what to think anymore.

She explains that – at her school – parents care more about having their student in a particular teacher's class, as opposed to the overall school grade. She says that parents "shop" for teachers and mentions some teachers, including herself, who are more highly valued. "It creates a weird situation," she explains, "because the parents are like – 'don't worry about the letter grade, it's fine, the school's whatever, but you're great.'"

Jane, who also teaches advanced English students at Bryan, shares a similar experience. She does not think letter grades accurately measure school quality, but she cannot think of any symbol that could. I ask her what parents should know about how the school is doing.

How the school is doing? Mmm. How the school is doing...I guess my first inclination is to say what's more important? Is it more important what [classes] the school is offering and what the environment is and what the opportunities are for kids in that particular school? I guess I would say – *what*¹³ the school is doing...I guess I don't know what would be the better way.

¹³ My emphasis added

It interests me that Jane changes “how” into “what” – which seems to reinforce her view that classes, school environment, and opportunities matter more to parents than the school letter grade.

Alex also suspects parents increasingly find the school’s grade irrelevant. He describes grades as already “baked in.” He explains:

I just think that at this point, it’s almost, and I hate to say it, but I feel like it’s just baked in, how it is, and it’s just the go-to. [My former school] will always be perceived as a failing school. And I think even if it got to even a C, and it has been a C before – it was a C many, many years ago – but even if it got to that level, I think there might be a perception of – well, something must have gone wrong then, because it can’t be, this is a failing school. You know, it’s a D, F school historically. That, that’s what it should be, in the mind’s eye of the community.

Jessica, who succeeded him as principal, also thinks that parents mostly ignore grades, but sounds more emphatically positive. She explains – as do some teachers at her school – that their school is a “legacy school,” that students’ parents and even grandparents attended. She says this family history gives the school a “familial quality” regardless of the grade.

These thoughts differ from what school employees have told other researchers. For example, one A school principal told Goldhaber and Hannaway (2004) that parents cared about school grades and applied substantial pressure on teachers to “do whatever it might take to perform well on the test” (p. 601). The principal added that “even though they know better than to believe that any one test can show how good we are, they still want us to get an A” (p. 601). Parents did not want excess time committed to test preparation, as in the case of one elementary school that cancelled all field trips and projects scheduled prior to the FCAT.

Yet, “while parents lamented this and knew their own child’s life chances would not be affected by his or her test results, they nonetheless claimed to be ‘type A personalities’ who ‘still wanted to see their school with the highest grade’ (p. 601). They report that one school dropped from an A to a C and its principal had to “expend considerable effort” (p. 601) convincing parents that it was still a good school.

Certainly, each school is unique and much has changed since Goldhaber and Hannway’s interviews in 2004. Maybe, as Theresa suspects, parents have become “jaded” about grades, or as Alex says – grades are “baked-in.” In 2013, NPR’s State Impact reached out to Floridians online and received a myriad of thoughts on school grades, all negative. One parent believes, “there is no emphasis on critical thinking skills, creativity, ethics, logic, and appreciation for learning.” Another says that her son attends a “ghetto” school but “in spite of the fact it is an excellent school” and that “they know my son by name.” She wants a “Zagat of schools.” Other parents say that the grades are “skewed” and “meaningless” and “a disservice to the teachers” (Mack, 2013). Teachers I speak with at a failing middle school hear many of these criticisms from the few parents they talk with, but they say that while some parents are hostile, most appear apathetic.

Sandra teaches seventh grade science at a persistently failing, urban middle school – the school Jessica now serves as principal. This is her second year at a job she calls “interesting to say the least.” She is a young woman of color who ultimately wants to attend dental school but entered education temporarily via Teach for America, a nonprofit organization that places teachers in struggling schools. I ask her about parents.

Oh, so hard to reach them. They either work, their phones are disconnected, wrong number, never answer, or [they are] just unresponsive. In terms of – if I talk to them – nothing changes with the student.

Sandra believes that “nothing changes with the student” partly because parents distrust the school – a very different perspective from her principal who speaks of the school as a “legacy” and “familial” school. She describes one student with whom she has “a really good relationship” but when Sandra tried to explain a problem to the child’s mother, the mother blamed the school and said, “you teachers are always picking on my daughter!”

And I was like – I’m sorry if you’re under that impression, but that’s not the case. I actually gave her a journal to write in so she could write when she’s upset. And her mom was not having it all, and I was like – wow – that’s shocking coming from this girl’s parent. But then again she probably also has been getting calls a lot about her child.

Rhonda, who teaches next door to Sandra, also has trouble getting through to parents. Parents she does see distrust the school. She explains:

I remember at the beginning of the year there was some parents that were coming to withdraw their students, and so some of the parents wanted to know: why should I bring my kid to this school, or why should I leave them here? Some just knowing that they were getting ready to withdraw the student.

Rhonda says that these parents usually do not know anything about the school, beyond its grade. She thinks this deprives kids of opportunities. Unlike Alex, she thinks that parents do notice the grade and that it leads parents to think that the school is homogeneous. In fact, she

argues, there are essentially many schools in one, with different levels of academic success.

Rhonda continues:

One student in particular, she was a part of the Calculus Project, and so she would be with a cohort of, of students for her algebra and geometry class. [I was] just trying to let the mom know, “Hey, you know, she is a part of a community of students, and it’s a small community.” If you make this school small to the students, then it makes it easier for them to be able to make connections with other people, and just get involved with different things around campus.

Rhonda believes that the student’s mother could only see the F grade and consequently, the student lost an important opportunity.

Alex – who calls grades “baked-in” – started the school year as Sandra and Rhonda’s Principal. He acknowledges that the F grade can blind some parents to their children’s opportunities and individual challenges. He explains:

It’s almost like a go-to, like if there’s a discipline issue at the school, a parent comes in and, we’re like, “You’re kid’s getting in-school suspension. You’re kid’s getting out of school suspension,” whatever the situation is. The parent gets upset and says, “Well that’s why this is an F school!” And it almost becomes – well, I’m just going to rely on that as, you know, a defense mechanism.

Alex believes that, rather than deal with a child’s problems, parents sometimes blame the school, and an F grade gives them ammunition. He continues:

But they don’t really understand what [the grade] means. And, it’s partially because it is kind of confusing. I mean, I know educators who are confused with how the whole system works, because there’s so many different components to it, you know? With your different point categories, with your buckets per gains, your different groupings of kids...it’s confusing for educators.

Jane agrees. She asks, “How could a parent understand it if we can’t even understand it?”

Alex continues:

Parents care about the bottom line. They don’t care about points. They don’t care about what your learning gains were for reading or math, or how civics did. In my experience, they really just care about the grade. So, an explanation can get lost on them.

Some argue that parents do not really need to know how schools are graded, but many teachers feel that not knowing, or not understanding, creates confusion, distrust, and blindness to the nuances of a school community – blindness that deprives students of opportunities, or breeds apathy.

Ignoring Growth: The Failure of School Grades to Capture Within-Year Learning Gains

The second theme that emerges from my discussions of grade reliability, is a perception that the metric poorly measures student improvement. Nearly every teacher I speak with offers some version of the refrain: we do not get credit for student growth. For example, Rhonda calls student growth expectations one of the largest frustrations of teaching. She explains:

I think a lot of times teachers become frustrated. When you’re talking about students that can’t read, not to say that all our students can’t read, but we have a percentage of students that are taking a reading class, or you have students in double block reading.¹⁴ If I have this kid that’s in eighth grade that’s reading on a fourth grade level, and I take them from reading on a fourth grade level to a fifth grade level,

¹⁴ In some counties of Florida, students who score below a certain benchmark on the FSA must take an extra reading course in place of an elective. When Rhonda refers to “double block reading” she means a student with two reading classes.

they've had a years' worth of growth, but that's not measured on the FSA, and so that's where the frustration comes in.

Rhonda's principal, Jessica, echoes these frustrations. I ask her if she agrees with the state, that the school is failing. She says:

No [I don't agree]. I'm here every day. I'm seeing the good that my teachers are doing. We have kids who might be on a first or second grade level, and we're moving them to a third grade level, but that's still not enough. And, you know...you can't go from first grade to sixth grade or eighth grade. It's just not possible, but we're moving kids, which is why I don't think the school is failing.

Both Rhonda and Jessica feel frustrated that the progress they make with students is not recognized.

Florida does try to recognize that progress. Dating back to at least the 1970s, educators have distinguished between criterion and norm-referenced tests. Criterion-referenced tests assess student performance on a particular task before and after instruction; while, norm-referenced tests rank students in comparison with other students (Mehmetoglu & Jakobsen, 2017). Florida's FSA is both criterion and norm-referenced. Each student's score falls into an achievement category but is also compared to their score the prior year. When used to grade schools, this method is referred to as "value-added", a concept that rose to national prominence with the No Child Left Behind Act (2001). The term refers to the value added to children's scores by each teacher or school (Goldstein, 2014).

Florida bases a quarter of a school's grade on "learning gains." Learning gains are measured relative to a five-level scale. Five is the highest and three is passing. There are four

ways to sufficiently make learning gains. A student can: (1) move up an achievement level; (2) move up a “sub-category” within levels one and two; (3) students who scored at level three or four may stay at that level while increasing their total points; or (4) a student who scored at level five can remain at level five. Despite these multiple options, participants in my study argue that growth is not sufficiently measured.

In 2017, the lowest score a student could receive on the FSA in eighth grade was 274 points. In order to sufficiently achieve “learning gains”, that student would have to score at 290 or higher the following year. What if a student begins so far behind that they cannot make that increase, no matter how good their teacher and school are? That is the question teachers in my study want answered. I ask Joy if she believes the state recognizes the points that students start at. Initially, she replies, “no, they don’t.” Then she stops herself, and elaborates:

Not truthfully. They say they do but they really don’t. Like, they don’t see if I had a child and they, at the beginning of the school year they’ve scored at the second grade reading level. The state, when they get that score, that two, three, four, five for FCAT or FSA, they’re not gonna go back and say, “Well, maybe we could drop this [target] score because actually at the beginning of the year this student was only at a second grade reading level, they’re at a third grade reading level now.” They say they do the gains, but I don’t think the gains count for anything if you aren’t showing proficiency at grade level.

Rhonda believes that Florida’s value-added measures are skewed in favor of affluent schools. She explains:

I was looking at the data for some other schools, at more affluent schools. What you typically find is that yeah, the school might be an A, but let’s look at the bottom 25 percent of the school, and this is where a great majority of your minority students lie.

And when you see the dynamics of the school shifting, it's because that bottom 25 has grown. And so people don't take that into effect. Yes, this is an A school, but how are your lowest performance kids doing?

Florida does calculate learning gains for the lowest 25 percent, but that component only accounts for 20 percent of a school's total points. Matthew DiCarlo, of the Albert Shanker Institute, reports that the gains of high performing students are "double counted" by additionally benefiting already high-performing schools (DiCarlo, 2017) So, schools that succeed in other categories can hide their lowest 25, while already failing schools cannot.

Florida has also faced criticism for attempting to "hide" the scores of English-Language Learners (ELLs). In 2014, Florida determined that while ELL students would count towards a school's test attendance, their scores would not count until they had been in the system for two years. This runs counter to federal requirements, a conflict that required Florida to seek waivers in its ESSA plan. DiCarlo points out that, ironically, the same arguments used by Florida to support ELL exclusion, undermine the way it counts learning gains. DiCarlo quotes Miami-Dade superintendent Alberto Carvalho: "[to expect] that after one single year of English-language instruction, a child who does not speak English would be ready to sit for an exam and demonstrate equal proficiency to one who was born in a community that spoke English is not only irrational, it is unfair and unreasonable." DiCarlo calls this "an excellent point" but asks:

By the same logic, however, what about a child from a disadvantaged background who is a native speaker but enters the school or the year far behind? Moreover, is it fair to punish schools or label them as "failing" if they happen to serve large proportions of these struggling students and fail to help most of them achieve well

over a year's worth of growth, and often 2-3 years, in a single year's time? (DiCarlo, 2014).

I did not visit any schools with large ELL populations, which might explain why no teachers explicitly discuss ELL scores with me. But, clearly, DiCarlo's questions about struggling students and yearly growth resonate with educators like Rhonda and Jessica.

One teacher at Rhonda and Jessica's school disagrees, at least in part. Lakisha is a black woman who has been teaching for 16 years but is new to her current middle school. She tells me that FSA is "in general" an accurate measure of her students' learning. She continues:

I personally feel that standardized testing is necessary, because at some point or another, when you leave high school, you're gonna have some type of standardized test. I do think it is necessary to really know for sure if we have comprehensive readers. If they are reading proficiently enough to be able to receive a high school diploma. Because you know what? The whole No Child Left Behind...not that we still completely focus on that, but we want to make sure we have people who are ready to go out into the world and work.

Despite her endorsement of standardized testing, Lakisha does not believe school grades capture her school's quality. She says that in order to fully evaluate students and schools "maybe we should do something a little bit more authentic."

Ignoring Talent and Effort: The Failure of School Grades to Measure Faculty and Student Strengths

Some teachers believe that, in addition to not reflecting student growth, school grades ignore teacher efforts and abilities, consequently obscuring an important measure of school

quality. Every teacher and principal in my study tells me that – in some way – school grades do not accurately reflect teacher work. For example, I ask Jane if the F grade reflected the abilities of kids or teachers at her previous, failing high school. She responds:

I don't think we would think it was an accurate picture of the teachers because I think we had a very academic faculty and a very bright faculty. We had a very diverse population [of students] so I would say we had a magnet program, engineering, and we had an AP academy that attracted kids. We did have a very intellectual faction and then we also had kids who were living in poverty and in gangs and being bused in.

Jane pivots from talking about the “academic” and “bright” faculty, to discussing special efforts the school made to deal with a “diverse population.” Other teachers also frame the issue this way. Sandra says that what the F grade for her school leaves out is:

The fact that our teachers work really hard, and it's so crazy because, in college...I wanna be a dentist, that's pre-dental, I never really gave teachers as much credit as I should have. Like I never realized how much work and time goes into teaching, and I think that the school grade doesn't put in, take into consideration, how hard teachers are working here and how much love there is.

I talk with several teachers at Sandra's school, and they all bring up the efforts and “love” of teachers.

Lakisha works with Sandra. She says that she intentionally transferred to a failing school, to make a difference. She explains:

I do like working with children who are struggling with reading. Necessarily, they just need a little bit more help and assistance and teaching them to be more or better comprehensive readers. So, I felt that this was a great place for me to be, because I have a lot of experience in [teaching reading]. I think that my students would definitely benefit a lot from it.

I ask her about other staff members. She says they are:

Caring, loving, hard-working, dedicated, motivated, willing and eager to learn or to help their students to do better. That's what I've experienced since I've been there, from the very beginning. Teachers who are energetic and excited about helping their children in any way possible that they can to be better readers, better learners, more motivated to learn and just be better, well-rounded students.

School grades aim to measure – in part – the abilities of teachers, but teachers believe their real contributions cannot be measured. Sandra says she is “not sure” how to reward teacher effort. She suggests:

I know there's different things they have, like the Best and Brightest for teachers that are rated highly effective...I don't know if there is a way to do that [grade teachers]. Maybe lessen the load of teachers, hire more teachers so class sizes are smaller. Don't, well like the fact that some of my classes have 23 students in them, 22 students even, it's like if they're a low class I need eight kids in a class...my smaller classes are the ones that if you look at my data, like my smaller classes, my fifth period, 87 percent on the last test whereas my sixth period has 60 percent because there's 22 kids in that class...I think that's something how you could help recognize teachers, by lessening their load, showing them that – okay, you are doing a good job.

For Sandra, large class sizes prevent her from fully employing her skills, which obscures her efforts. So, she thinks the best way for the state to recognize and understand her efforts, is to change policy.

Others mention an obfuscation of student efforts. Jennifer, the principal at Sandra's school, describes the school as “a diamond in the rough” because “preconceived notions people have about [our school] are not the same as what's here.” I ask what the difference is.

I think people don't think that we're achieving and we have lots of kids here who do achieve. But we have some kids here who don't, so it's a kind of even mix. Or that our kids are misbehaving and actually less than five percent of the population misbehaves.

Her co-worker, Joy believes part of that public misconception is due to limited measures of student success. If all the public hears about are FSA scores and police reports, then people remain blind to other forms of progress. She explains:

These kids, they're strong in what they're strong in. And sometimes, multiple choice, bubbling the answers, is just not my strong suit. So that does not mean that I'm not, I can't do well, I can't understand what my teacher is saying. That just means that my anxiety kicks in when I'm sitting here. And if that's the only thing that you're gonna look at to write whether my child is performing or not, or my student, per se. I call them my children because I get so close to them. Whether they're performing at a high achieving rate or not, then that's not fair to them.

Joy says that by ignoring her students' talents, by focusing on "only" tests, the state discourages students. She continues:

Fortunately, I had a parent that, you know my situation wasn't so terrible. My mom did go to college so I had that one person to always reflect. But they [my students] don't have that. I may be the only person they know to look like them, relate to them, that actually has a [college] degree. They may not have that so it's like – with them not making that connection [and] then all they're being graded on is a test, that shoots them down a lot. They feel like, "Well, you know..."

Joy thinks that, school grades, by only recognizing one measure of ability, ignore the efforts of teachers and students. Teachers feel invisible and misunderstood. They feel unrewarded. Some believe that students feel discouraged. Elsewhere in my dissertation I discuss standardized testing in more detail; here I most want to emphasize that teachers feel hidden behind scores, and by consequence – behind school failure.

Different Schools, Different Challenges: Teacher Advocacy of Differentiated School Grades

If school grades hide student and teacher progress, they may also – at a more holistic level – misdiagnose a school’s true challenges. This misdiagnosis may in turn prompt poor policy remedies. This idea is well-articulated by Martha who is 40 years old and in her seventh year of teaching. Martha teaches middle school civics and social studies at a failing school, but for the prior six years she taught in a different county, at an A school. She explains some profound differences between her past experiences and the experiences of her students. For one, Martha identifies as white; 100 percent of her students are black.

Martha is not from the school’s community, but many of her students have never left it – so their exposure to the world is limited. She tells me a startling story to this effect. When she saw her new classroom at the start of the year, she assumed it had been a literature classroom. A fictional map adorned the back wall – a fantasy map – from *Lord of the Rings* or *Game of Thrones*. Students asked Martha to point out their school on that map. She thought they must be joking with her, but soon realized the question was sincere. She says that some students had never seen a map of the real world or were not taught to distinguish real geography from fiction. She knew that she would face new challenges moving to an F school but did not realize the extent of students’ isolation from the world. She feels frustrated that their isolation, combined with other factors, is not considered in the state’s grading formula. Martha goes on to discuss those other factors:

I don't think it [the school’s grade] truly reflects what the children can do. It’s just on that day, that's what they did, but we don't know what happened the night before.

Maybe their mom got beat up by their boyfriend. Maybe their grandmother died. Maybe they didn't get breakfast because they didn't have any food.

Maybe they broke up with their boyfriend or their girlfriend or maybe they suffer from depression. None of those things are taken in to effect when they do a school grade.

I ask about the overall school.

I feel like [our school] absolutely can be above an F, but I think it's because (a) they're in an old building; (b) the lack of technology. Some of these kids don't even properly know how to use a computer. They don't have the exposure to outside resources. In order to even excel...there's no way for them to excel if they're not exposed to certain environments. You know, as far as field trips, those are pretty much non-existent. Even just getting new equipment, getting new library. Anything like that, it's non-existent.

Martha and others pair student isolation with other challenges. I address these challenges elsewhere and other researchers have extensively explored the impacts of nutrition, housing, safety, discipline, and a host of other factors (Amrein & Berliner, 2002; Anyon et al., 2014; Berends & Koretz, 1995; Borman & Dorn, 2007; Conley, 2001; Ferguson, 2008; Figlio & Lucas, 2004; Gosa & Alexander, 2007; Grodsky, Warren, & Felts, 2008; Jacob, 2004; Johnson, Kraft, & Papay, 2012; Murnane, 2007; Parke & Kanyongo, 2012; Porfeli, Chuang, Audette, McColl, & Algozzine, 2009; Rossman & Rallis, 2003; Schwartz, McCabe, Ellen, & Chellman, 2010). Teachers tell me that grades cannot be accurate if they ignore all of these – often external – challenges faced by students and schools.

Teachers do not use these external factors as excuses, rather they see the missions of teachers at struggling schools as fundamentally different from the missions of teachers at

more affluent schools. David – who teaches at an A-graded high school – clearly articulates this idea:

I teach at an A school but I think that if I was at an F school, or a D school, my job would be totally different. Here I can come to work and put the desks in a circle and say, “Hey, today we’re going to do a Socratic seminar so what do you guys think about this?” My job is to ask hard questions and keep them productive. But at an F school maybe the kids don’t even know what a Socratic seminar is. I guess what I’m getting at is, it doesn’t make sense to evaluate me the same way you evaluate those teachers [at F schools]. And I don’t need school grades to tell me that. I’m relatively new at this but I can walk into any classroom and immediately have some sense of what those kids need.

Several teachers tell me that if the state understood the difference David talks about, the separate missions of A and F schools, then policymakers could better support teachers. Jane – who experienced the transition to school grading in 1999 – speaks to this:

I think what pervaded then [in 1999], and often pervades now, is that they’re spending millions of dollars on what we already know: that we have a population of kids that are lagging behind because of social problems that go much deeper and beyond the scope of the schools and the classroom they’re in. And that there would be a much better use of those funds than to just have tests. That was the beginning of when funding was diverted as a result [of grades].

This funding diversion is mentioned by six of the teachers I speak with. They all emphasize that the money from testing and monitoring and publishing grades, would be better spent elsewhere.

Jane goes on to explain that not only can grades not capture the “social problems that go much deeper” but that using the same grading scale for all schools “precludes the state from recognizing differences in learning opportunities.” She continues:

[Schools] just want to finish the race first. They don't consider the true needs of the students and [the state] just hangs the carrot in the same spot for everyone. The state is concerned with economics. The state is concerned with attracting business, with being viable. It then just sets unachievable expectations and doesn't necessarily distribute the support needed to achieve those expectations. Can they really be achieved by schools?

Jane's question – can the state's goals really be achieved by schools – will guide much of the following chapters. Jane starts to explain why she believes that the answer is “no”:

Because the needs are greater of the population ultimately. It could be building needs too, but the needs are greater. There needs to be more support in place. You get a kid who's raised in a household of safety and nurturing and education as a true value, not one that's just wished for, the level of vocabulary is high, they're read to from the time they're infants or toddlers. You could give them a book and an assignment and they could learn.

Martha, Jennifer, David, and Jane all suggest that A and F schools face entirely different challenges – challenges ignored by the grading system. Maybe schools cannot address all of the social problems kids face, but these teachers believe that by ignoring the problems altogether, the state hinders education and creates unfair expectations. Martha, for example, suggests that her school could be an A if students had access to technology and travel to expose them to more of the world. David's students already have those things.

Final Thoughts: Chapter Four

Can a policy succeed when those charged with implementing that policy do not support it? Teachers I speak with support evaluation but not the school grading system as-is.

Some of them do not understand it. Some think it changes too quickly to understand, particularly for parents and students. Grades are meant to communicate a school's quality to parents, but teachers and the media suggest that the interpretation of grades by parents is more complex. Apparently, some are hostile to grades; others do not care. Some use grades as justification for distrust of the school. Teachers think that parents need to understand how grades are calculated, so that they can judge for themselves the holistic value of the school. Otherwise, their children might miss out on valuable experiences.

Confusion and shifting policies lead some to believe that the true intent of state education policy is to wage a war against public schools. Many teachers believe that politicians, especially Republicans, want their schools to continue to fail so that there is justification for closing them. They do not see charter and private scholarships as providing families with choice and opportunity. Rather, they see these policies as willful attempts to undermine public education. In my experience as a teacher I have heard different motivations attributed to these anti-public school lawmakers, including hostility towards the education association, suspicion of all public institutions, and attempts at further racial and class segregation. This fear is only one example of a poor relationship between teachers and the state.

Earlier in this chapter, I also ask: what constitutes a good school? Teachers say that their efforts and the efforts of students matter. They also say that student progress is important but must be understood in context: students start a school year with different education levels, learning opportunities, and spoken languages. School grades – teachers think – ignore some strengths of schools not measured on standardized tests, as well as

student progress that does not fit into universal time-tables for growth. When grades do not accurately portray strengths and challenges, teachers worry that the state will not craft helpful policy solutions.

The number of lawsuits and public campaigns waged by the Florida Education Association suggests that teachers do not feel that – hostile or not – policymakers or politicians listen to them. The wisdom of school grades has not been among the various education debates of state policy-makers. Nevertheless, grades represent a significant policy question. Jacobsen, Snyder, and Saultz (2014) compare school grades to a range of other evaluation metrics used by other states: performance index ratings, rankings, and the percent of students meeting goals. They find that “the format of the data may play a significant role in shaping perceptions of school quality when schools are either excelling or struggling” (p. 14). In their study, people randomly assigned to see grades, rather than a different metric, perceived a larger discrepancy in school quality between strong and weak schools. The researchers conclude that “people see greater differences between letter grades than they do between words such as ‘advanced,’ ‘proficient,’ and ‘basic’” (p. 18).

In 2001, Florida’s policies set a model for the Bush Administration’s NCLB reforms and they continue to influence the policy reforms of other states (Borman & Dorn, 2007; Herrington, 2005; Jones, Jones, & Hargrove, 2007; Ujifusa, 2012). Conversely, other states can illustrate the importance of school grades in policy success or failure. Louisiana, for example, has a voucher program tied to school grades that in 2016 *the New Orleans Times Picayune* called, “devastating” when students performed worse after leaving public schools (Dreilinger, 2016). Arizona’s school grading policy became so confusing in 2017 that the

president of the State Board of Education announced that the grades were flawed and needed revision, catching by surprise other board members, county superintendents, and the Arizona Education Association. The Association's president said that parents could learn more about hotels on hotels.com than they could learn about schools via school grades (Stephenson, 2017). In 2007, New York City Mayor Michael Bloomberg helped promote record increases in scores and school grades. The city's schools won the Broad Prize and numerous publications heralded the achievement. Soon thereafter, to the embarrassment of schools, the National Assessment of Educational Progress released its scores for New York City, which showed that – in reality – students showed no gains on national tests (Rossman & Rallis, 2003).

CHAPTER 5

TEACHER RECOMMENDATIONS: IDEAS FOR IMPROVING SCHOOL ACCOUNTABILITY

Some teachers describe feeling like untapped resources for policy-makers. In Chapter Four, I explore ways that teachers think school grades mischaracterize school quality. Building on these beliefs, teachers offer ideas and guiding principles that they believe must accompany an effective evaluation system. In this chapter, I aim to summarize some of these alternatives, which coalesce into three themes: (1) grades must account for environmental factors; (2) the state must differentiate school grades in the same way teachers differentiate instruction – to account for unique student needs; and (3) grades must be understandable by and useful for teachers, students, and families. Perhaps most of all, teachers want to feel that the state considers them professionals and values their input.

Teachers feel profoundly disrespected in a variety of ways, from wasted time in faculty meetings that could have been e-mails to being cut out of policy discussions. When I began teaching in 2009, our district had just adopted as its guiding philosophy, the work of Robert J. Marzano, an education researcher known for his ideas on school reform. For the

following three years, I do not remember a faculty meeting at which Marzano was not mentioned. Yet, teachers took great umbrage to the fact that Marzano had never taught in a classroom or worked in a public school. This same grievance permeated many of my interviews with teachers: why is our district listening so closely to people who have never taught and do not know our students?

Former Assistant Secretary of Education, Diane Ravitch (2010) makes a similar observation about the 2007 New York City school reforms and connects the ascension of “non-educators” with standardization and corporate-style accountability, writing: “No school-level supervision was needed, because the Department of Education intended to judge every school solely by its results—that is, whether it raised test scores...this tactic [or replacing job titles with corporate-like positions] enabled people to hold jobs without meeting state certification requirements. Most high-level officials at the DOE were non-educators” (p. 80). I could devote an entire chapter to Marzano’s influence on public education, but his importance here rests in what he represents for teachers. Teacher and writer Paul Murphy, on his website *Teacher Habits*, writes that Marzano “has been an outsized part of my professional life for longer than he deserves” and that “for most teachers” his “name is mud.” Most emblematic, Murphy says, “we’re really sick and tired of being told we’re failures by people [like Marzano] who don’t have the courage to do what we do.” (Murphy, 2018).

Teachers see the three themes of this chapter as common-sense realities about schools: schools are different, with different students, in different neighborhoods. In teacher trainings and faculty meetings, teachers report being told to differentiate instruction and that

every child is unique with her own learning style. Why then does the state ignore those truths when grading schools? Teachers suspect the answer lies with those like Marzano who purport to layer teachers and schools with accountability without having “the courage to do what we do.” Teachers sound surprised in some of my interviews that anyone – even a researcher and former teacher – cares about their opinion on the issue of school grades. No one has asked them before, even though it determines their evaluation and employment.

Teacher Arguments for the Inclusion of Environmental Factors in School Grades

Teachers ubiquitously tell me that an effective school grading policy must account for unique environmental factors. These factors include quantifiable things: the percentage of English Language Learners (ELLs), the school’s racial and income demographics, the experience levels of teachers; the age of a school’s building and the safety of its neighborhood, among many others. Some other variables – a culture of success, family legacies, and school rivalries, among others – are difficult to measure.

Teachers do not use these factors to excuse poor student performance. To the contrary, teachers like Jane condemn administrators who they think cynically use demographics to excuse school grades. In 2017, Jane’s school dropped a letter grade. I ask her how her administration addressed that drop at their next faculty meeting. She explains:

[The school grade] was surprisingly absent from the dialogue. It was justified with comparisons to other schools that had also dropped. It was the typical thing that we do, which is when we’re an A, we’re amazing, and when we’re not an A, there are other things that need to be considered as to why we’re not an A. To me that reeks of

something that's not valid, a system that's not valid. When – “Yay! We're being successful!” and – “Oh, we dropped, well there are reasons why.”

Jane thinks that her school only considers environmental factors when the school is struggling; she wishes the school grade always communicated those factors. This would – in her eyes – increase the grade's validity.

For some teachers and parents, this concern over accuracy is not merely about making grades “valid.” Some feel that if grades do not recognize the unique needs of schools, then policymakers will ignore those needs. Angela, a parent of state-scholarship recipients, says she sends her kids to a private school because the public school's C grade is unacceptable to her. In addition to parenting, she works as a service coordinator and case manager for low income families. Angela describes her career this way:

I assist low income residents with services to help them to become self-sufficient. So whether it be going back to school, getting their high school diploma or GED or going to secondary school or vocational school, or wanting to become a first time home owner, or trying to get a job, or just basically trying to get their life together, get them off drugs, alcohol, whatever it is. Then I refer them to services that are able to help them.”

I ask her what the public school would need to change to make it more acceptable to her. Her answer – based partly on her experiences with struggling families – does not concern academics so much as factors not directly measured by school grades.

I think honestly if they want to make a change in those schools, certain programs have to be in place for those changes to be made. So, if they're having a lot of kids with behavioral problems, then there should be a program, a successful program with therapists who are certified, licensed, whatever, who know what they're doing to be in those schools and to understand those kids.

The thing is, the kids who actually go to those schools, they have lots of problems. The environment that they live in is not the best. Just horrible. Just from me dealing with these families, the environment is horrible. It's mostly single parent families, which I'm a single parent, but I came from a household with two parents.

The thing is with most of these single families, in these neighborhoods, they come from broken families and they just basically follow the same cycle and there's a lot of drug use, there's a lot of abuse. A lot of these mothers are really young because, just from experience talking to some of them, a lot of them have been sexually molested when they were younger, so they become promiscuous. That's just some, and others just didn't have a role model to follow.

Angela is not aware of any standardized metric or signal to parents that measures school efforts to address behavioral problems or poverty, even though these are things she cares deeply about.

Joy is a black woman in her fifth year of teaching. As a college student, she did not plan on teaching. She originally wanted to work as a clinical psychologist but it required more time than her "responsibilities would allow." She does not regret her choice since teaching helps her in "understanding the human body at an adolescent age" as well as with her own parenting. Joy discusses some of the same concerns as Angela, from the perspective of a teacher. She says that when dealing with struggling students – 100 percent of whom are black – she has many "fearful moments." I ask what she means.

Sometimes you do get children that deal with things that you may have never had to deal with as a child. So, the anger that it brings them can kind of...you know you want to deal with them accordingly. You never want to say the wrong thing because you don't want to trigger a child. [You don't want] them to get home and do the wrong thing to themselves or to others. So, when you hear red flags, those warnings...because we have a lot of children that are so emotionally distraught that you don't want to [trigger] them.

Even though Joy identifies with her students racially and culturally, her lack of experience with trauma makes her cautious about saying “the wrong thing.” She feels that – even with her studies in psychology – she lacks the training to properly respond to the needs of some students.

Jane remembers feeling a similar struggle when she taught at a poorer-performing school. For the future, she believes:

We really need to be getting way more in the area of psychology and counseling than we are. We’re expected to be psychologists and counselors and parents with no adequate training. Now one of the elements of our [student] grading system is to what extent we ignore distractions and are unaffected by student outbursts and inappropriate behavior in the classroom.

I ask what kind of training she suggests. She says:

How to walk the line between not reacting to inappropriate behavior and being aware of potential violence and danger in the classroom. All these areas that one video or one after school training is not adequate to support. It has to be a bigger part of our teacher preparation, like the way the CIA trains people to be in the field, it needs to train teachers to be in the field. We’re in the field.

Like Joy and Angela, Jane believes that issues of psychology and counseling are essential to effective schooling, yet they go ignored by school grades and teacher preparation.

“A Sociological Epidemic”: Connections Between Demographic Factors and Academic Performance

Jane’s co-worker, Theresa, also discusses unique needs of schools and students, with a focus on external circumstances. She sounds angry describing what happens when those needs are “hidden” or obscured by a school’s grade.

If you were to take the entire faculty and staff of [a local A school] and put them in [a local F school] and put the entire faculty and staff of [the F school] at [the A school], then [the A school] might drop to a B but [the F school] is not going to change. I know that the teacher has a lot of responsibility about the success of students in his or her classroom but when a kid has seven classes a day and a myriad of other problems, we’re addressing the wrong issues. This is a sociological epidemic. School is the last great democratic institution that we have in our country, where we have to take everybody. We have a child in our school this year who weighs about 70 pounds, about this big [table height]. He’s nonverbal. He wears a diaper. His IQ is so low that it’s not registered, but he’s enrolled in Algebra 1 because state law says he has to be in Algebra 1. There’s an insanity, a disconnect between people making policy, people enacting policy, perceiving policy. There’s a big disconnect.¹⁵

I ask Theresa if getting an F could be a useful signal to the state that a school needs help with challenges like special-education students. She responds:

Getting an F, sure. The first year that you have an F, they come swooping in and they bring doughnuts and cupcakes and money, and they’re going to send you all to professional development. It’s all great. The second those letter grades come out though, the firings begin.

Theresa stresses that these “firings” are not a fair response because teachers cannot respond to all of the problems children face, most of which stem from their neighborhoods. She says

¹⁵ I have removed Theresa’s references to particular schools. I understand her references as examples of schools in general, rather than thoughts about unique circumstances.

that parents have “always known what neighborhoods were what neighborhoods” and that “school grades have been far more detrimental than helpful” partly because the consequences for failing schools do not address the “sociological epidemic” plaguing families.

Martha also discusses the struggle for teachers to respond to student needs. She works at an F-graded middle school with “100 percent African American students.” She is new to the school but previously taught “predominantly white children for six years.” She felt that she “needed to be where kids who had no idea and hadn’t seen stuff about civics, history, [and] things like that [lived].” Despite feeling a “calling,” Martha thinks the school’s socio-economic realities – “reduced lunch, Title One, and the opposite races” – are challenging. In Chapter Four, for example, I discuss Martha’s frustration with the relative isolation of her students. She says that students “don’t have the exposure to outside resources in order to excel.” These resources include technology, library books, and field trips. Martha says these are “non-existent.”

Martha’s principal – Jennifer – also talks about the need for exposure. She compares the relative experiences of her students versus her own children and children at more affluent schools.

My kids [students] have probably never been out of our neighborhood or some of them have maybe never been to Disney World. Whereas, in a more affluent school, kids are exposed to things. Kids hear things. Kids who are verbal, they talk over dinner. So, their words and their vocabulary – I have to give to them. I don’t have to give [them] to my own child because he’s always heard those words. Those are the things that aren’t accounted for and should be accounted for.

Martha and Jennifer's school will soon be moving to a new building, with new technology. Martha thinks this will be "revolutionary" for students. I ask if the F grade might have helped spur the county to modernize the school. She says, "yes," but feels the help is offset by the harm of having an imposed "scripted" curriculum that "frustrates" both her and the students.

Martha and others see some benefit to schools when their counties send additional resources, such as new technology. In this respect, perhaps a failing grade helps draw much-needed help. However, these same teachers dread negative attention – scripted curriculum and staff layoffs – that they feel is counterproductive. They would rather receive credit for the special work that they do, like working harder to expose isolated kids to different technology and experiences. Others think innovative interventions like psychology training are needed, but not provided. Even more teachers feel that nothing the county does will be enough to change schools' failing reputations. This feeling of futility is inspired partly by demographic differences that teachers want grades to take account of.

Several teachers tell me that, even without an official acknowledgment, school grades say more about neighborhood demographics – and race in particular – than about schools. Theresa says teachers at her A school call campus a "country club." She says: "I went from [an F] school that was overwhelmingly food insecure, free and reduced lunch status, Title 1...I went from that school overnight to a school where kids were handing me bottles of Calvin Klein cologne because they could." She continues:

I think the rating system creates a have-versus-have-not mentality, because it's definitely tied to neighborhood. [Local F schools] are never going to be high-rated schools. They're not. That sounds racist and horrible, but if you're trying to pay your 200 dollar rent for your house and then you're sending a kid who's not ready to learn

to school, sheer will and determination can do a lot, but a poor neighborhood feeds into a poor school. [A local F middle school] is a school of absolute blight and despair, literally the poorest of the poor. It feeds a school district that is probably the single most abjectly poverty-stricken in [the area]. It's 100 percent black.¹⁶

Jane agrees. She says that grades are “more accurate in terms of social stratification, socioeconomics...you've got this population that has this amount of success indicators and then you have this population that has this degree of at-risk factors.” She thinks that these risk factors cannot be fully addressed by schools. She explains:

[The state] just sets unachievable expectations and doesn't necessarily distribute the support needed to achieve those expectations. Can they really be achieved by schools? That's the problem too, we're dealing with social problems that go way beyond the schools. We do our part, we feed them, we sometimes even clothe them. We try to nurture them the eight hours we have them but we can't take the place of a safe, loving, nurturing environment from infancy. We can't do that. We can't follow them home and protect them.

Jane and others see these efforts to “nurture” students and work with the community as key components of a quality school. Yet, since school grades do not account for demographic factors, they cannot acknowledge schools' demographic supports.

Theresa says this happened at her previous, F graded school. She says that before 1999, it was “known as a rougher school, but it was also known as a very supportive school community and a sort of vanguard of the community [with] social services, parents came whenever they needed to come, we had meetings that were open until 9 PM.” However, once “the letter grade system kicked in and real estate agencies really started touting the school

¹⁶ I have removed obscured references to particular schools and to the city.

letter grades, the last few families that were pro-education fled.” She says that grades ignore that “institutionalized poverty is the real problem in this country and you can call a school whatever you want to call it but if you have poor children who haven’t been fed, they’re not going to learn.”

Alex, who previously served as principal at a failing middle school, says this applies to his school. When I ask whether the F grade is accurate, he responds that the school is “in a unique situation.” He continues:

Historically, the neighborhood, well there are just some challenges that exist within the neighborhood, some cultural challenges that exist, some socioeconomic challenges. But the school itself is almost like a hub for the community and we work within the community and have to deal with some community challenges that exist at the school. But overall, it’s a place where the parents want their children to be successful.

He tells me that teachers have trouble reconciling the school’s role as a “hub for the community” with its failing grade. They want the grade to instead, in some way, reflect that community role.

Cultures of Success: Recognizing Intangible Performance Factors

Alex and others tell me that socioeconomic challenges are more insurmountable when parents and their children do not understand what success means. He explains:

The children want to be successful, but what I found was, they didn’t know how to actually identify what that means. When you say, “do [you] want to be successful?” to children, they say, “yeah, I wanna be successful.” “Well, what do you want to do?”

“Oh, I don’t know. I don’t know what I want to do. I don’t know how I’m gonna get there.” They just have a vision of – I’m going to be successful – without a clear plan.

So they really look to the school to kind of frame what that means and what it looks like, which is a challenge. When you look especially at inner city schools, often times you don’t get all of the manpower and resources that you would get in some other school situations. Or often times you get first year teachers with less experience that weren’t picked up at other schools, that end up working the more fragile schools. Unfortunately, that happens sometimes.

Alex explicitly links student ignorance of what success “means” with the challenges associated with new teachers and less resources. This link is reflected in my interviews with teachers at his school, who say they find it difficult to teach children what it takes to achieve success. I ask Martha, for example, if her students have a clear idea of what success looks like. She responds:

I don’t think so. I don’t think they know what success is. They see stuff on TV, they say, “Oh, I can be a doctor,” but I don’t know that they truly know what it takes to be a doctor because they’ve not been exposed because many of their parents aren’t college educated. So, if parents aren’t college educated, then they may not know about college, which is something that I try to bring in.

I ask her how she communicates that to students. She says: “[I ask], ‘Hey do you want to be the person who drives up in the Mercedes and gets the gas, or do you want to be the one pumping gas? So, to pull up to the gas station you’re going to [have to] get that through an education.’”

Martha’s co-worker Rhonda is a geometry teacher of eight years who also does family and community engagement work for her school. She describes this challenge of teaching success as consistently difficult to overcome. She explains:

[In] any poverty-stricken neighborhood, the person who has the money often times is the drug dealer. So, if I see him with a nice car, and him with the girls, with money or whatever, that's who I aspire to be. If I don't have a doctor, or a lawyer, or whomever that comes to either talk at my school or that's in my family or who's trying to mentor me, what am I aspiring to? I have no idea about who I am and a lot of times we look at the parents and say: "Why aren't they teaching them this?" How is someone going to lead someone somewhere that they don't even know how to get to? So, that's the problem that I think we run into on a consistent basis is how do we teach our students to excel above what they see?

Rhonda and Martha's new principal, Jennifer says: "most of our students want to do well" and "less than ten percent" cause trouble. I ask her what differentiates her students from those at an A school. She responds:

Because of their home environments, where they're coming from, the type of life that they are exposed to on a day-to-day basis, not being able to see success around them. You know, having relatives or family members who are not as successful and what they experience every day – in my opinion – is a direct result of what they feel like they can accomplish. So, they have very little expectations coming from the home environment and I think that's the biggest challenge that we face.

As I discuss in Chapter Four, teachers at different schools see their missions differently. Based on my own experiences and my interviews with teachers at A schools, very few – if any – teachers say that the biggest challenge facing highly-graded schools is "very little expectations coming from the home environment." Teachers believe that these different challenges call for acknowledgment and for different measurements of success. Only by admitting differences and measuring them can policymakers figure out how to respond.

The Difficulty of Attracting Teachers to Failing Schools

Rhonda and Martha describe a need for successful role models. Unfortunately, teachers and principals report that successful role models can be hard to find. Alex says that socioeconomic challenges make raising the school's grade unrealistic, which in turn makes attracting quality teachers difficult. He explains:

Teachers who are more experienced, who've been around other schools, kind of see how things shake out. I've been in suburban school settings. I've been in inner city school settings. When I was in a suburban setting, we were an A school. When I moved into the inner city, you're a C, D, F kind of school. So, it's almost like, that's just how things are. And it's unfortunate, because people know. We have job fairs as a school district where they put out...we as principals say, "this is how many vacancies I have." They put out a bulletin and get people from all the universities to come in and it's a huge job fair. But when I'm sitting at an F school table, and I need a reading teacher, and the table next to me is a suburban, more affluent population type of environment, and they need a reading teacher, well where you gonna go? And it's almost not fair, because we're fighting for the same people.

I ask Jennifer if she agrees that the school faces unique challenges because of the neighborhood. She says: "No, I don't. I don't because I think there's poverty everywhere. I think that there's apathy of parents everywhere. I think that all of the social ills that we see here at [our school], they're everywhere." Yet, she does acknowledge that at a failing school, hiring can be difficult. She started after most of the year's staff had been hired, but I ask if she anticipates having difficulty in the future. She responds:

Absolutely, absolutely. I've only done two [interviews] since I've been here in the three months [but] there's very little interest, or when I sit down, we start to talk and then in the end I say [what school I work at] and they say: "Oh, you know, that's not really where I wanted to start." And then I say: "I'm telling you that this is going to give you what you need to be able to go anywhere you want to go, but if you give us a chance and you come here, I promise you 'll feel like you're changing the world

and you'll want to stay." But you have to get people to think they want to come here first.¹⁷

Jennifer says that she would have no problem hiring staff at an A-graded school, but at her current school she may have to rely on first year teachers. If so, she wants to hire those with the most promise. Her strategy is to convince them that they will gain valuable experiences for future jobs and feel like they are "changing the world."

Alex agrees, but says that even hiring the best new teachers can be difficult. He explains:

It's hard to staff an inner city failing school. It's almost a – just get what you can get. When all of those positions aren't filled, the district is on your back saying: "You have reading vacancies. You have math vacancies. Just fill them. Just fill them. Here take this candidate. Take that candidate." And that pressure, it almost turns into – let's just get somebody with a certification and do our best with them.

Alex stresses that new teachers can exacerbate existing problems for a school. He explains:

The danger in [just filling vacancies] is that you're pulling in the teachers who don't have experience teaching, teachers who don't have experience in a high-pressure situation, that you're pulling in and saying, "Go teach those kids." And the bottom line for the district is, [they] don't want to have any classrooms with subs in it. [They] want full-time, certified teachers.

But the problem that it creates for the school is [when] your boss comes to the school and walks in a classroom and is like, "What's going on in there?" It's like: "We're working with them. We're trying. We're coaching them." And you actually end up pulling your internal resources to support those teachers in those classrooms. You're pulling your reading coach or your math coach to go triage in a room with a teacher who's struggling. Which means now that coach can't support other teachers.

¹⁷ I have removed Jennifer's reference to her school's name.

According to Alex, all new teachers struggle; the problem for a failing school is the number of them:

[Last year], I had 53 teachers total. 28 of them, it was their first year [at the school]. 22 of them it was their first year teaching, ever. When you have 22 first year teachers on one campus, in a struggling school, in the inner city, it is a challenge to be able to do the kinds of things that a school should be able to do in a year. Just because [new teachers] don't know. [New] teachers don't know classroom management, curriculum standards, all of those things. That's a challenge that is created internally, because of a school grade, that we have to be able to work through. Along with that comes discipline, if you have that many first year teachers – I don't care what setting you're in – you're gonna have more discipline issues.

The district tries to respond to Alex's concerns, but he argues that its response is insufficient.

He explains:

The district tries to supplement teachers – give them extra money. In [our] county, it's 2,600 dollars to work with what we call "targeted schools." But the teachers get that 2,600 over the course of a year. Then there's a Title One bonus, which is five dollars a paycheck. So, [teachers] get a little bit more money, but that's not really incentive enough to pull your more experienced teachers.

Funding (In)Equality: Teacher Arguments for Recognizing Differences in School Support

Regardless of staffing and any out-of-school challenges, Florida wants students to receive equal educations. The state's funding formula ensures that schools receive equal base funding proportional to the number of students they serve. Theresa tells me that despite the "have-versus-have-not mentality" of grades, "we are a state that practices equal property tax distribution so theoretically, [each high school] has the same per capita budget," but she quickly adds that, "you know that's not real." She continues:

We have a hyperactive PTA that supplements things like crazy. [Another school] could never have done a [fundraiser] like us. They could never have raised 1.9 million dollars in capital expenses in six months. We did, because of our community. Those extra layers provide a different learning environment. Every kid in our district has a laptop right now, but at [another school], if you lose your laptop, you've lost your laptop. Those kids can't pay for a replacement. Our kids can. There's still an unequal playing field.¹⁸

Her co-worker Jane agrees:

[Funding can't be equal] when you have a school like [our rival] that has a pool and has a restaurant grade kitchen for students to learn in, that has the labs and everything is stocked. There may be a certain portion of funds that come from taxes that are equitably distributed, like for example in buildings. We may all have budgets and so forth, but our buildings might have different demands. Sure, all the kids have computers, but do all the kids have internet? What are those kids having to do after school that some kids who live in better neighborhoods don't have to do after school, that can focus on their education?¹⁹

Like Alex, Jane says that these differences are exacerbated by “defacto staffing as a result of school grades.” I ask her what she means. She says: “you have teachers that for different reasons don't want to go to those D and F schools that are going to be more attracted to the A, B, and even C schools.” One of those reasons is – according to several teachers – the inequality in resources. In our interview, Rhonda tells me that “right now it's pretty cool” but says:

I remember days when the AC was broken and you have kids in class – and the AC's broken – then I may [bring] water bottles [for students] or have...just have to do different stuff. Like, the ceiling tiles, the people are here like every week replacing them, it's all putting patches on different things. I said if the AC repairman, if he got paid by how much he was here, for every work order he had to fill, he would make a

¹⁸ I have removed Theresa's references to specific schools, as well as a specific fundraiser.

¹⁹ I have removed Jane's references to particular schools.

lot of money. He would make a lot of money because they're just patching different things.

Rhonda reiterates several times that rather than permanently fix problems at her failing school, the district "patches" things. Next year her school will have a new building, but she emphasizes that for a long time she doubted that outcome, which made it difficult for her to teach and for students to learn.

Politically there were certain things that were going on. They were thinking about turning this into a charter school, or just shutting the doors and different things like that. But still, when you talk about having the students in a school where [they're] like: "What am I supposed to be proud of? What am I supposed to look forward to?" Like the bathrooms get broken and it's not a matter of the kids always doing something, it's just because it's old. So what pride am I supposed to take in something that's ridiculously old?

Rhonda suggests that when kids cannot take "pride" in something, they feel less incentive to behave well or focus on learning.

Others point out that some schools cannot offer students equal class opportunities.

David – who now works at an A school but previously worked at a smaller, C school – explains:

Here students can take as many AP classes as they want. We have more than most could possibly finish. We really encourage that because it is supposed to show college readiness and that's good for kids, but it also helps with [our] school grade. But, at my [former] school we barely offered any. We were just smaller and we didn't have such a big staff who did so many different things and we didn't have all of the extra funding. So, kids missed out but we could also never get those points, for the grading formula.

For some schools, this creates a circular problem – they cannot afford to offer a wide variety of “college-level” classes, but A-grade bonus money that could help them do just that is out of reach. (Gonzalez, 2012). Moreover, research has found that the bonus money has the effect of re-allocating money away from poor schools and schools of color (Dittmer, 2004).

Summarizing the Inequality Challenge

There is a pronounced feeling among my study participants that policymakers do not understand these links between school grades and external circumstances. I ask Martha – who felt a “calling” to move to a failing school – about Jeb Bush’s argument that it would be unfair to not hold all kids to the same high expectations (Cobb-Roberts and Shircliffe, 2007). She responds:

Jeb Bush went to private school. Jeb Bush never filled out a FAFSA in his life. Jeb Bush’s kids never did. So he’s talking from a perspective he knows nothing about. He doesn’t know about public education. He can’t possibly know about it, unless he’s here in the trenches, seeing it. So [his argument] kind of insults me a little bit because each kid is different. Each teacher is different.

David agrees. He describes the idea that “politicians understand what students need” as “ludicrous” and “absurd.” These teachers agree that grading every school in the same way, regardless of circumstances, results from policymakers’ ignorance, if not outright hostility.

All of the teachers who I speak with address the impacts of demographic factors on school grades. These teachers do not think student challenges fully explain school grades, but they do think grades should capture or account for those challenges, as well as schools’ efforts to address them. Some worry that – by leaving out demographic factors – grades

encourage policymakers to ignore unique needs like trauma counseling. Some tell me that a one-size-fits-all grading system also ignores the differences in what small rural schools and large urban schools have to offer. They think important components of a “good” school, like community engagement, go ignored.

Others believe that a failing grade draws helpful attention to the school, but simultaneously fear interventions they see as counter-productive, like scripted curriculums and staff layoffs. These interventions are seen as unfair because they punish teachers while ignoring the real sociological challenges students face. These teachers think that school efforts to assist students with food, shelter, or counseling services should factor into the grading system. They say that a failing grade makes it even more difficult to attract high quality staff, as do seemingly ignored resource and facility differences. It is unclear how a more inclusive grading system would change this dynamic, but teachers feel certain that their schools do a lot of good and overcome a lot of obstacles that grades do not reflect. Teachers report that when that good is not reflected, teachers and students able to leave do so and their schools decline even further.

Differentiating Grades to Match Differentiated Instruction

Lakisha is a black woman in her sixteenth year of teaching. She teaches eighth grade intensive reading at a failing middle school. Although new to the school, Lakisha says that she has taught “children like the students [here] for a number of years.” When I ask what kind of student she means, she replies, “challenging, coming from low socio-economic

backgrounds, not necessarily as motivated to learn.” Lakisha believes that these students “just need a little bit more help” and that “this was a great place for me to be, because I have a lot of experience.”

One of Lakisha’s goals is to prepare students for the Florida Standards Assessment (FSA) which they must pass to move on academically. I ask her if she believes FSA is an accurate measure of student knowledge. She responds: “I would honestly say yes.” She continues:

I personally feel that standardized testing is necessary because at some point or another – when you leave high school – you’re gonna have some type of standardized test. I do think it is necessary to really know for sure if we have comprehensive readers. If they are reading proficiently enough to be able to receive a high school diploma. Because you know what? The whole “No Child Left Behind” – not that we still completely focus on that – but we want to make sure we have people who are ready to go out into the world and work. College and career readiness, you know, we want to make sure we do it. If you don’t go to college, then you do have...you’re ready for your career enough for you to at least support a family of four. You’re educated enough. So, I do think it’s necessary.

This answer surprises me for a couple of reasons: first, Lakisha is the first – and only – teacher who speaks positively about the FSA or standardized tests; second, Lakisha herself tells me, earlier in the interview, that she does not think her school’s F grade is accurate. I ask her: if the FSA is accurate, but the school grade is not, what else does the grade need to measure?

Lakisha responds:

I’ve always kind of felt like maybe we should do something a little bit more authentic, because the thing about teaching is we teach in the multiple intelligences. We feel like we could test them with STEM programs and different things like that.

So, why not test them in that way? Everything doesn't necessarily always to be a paper-pencil. So, sometimes certain students, or a group of children, maybe they could have a school put together a project or a group to do a project, kind of like the STEM thing. I don't think necessarily all children can perform the same way on a test. But I think it should be a combination of things.

Because we teach them using the multiple intelligences, but then we don't want to test them with the multiple intelligences. So, it should be a combination.

Lakisha is not alone in this belief. A majority of the teachers I speak with – and both principals – talks about “multiple intelligences.” Martha, for example, tells me that “if you have to differentiate instruction in order to teach each kid, you have to differentiate [school] grades.” Teachers discuss the concept in two ways: suggesting that students are evaluated differently and that schools receive grade points for innovative efforts to reach students. These concepts are not unrelated, since standardized tests make up the majority of school grade points.

Lakisha discusses the former concept. So does her colleague, Martha, who tells me:

I think [testing] needs to be differentiated. It needs to be on levels. We've got IB or gifted kids up here. Let's test them on their giftedness and their IB. General education kids, test them on what we expect general education kids to know. The average kid and then the ESE kid – let's test them on that level. It has to be, it has to be on different levels.

It has to be differentiated or there's no true way to get a good evaluation or a good grade about a school because there are so many different learning styles. Some kids don't learn by reading a book and writing something down. They learn visually. They learn by acting it out. They learn by writing it. If Bloom's is a true theory, then that's how we need to test.²⁰

²⁰ Bloom's Taxonomy refers to a six-tier model of learning objectives based on complexity.

Sandra – who teaches seventh grade science – agrees. She tells me that there is a disconnect between how teachers are taught that kids learn – via multiple delivery methods – and how the state wants kids evaluated. I ask her if improving the school grade has become more important than good strategies for teaching students. Sandra replies:

I think so, to be honest. Like [the grade] is just everything, you know what I mean? I remember when I was in – and it wasn't too long [ago] – I was in school, and our teachers taught standards, but didn't teach standards like this...like, only standards of basic instruction. We need to get everything down that could be on this test. I'm like, "Okay well when are they actually going to master them?" They're like, "Oh we're not about mastery, we're about exposure." [Sandra laughs].²¹

Sandra works at a school where she receives scripted lesson plans from the county. She feels incredulous that while the lesson plans outline different learning strategies and mastery goals, her supervisors want her to "teach standards like this," with just enough "exposure" to raise the school's grade.

Other teachers discuss different intelligences at a more macro-level, as they relate to scoring the school. For example, Joy – a science teacher who teaches next door to Sandra – says she wishes that the state paid attention to how the school engages with different groups of students. During our interview, a group of students is working on a science project in another part of the classroom. Joy explains:

The SECME club is meeting now and they're gonna be going to SECME competitions. [The school should be evaluated] on things like that, seeing students with district competitions for different subjects. We don't really get the resources here, but I don't know if [the county] has a model UN or a debate team that would be geared towards history and social studies. I know we have the VEX and robotics and

²¹ Sandra refers to the Sunshine State Standards, which can be found at: www.fldoe.org/academics/standards. Several teachers tell me that they are required to connect each of their lesson plans with the standard they intend to teach.

SECME, we're developing our math team and science so they [should] be looking at that, seeing how we'll perform when we go to these things, the fact that our students are getting involved with those, using that as a way to say, "Okay, well let's see how this specific group of children did." We really need to separate our IB students from our regular students so they know which group of students tier where, like really get that down pat. So, they can look at them separately and not altogether as a whole school.²²

Joy suggests that programs like SECME and Model UN are better measures of student achievement than standardized tests. She wants school grades to consider student participation in programs like these, as well as different groups of students – IB and "regular" – rather than the "whole school." She tells me that she worries the F grade gives the wrong impression of the school and scares off parents whose children might find success in the IB program or in an academic club.

Joy admits that this might be a difficult thing to measure. Her principal Jennifer agrees. She says: "I know that there's no way to say, 'Okay, we'll grade you like this and you like that and you like this.'" However, she says several times that "one size fits all doesn't work." She concludes that the state has to take into account where different groups of students start. She explains:

I've thought about [alternatives to school grades] and I've thought about the performance versus growth. I know we have to figure out a way to say we're doing what we're supposed to do. When we do what we're supposed to do with the set of kids that we have it has to look totally different. And I don't know how to make it where you take us where we are and where we actually move our kids and then you take them where they are and where they actually move their kids.

²² SECME (The Southeastern Consortium for Minorities in Engineering) is "an alliance of K-12 educators, universities and industry/government partners committed to engaging and preparing more minorities and girls for college and careers in STEM." For more information, see: secmeinc.wixsite.com/secme.

Unlike some teachers, Jennifer is not looking for an alternative to standardized testing; however, she thinks that grades have to give the school credit when “we’re doing what we’re supposed to do” and “we actually move our kids.”

“A Mechanical System”: Reducing Students to Points in the Grading Formula

Jennifer and others emphasize that they want to move all kids, not just some. Some teachers worry that the grading system encourages schools to focus on points instead of the students they represent. Jane speaks clearly about this. She says that the difference between working at a C school and a B school “is an impossible thing to communicate, even with school grades.” In trying to peg that difference, she says: “What I think starts to happen [at a C school] is that you start to detach from the humanity and it starts to be about competition and numbers and data rather than individuals and children and guidance and modeling.”

Alex – a principal formerly at a failing school – also laments this focus on “numbers.” He explains:

I don’t think it [the grading formula] accurately measures learning. Because, just because one kid, maybe they needed 16 points to get to the next band to count as a learning gain, well maybe they got 14 points. Well, they don’t count it as a learning gain, but you might have another kid who needs one point and got two, so that kid went up two. That’s a learning gain. This [other] kid went up 14 but it wasn’t enough and it wasn’t, doesn’t count.

I think they wanted something that was fair for all of the schools, an equitable system, but I think a lot falls through the cracks with measuring that achievement. For a middle school, it’s 900 points. So, 300 of your points come from reading, 300 from math. Then it’s 100 for civics, 100 for 8th grade science, and 100 for acceleration. So,

when you look at all those pots of kids, it almost sets up a system where we as schools start looking at kids and saying: “Well, how many points is this kid worth? Okay. For reading, is this kid gonna count for us or against us?” And it puts more pressure on the school, I think. Especially as a principal, because you start thinking in terms of numbers. And, we sometimes begin to lose sight of children and their learning and what we’re trying to do for them.

[It has] become more of a mechanical system where it’s a grab bag for points and where can you get your points? And how can we get as many points as we can, because we need to get able to sustain a certain level?

Alex adds that schools “do a good job of making it about kids” but says of the point calculation: “you can get lost in it because there’s so much pressure coming from external forces.”

Educational Triage

Booher-Jennings (2005) describes these practices as “educational triage” – a focus on students who might be on the “bubble” or close to passing, as opposed to those students at the top or bottom levels of performance. Sandra, a teacher at Alex’s school, gives a specific example of such a policy. She believes that an F grade accurately describes her students’ performances, but she holds mixed opinions about the resulting policies. She explains:

One of the initiatives now is targeting 18 students. Each teacher gets assigned 18 students to target to push them and show them that they’re FSA score can move by like...well they’re in a perfect spot where they can get moved by like seven or eight points. It’s just those seven or eight points that will change a school grade. So, I think that’s very unique to an F school, like at another school you might not just target 18, you’re probably told, “Okay, we’re gonna try to get all of our kids to move.” But, here it’s like specifically 10 points.

I ask if focusing on 18 students has created any problems. Sandra says that she worries about the kids who are not directly targeted.

You might ignore...well they told us, “You’re not leaving out the other kids, you’re just putting a special focus on them.” And then it’s like – okay, well if I’m putting my effort into those kids, what about my kids that are really, really far behind? Am I not extra focusing on them? Am I only targeting these 18? I don’t think it was meant to be like that, but as a person, that’s how I took it. I was like – that’s kind of interesting that you would do that. We had a whole meeting about it, and we had to write down the names of 18 kids that we are choosing to focus on. I have a list and I had to pull all their data for them and I had to look at what are their scores on my past tests and I think that’s interesting.

Sandra seems torn between her support for giving students extra support and her concern that her school is not focusing on getting “all of our kids to move,” especially those “really, really far behind.” Later in our interview she says that “moving the school grade has become more important than actually educating kids.”

Sandra’s co-worker, Lakisha – a 16-year veteran teacher – shares her concerns. I ask Lakisha if the emphasis on students who “are on the bubble” pushes teachers to leave out high and low-performing students. She answers:

Yes, definitely. Because it can be rather difficult to move a student at the top, because they are already at the top. So moving those students would require a lot of work. Most often, hopefully they’ll maintain where they are. But the students who are at the very, very, very bottom – the lowest 25th percentile students – we do want to make sure that we pay close attention to them as well but we definitely have to make sure we direct more direction to those students who are in the grayer areas. But it could be a huge concern later on if we don’t still focus on those [lowest percentile] students as well.

Like Sandra, Lakisha seems genuinely conflicted – committed to focusing on students “in the grayer areas” but concerned about not giving enough attention to other kids.

Teacher Calls to Make Grades More Constructive

None of the teachers I speak with believes that school grades are accurate, even though some trust the underlying data. They do not see this as a contradiction, but instead as a sign that the current grade indicators are wrong. In each interview I ask: if grades are not useful, what could the state or county provide to you or to parents that would be?

The majority tell me that they do not need any data from the state. They say that no one can know better than they can how their students are doing. Sarah – a 19-year veteran teacher – feels doubly upset that administrators presume to know better and impose on her planning time. She explains:

What they’re doing now is pulling yet another planning period, so that we can sit there and [they can] tell us what teachers are supposed to do anyway. I mean, if you're worth your salt as a teacher, you know where your kids are, whether that's from grading papers, or observation. I don't need a number from a score, for a test that they took in the middle of the year to see how my kids are doing

Teachers argue that instead of sending data, or interrupting planning time to deliver it, the state should listen to them and defer to their assessments. When I ask how to inform parents, several of my participants lament a loss of community functions. For example, Sarah – who now teaches in a small, historic town – tells me:

I saw this a lot [when] I used to live in South Florida and a lot of the schools were community schools where they had programs and different things that...you know, some adult learning opportunities and things like that. So teachers, parents, the community could [all] be there and participate. Some of the teachers would teach those courses. So, people got to know each other. That's what I would like to see more of: a daily commercial, [instead of a grade] from someone who has never stepped foot on the campus.

For Sarah, community classes and a “daily commercial” seem more useful than a letter grade.

Earlier in this chapter, I quote Theresa who – when discussing demographics – says that before school grades her prior F school was better known as “a very supportive school community” with “social services” and “meetings that were open until 9 PM.”

Several teachers at a failing, urban middle school also suggest the “community school” concept as an alternative to school grades. They tell me that most families know better than to pay attention to the grade because in many cases parents, grandparents, and even great-grandparents all attended the school and live next door to its teachers. I ask the current principal – Jennifer – how she would describe the school's relationship with the community. She says:

I think [it's] fine. We have four crossing guards. I went out on Friday and I gave them each a bag of snacks and we talk and then I talk to people in the neighborhood. Nobody, and I don't, see a problem [with the school] now.

Jennifer has only been at the school for three months, but she says that community members care more about her interactions with them than they do about the grade. She says the school has a “familial” atmosphere. Later she talks about interactions with parents:

We had a meet-the-principal night, which was well received. I make it a point to go out and introduce myself and if I hear parents come in or if I'm walking out of my office, I'm always going to say "good morning" or "good afternoon" and introduce myself and say who I am. So, I have a lot of parent interactions.

Jennifer's predecessor at the school – Alex – echoes these same themes. After discussing parents, I bring the interview back to teachers and press him to tell me: if the state is going to somehow grade or send data to the school, what should it do?

Alex responds:

What I think would be helpful, probably more so to teachers, and I think that if you provide this to the teachers it'll alleviate them a little bit more and just make their year easier, is [to help] them better understand the VAM data. Now, I think that the VAM is some kind of secret science that we're not allowed to know all of the components and pieces of, but I know that with the VAM, from the research I've done, is all of the kids are at different points and their growth is based on where they are.

Alex focuses on value-added model (VAM) data – the growth that teachers "add" to a student's test scores from year to year. This data factors into each teacher's evaluation and is supposed to reflect the school's learning gains in Language Arts and Mathematics – which effect the school's grade. He continues:

If they [students] were not on grade level or they are on grade level [or] they're above grade level, they all have different starting points. But the teachers don't necessarily know what that means. And if that was something that we could articulate to the teachers, provided through the state, I think that would help the teachers have a clearer conceptual understanding of what kind of learning gain the students need.

Because, in my opinion, learning gains are important, but the goal is [to be] on grade level. So, what can we do to get that gain for them to be on grade level? And unfortunately you can't always in one year, meet the expectation. And then the next year you have to continue to grow higher and higher. But that would, I think, help

alleviate for the teachers and for the schools some of the confusion as to where the kids are.

Elsewhere in the interview, Alex tells me that it is unrealistic to hold all students and schools to the same standard, but if the expectation is going to remain that all students achieve at the same level, then teachers need to understand where students start at. Unfortunately, teachers only receive a VAM score determined through a “secret science.” Alex explains:

I don’t even know the report that we get back with VAM. It has a number on it. Like, “Here’s your VAM.” It’s like -0.0078964. Well, what does that mean? You know? How did that number happen and where did it come from? Because it builds in more anxiety with the teachers. So if the state could do a better job of explaining that, since it’s a component of their evaluation, it’s a component of a measure for the school, and do a better job of explaining how that happens. We just call it magic math that happens and we get numbers back and [they say]: “Here you go. Here’s your VAM. You’re effective, you’re not effective.” That would help to alleviate some of that stress that teachers experience.

None of the teachers I speak with brings up VAM specifically, but almost all of them complain about attempts to link student test scores with their professional evaluations or salaries.

These teachers and principals are in agreement that grades are not – but should be – useful. They do not offer concrete responses for how to make them that way. Most resent the idea that the state should tell them how their students are doing. They feel like the experts most qualified to gauge that. To inform parents, some encourage parent involvement in the school and vice versa, creating a “community” or “familial” relationship between the school and families. If grades are to be useful, at least one principal believes the state needs to better communicate how teachers affect them.

“I’m Just a Teacher”: Teacher Calls for Recognition, Influence and Respect

In the previous section, I write that teachers want the state to listen to them. This is a sentiment that crops up in every interview, even when I do not ask about it. Teachers feel disrespected and take it for granted that policymakers are not interested in their input. As Joy, who has taught for five years at a failing school, tells me: “I’m just a teacher. I don’t have the credentials,” or Rhonda, who values her influence over kids but wishes she could influence her administration and the county. She says: “I was thinking about going for a doctorate in public policy [so they will listen].”

David complains that “in five years of teaching” his county “has been through two or three evaluation systems all created by education experts with one thing in common: they have no experience teaching.” Joy says that reform ideas “come up often but it’s always amongst teachers because we don’t feel like we’re being heard as far as those concerns to administration.” She adds that, “I try not to blame them [because] I don’t know if they’re being shut down by the person they’re going to or they’re just feeling like it’s not worthy enough to present it.”

I ask Alex if teachers and principals are cut out of the policy-making process. He responds:

You get an impression that input is occurring [but] I don’t know of any such input. And I was even privileged enough to go to Tallahassee maybe two years ago when the state was making changes to evaluations and VAM and all these new components were coming in.

And we sat there in the room with individuals in the Department of Education and they were explaining the teacher evaluation and why we needed to change things. And the argument that one of the individuals was using was: “Look how many

teachers are effective, but look at where our reading and math scores are. These don't align, so it can't be correct."

This state meeting on educator input felt insincere to Alex and worse – seemed to blame school failure entirely on teachers. He continues:

I was with a group of educators – some district people, some school-based people. We were all equally offended, because it almost seemed as if they were saying, "Well, too many people are doing a good job, so there's a problem somewhere." As an administrator, I felt like they were saying, "Well, the administrators don't know how to evaluate teachers."

The district personnel, they kind of had a similar reaction. But it almost seemed like it was just a blame game. Like, how can we...how can this be? But at no point did they ask us what [we] thought. At no point did they ask us: "Hey, look at this data. Interpret it. What do you feel?"

Because in a situation where 98 percent of the teachers are effective and our graduation rates are not at 98 percent, our reading rates are not at 98 percent, our math is not at 98 percent, I think that does beg a question of – why is there a disconnect? But, I don't think that one necessarily has to do with the other. I think that they're in a lot of ways independent of each other.

So in that interaction that I had when I was up in Tallahassee, it left a bad taste in my mouth. Because I didn't think that they really care about what we thought. I didn't really think that they were soliciting any feedback from us. They were just telling us that something's wrong, so we're gonna fix it and this is how we're gonna fix it. I know that most of our legislature, state legislatures are not comprised of teachers.

I quote Alex at length because he captures so many of the themes that arise in my other interviews. In Tallahassee he felt ignored, patronized, and blamed for problems beyond his control. He felt like legislators did not include or understand educators. He continues:

I think that the approach that they're taking to education is similar to a business approach. "We have a bottom line, this is what we have to produce, and if the teachers aren't producing it, let's cut...they can't get the incentive bonuses." But they're not really operating as educators. They think it has been taking on more of

like a business approach: “We’re stewards of tax dollars and we have to be mindful of how we’re spending things and allocating resources.” I understand all that, but I think that’s also where you lose kids.

Several teachers I speak with disparage this idea that legislators and other politicians want to run education “like a business.” David, for example, tells me that he “constantly hears people say that education should be run like a business” but that “helping kids learn isn’t the same as making money” and “if any of them [legislators] were teachers they’d understand that.”

Unrealistic Expectations, Low Pay, Unfair Accountability

Teachers also think that this “business” model makes them solely responsible for results they cannot possibly achieve. I ask Jane – who works at a B school now but previously worked at an F school – if she would want to work at a failing school now. She says:

No. Because my pay is directly connected to the success of my students and the expectations are inequitable and you don’t have the support you need to work with a population that has been deemed a D or an F school. Frankly, I just don’t know that I would have the energy at this point in my career to do that.

Jane is not the only teacher to connect low pay with unrealistic expectations.

Rhonda – who teaches at a failing school – tells me that she knows “a lot of people who give teachers a lot of flak for making too much money.” She continues:

If they knew and understood how much a teacher...like who and what a teacher is, [then] they would easily give a teacher 100,000 dollars because in one day, in one

class period, I can be a teacher; I can be a doctor; I can be the counselor – all these things at one time. And you're taking 22 kids' sense of normalcy and I'm trying to conform it to one sense of normalcy, not one time a day, [but] six times a day. And so if people were paid more for what they did that in itself would probably make a difference because you're talking about taking stuff home, and you're grading papers, on time that you're not being paid for.

Rhonda goes on to say that she works extracurricular events with no pay and buys supplies with her own money but does not “need to be paid for everything.” She is more upset about the respect she thinks the money represents. She feels similarly about her time. She explains:

Another thing is meaningful professional development. A lot of times you'll say, “On Wednesday we're gonna meet in the media center for professional development.” If I'm a teacher who doesn't struggle with classroom management, why is my time being wasted sitting in this professional development? That's not my issue. Give me something like differentiated instruction or something like that where that's a component of teaching that I may not have down pat or I struggle with. Don't sit me [in] classroom management because when you come in my room that's never an issue. If I have that down pat why am I sitting here at this...it doesn't make sense and it's like we already don't have enough time. Then you want me to go to these professional developments and it's just like I'm not paying attention because I don't need to know.

For Rhonda, low pay, high expectations, and wasted time all make her feel as if she is not respected as a professional.

Sarah, who works at a C school, sounds genuinely angry as she tells me: “I'm not willing to work 55 hours a week and make 47,000 dollars a year anymore.” She adds that:

In spite of me being high-impact in the state, highly-effective in the district, well-respected among my peers in the school and my administrators, we were offered – teachers who were tenured – 350 dollars bonus, one-time bonus. No raise. No nothing. So, I am disgusted with the whole system.

Like, Rhonda, Sarah feels most angry about the respect she feels the money represents. I ask her if a lack of respect for teachers is a problem at her school. She responds:

Oh, absolutely! It's like at the meeting that we had the other day [about] monitoring our students. I don't think there's anything wrong...I wish I could put it in words exactly how it was presented to us, because it was so unnecessary. You know, we're not stupid. And that's basically how I feel this district and our administration looks at most of the teachers.

I ask her what exactly makes her feel that way.

First of all, we have to clock in. We are salaried employees and we are clocking in. I mean, that's just one of numerous issues that have come up in terms of professionalism. Offering people who are at the top of their field a 350 dollar bonus and – let me add this – in terms of professionalism, I'm a tenured teacher with 18 years experience. I dug my heels in and refused to give up my tenured position. Well, the new teachers that work here? They're going to get...like the ones I told you have one year, two year, the teacher who piggy-backs off of my plans? They're going to get a 500 dollar bonus because of the way the state and district set it up. So, if that's not being treated unprofessionally, I don't know what is.

To Sarah, professional development feels “unnecessary” and she feels like the district and administration see her as “stupid.” She resents having to clock in and receiving a smaller bonus than the new teachers that she supports.

Theresa tells me that resentment between colleagues happens because teachers know they cannot get respect or support from superiors. She explains:

I guess what is bothering me right now is – it has become so very bureaucratic and it has become so very limited and the biggest fights that happen at my workplace are between colleagues. We have no power, so we fight against each other. It's trivial and stupid, but that's what happens. People [feel] very territorial and they're fighting over crumbs. They're fighting over absolute crumbs. I'm one of them. I don't want to lose

AP [advanced placement]. [If] I lose AP, I leave, because I don't think I would do well in a ninth grade regular class. I know that sounds really faulty, because we're in a system where people try to work their way up and out, but...you can argue that the kid who needs me the most is a 14-year-old who can't read, but I did that. I did that. I got called a fucking bitch for years. I had my car stolen. I've had my wallet stolen. I got pushed up against the wall. Let me have my AP kids.

Theresa says teachers are fighting like this at the bottom of a “top-down, demoralization process” – under the governor, their superintendent, and the superintendent's assistant. Of the latter she says: “All he does is visit high schools to shriek at teachers.” She describes a local school that dropped from an A to a C this way:

Those teachers are no longer allowed to give any assignments without administrative approval. Every assignment must be judged out of one hundred points, no matter how minor or insignificant, and every assignment must connect with standardized results. Teachers are leaving in droves. [A local F school] had an 85 percent turnover last year. 85 percent of their staff left. Most of them did not leave for other schools. They left. They left the profession. It's a revolving door because of the letter grade system that was contingent upon a standardized test that started in the late '90s. This is the fruit of that.

Theresa paints a dire picture of Florida education in which teachers like her – who teach advanced students in A or B schools – cling to their positions as others who feel micro-managed and disrespected leave.

I ask Sandra – who works at an F school – if she knows any teachers who are motivated to leave. She describes some as “desperate” to do so. She says:

My co-lab is [desperate]. He – just it's everyday something with him. Like, they'll spring things on you last minute, like this morning the new dean came to him five minutes before the bell rang and said, “I need a list of your five lowest students based on data.” But, [this was] five minutes before class was to start and they got mad at him when he told him no. And he, everyday he said he's about to explode, he's about

to blow up. It's every day. And he's like, "I hate coming to work." He's already missed all of his sick days because he's so stressed out coming to work.

Sandra's colleague is "desperate" to leave not because of the difficulty of the job or the students, but because the administration has unrealistic expectations of him and does not respect his time. I e-mail Sandra's colleague to ask if he will talk with me. He does not respond.

I ask Theresa if her principal respects the work she does. She says:

Yes and no. I received my evaluation this week from my new supervising administrator with a comment: "Needs to be more involved at [the school]." I fired back with another copy of my three-page [professional participation] log and a request to meet with him and his principal to discuss my resignations from all of my committees. He apologized and said he did 48 evaluations in an hour and must have made a mistake. He's made a lot of mistakes. The evaluation system itself is flawed, so if you're not an athletic coach I don't know if they're aware of what you do.

Teachers like Theresa feel like the efforts they make – aside from standardized tests – go unnoticed, just as their schools' most important contributions go ignored by school grades.

I include these anecdotes about low pay, lack of respect, irrelevant professional development, micro-management, and ignored contributions for two reasons. First, when I ask teachers whether their ideas about school grades and school improvement are valued – these are the ways they respond. Second, these things help contextualize teachers' feelings of frustration. Teachers feel like experts in educating students, but also feel not-valued by policy-makers.

These frustrations are not unique to the teachers I interview, or even Florida teachers. Jones, Jones and Hargrove (2003) cite stress and frustration, low salaries, "lack of influence

over school policy,” “inadequate administrative support,” “lack of career advancement opportunities,” and “lack of respect and recognition” as among the top reasons teachers leave the profession (p. 137-138). They cite a 2000 survey by the Council of Chief State School Officers, including a teacher quoted as saying: “These tests, and all of this pressure to make kids do well on the tests...it’s an insult. It’s saying that we aren’t a profession and we can’t be trusted to do our jobs, so high-pressure tactics are necessary to make us behave. They’re treating us like stupid children, they’re turning us into bad teachers, taking away every bit of pride” (p. 142). Another teacher told them: “Feeling like less of a professional is an understatement” (p. 142).

The *Orlando Sentinel* columnist Scott Maxwell writes that Florida drives 40 percent of its teachers out of the profession with “policies that are top-down, bureaucratic and generally bone-headed.” He suggests that legislators start listening “to the men and women who have dedicated their lives to education” because: “When legislators draft gambling bills, they court casino execs. When they change insurance regulations, they seek out insurance CEOs...Yet, when it comes to education, these same politicians turn deaf ears to public school teachers and often show disdain.” According to Maxwell, one of the “most-cited desires” he hears from teachers is: “Treat us like professionals” (Maxwell, 2017).

Teacher Experience and Policy Input

Joy – the teacher who says she lacks “the credentials” – tells me that no one at her county would listen to her policy ideas. I ask her if she thinks they should. She says: “Well,

I'm working towards my educational leadership specialist's degree, but no, like that's what I meant by the credential-side. Because I don't hold a doctorate, I'm not in a certain position. I don't think they will listen." I clarify by asking – "do you think that they *should* recognize your input as valuable?" She says: "

Oh, absolutely because I'm in the trenches everyday with the children. If you're only stopping in for 10 minutes once a week how are you gonna tell [me] what's gonna work best for 100 and something students who I've seen for 250 minutes a week?

Her former principal, Alex, agrees. He says that people from the state and county try to tell him how to fix school policy "without actually spending quality time there to, to take in and assess and measure things happening at the school."

Joy has ideas about how to improve the school's performance overall, but feels her ideas are ignored. She explains:

It's frustrating [to be an F school]. It is, especially right now. Other than [another teacher] and the guidance staff and the janitorial staff, the media specialist, and the cafeteria staff, I've been here the longest, as of now. And it's frustrating because I know how hard and how much I put in, but science is only tested in eighth grade.²³

Joy teaches seventh grade science, so her students do not take a state standardized test, but she thinks that she could contribute to tests they do take. She explains:

Reading and math, they're testing in all grade levels and they make up the bigger percentage [of the school grade] so I wanted to do reading across the curriculum. I wanted them to push for every class period once a week doing some type of reading assignment. That – maybe you read it, maybe you identified the vocabulary words

²³ I have removed Joy's reference to a specific teacher by name.

and created some type of word description in science but then you went to history and found a way that they could maybe look at all the site words or something and you did a portion there. Then you got into math and you read deeper into it and identifying key details so then when you get to writing or language arts you can write a summary about it.

Joy says that her school does this with math. It has math days when all of the classes somehow focus on a math problem. She thinks reading and science could be integrated in a similar way. She continues:

I just feel like we don't – they don't – push for that cross-curriculum enough and if we did the kids will be so used to seeing something...you know, when you see something over and over again you can't help but your memory starts to kick in and that's kind of like what needs to happen.

I ask if her administration is open to her suggestions. She says:

Yes and no. Because at the end of [the day], they have someone over them telling them certain [things]. It's like we don't have as much free will. So, yes and no, because they have to answer to someone as well and that person may not agree. Even though it sounds good they may be like, "naw." You know, people shut down ideas for their own personal reasons, you never know. You may mention an idea one year and then it comes back and it's finally in place two years later.

I feel like there is more they can do [to engage teachers]. A lot of times people, especially I see in [the county office], they get to these positions and they use those policy handbook answers, like someone's really programmed them to say just this. You know, "This is how you respond. Respond like this each time." It's kind of like a vague answer that never really solves a problem but you've studied the language so much it sounds good when it comes out of your mouth, so it just shuts the person up who doesn't know any better at the time. I see a lot of that going on and it's like – okay, I can't...you're not gonna listen to me.

Joy – despite having more experience than most in her school, including the administration – feels like supervisors, especially at the state and county level, do not value her input. She has ideas about how schools should be graded differently but she also has ideas about how to raise the grade within the existing framework. Yet, she believes her input only receives vague, programmed answers meant to “shut the person up.”

Joy’s school is operating under a turn-around plan and monitored by the county. Her district, like most in the state, has a special office devoted to school interventions. Teachers receive “progressions” – daily lessons plans that the county expects them to follow in order to raise scores. Some teachers find this helpful. For example, Sandra – a second year teacher – says that “as a first-year teacher that’s what you want, you want to be told what to do.” She even suggests that “that’s something maybe all schools should have for their first-year teachers.” Sandra also finds county officials helpful – “they help write the progressions, and they kind of oversee and make sure we’re doing the right thing.” Two other young teachers express similar opinions.

The more experienced teachers at the school who I talk with all disparage the turn-around plan. Martha, for example, tells me that the county officials – who she identifies as coaches – could be helpful, but “their hands are tied to very specific scripted progressions.” I ask if such progressions are ever helpful. She responds:

No, I do not. Maybe for a new teacher but not for one who’s taught for years and knows what they’re...how to teach the material. It’s not an asset at all, but for a new teacher – you know – maybe who’s been one or two years it may be an asset, but for me it becomes very tedious when I’m looking at stuff and they’re telling me to do this, to spend 20 minutes going over these 20 vocabulary words. Then stop. Then

have them get out a piece of paper. I can do that without them giving me that script to read.

Martha believes this frustrates the students as well.

I think they are frustrated because they want interactive activities. When I bring in some interactive activities for them, they love it. They love doing stuff like that, and it almost like brightens their learning, their brain, because it just gives them the opportunity to express themselves creatively.

She contrasts this situation with that at her previous, A-graded school.

I had total autonomy on my teaching: teach what I want...as long as I covered the curriculum. Whatever I want, however I wanted – show videos, have parties, do nothing...you know, whatever I wanted to do, I was able to do. Here, it's very scripted, very specific. I do very specific lessons that they provide [for] me and I can do some enhancing but it's very, very scripted.

This limit on Martha's autonomy contributes to a frustration that she is not treated as a professional with qualifying experience and training.

Even younger teachers, who may like the progressions, believe aspects of the turnaround plan are counterproductive. In her classroom, Sandra has the averages for each class posted on the board. I ask her why she posts them, and she says: "We're required to." She goes on:

Nope [I don't know the reasoning], but I'm required to put those up. I was told I had a hard deadline, I had to put them up, the posters that had the [grade] by class period, [also] which students are on target, need improvement, or need much improvement, and the class average, and how many students are in that class.

She thinks this policy has mixed results:

I think, for example, my sixth period where it's 50 percent on the last test, they're like the odd ones out. They see them and [think], "Well, those other kids are smarter than me, I don't really care." Whereas, my first and second period, I have IB and they compete with each other. So, I think for them it's really good. They say, "Oh man, second period got 88 percent? We have to raise our scores, what we're gonna do next time." But for some of the other kids, I feel like it could be very disheartening. They [the county] also started a new program [called] S.W.A.G. – it's "Students With Awesome Grades." In every class after a test they print these out for us and gives us candy and we have to announce to the class and give out rewards in front of everyone for kids that scored 70 or higher.

Sandra believes that public competition benefits some students, while "disheartening" others.

To her frustration, she is not allowed to use her discretion but must apply the same policy in each class.

Final Thoughts: Chapter Five

My interview data, as well as less formal interactions with teachers, strongly suggests that teachers do not like school grades. Teachers resent what they perceive as inaccuracies and, while not clear on concrete alternatives, they have general ideas about how grades might better reflect school quality. They want grades to account for environmental and demographic challenges, as well as school efforts to address them. Some believe that if the state ignores such challenges it will never be able to intervene in helpful ways. They admit that some challenges are difficult to explain or measure, especially when students lack exposure to successful role models or the world at large.

Teachers describe stressed relationships with the state, their districts, and in some cases their administrations. The stigma of a failing grade makes it difficult for principals to attract teachers, especially those with experience. District efforts to incentivize teachers supposedly fall short. Teachers feel frustrated by unequal funding, describing some schools as “country clubs” with budgets supplemented by wealthy parents and donors, while others struggle to pay utilities. Worst of all, they do not think that policymakers understand the links between failing school grades and these unequal challenges.

Principals tell me that grades are not useful enough for teachers, especially when student growth data is kept esoteric or unexplained and teachers take offense when state and district officials tell them what is best for their classes. Teachers point out that often these officials do not have classroom experience or – at best – have not spent time at their schools. There is a disconnect, teachers tell me, between the way these officials ask teachers to appeal to multiple intelligences in their classrooms and the way students and schools are evaluated in standardized ways.

The media reports that teachers are leaving the field because they feel disrespected and micromanaged. Teachers who I speak with agree. They feel as if policymakers listen to supposed “experts” but that teachers’ professional experience and training is not valued. They feel underpaid. They feel as if any policy ideas they have would be ignored. They feel like school grades exclusively blame them for factors they cannot control while ignoring the great things that they do, such as overseeing significant student growth. While some district interventions, like technology upgrades, aid teachers in their craft, others, like scripted curriculums, are stifling. They think the state should know all of this, but first it has to listen.

CHAPTER 6

THE ROLE OF ECONOMIC AND RACIAL SEGREGATION IN SCHOOL PERFORMANCE

Undergirding my other dissertation chapters is a demographic – racial and economic – segregation of Florida's families. It seems remiss to not try and capture part of that reality, as reflected in the words of teachers. Every teacher I speak with brings it up, as do both parents, both principals, and a nonprofit representative. Some of them speak of it with resignation, while others call it "depressing." All agree that it is inexorably linked with the distribution of school grades as well as with other non-academic measures of wellness. In this chapter, I will briefly examine how teachers and parents feel schools are separated by class and by race. Additionally, I will explore why some teachers and parents believe these separations are growing more extreme. Some fear that school grades further increase segregation while others fear that segregation consigns certain schools to failure.

In Chapter Five, I explore teachers' wishes that grades would account for environmental factors. Several unequivocally state that schools cannot solve, nor be expected

to solve, all the problems that students face. Jane, for example, gives a list of broad reforms needed both in schools and beyond them:

I think superficially there needs to be certainly better pay. There needs to be more equity in the way people are paid in the workforce so that they can provide for their children and give them a sense of safety and do so within a normal 40 hour work week and not have to work multiple jobs. That healthcare at a reasonable price, if not free, needs to be accessible so there aren't those generational worries. That the tests or assessments need to be diversified, there need to be multiple assessments that address students' abilities that can't be based on one single test, one single day of the year. There needs to be more of a portrait or a portfolio type of assessment. If all of those social needs were met – food, shelter...just basic Maslow's Hierarchy.

If we provide all of those needs and get them to the point where education can truly be a value, by that I mean one that you actually have the time and energy to fulfill. If we could somehow assess the whole child.

Jane urges educational reforms, but she places the need for them alongside the disparate needs of families and students: work hours, income, health care, and housing. Without these needs met, she implies that education cannot “be a value.” Jane illustrates how – for some teachers – segregation is part of the story of school grades. When different schools serve students at different levels of “Maslow's Hierarchy,” then those schools see their missions and interpret their outcomes differently.

Limited Options: Student Segregation by Family and Community Wealth

Martha, who teaches at a persistently failing middle school, tells me that the students are “brought down” by both the school's grade and their relative poverty. She explains:

I think if [students] constantly are getting, “Oh, you have an F school. You have an F school. Your school’s in a very poor neighborhood. Your school doesn’t have these assets.” I think that brings them down and that’s [not] what they need. They need to be brought up.

She continues:

I think [poverty] does bring the students down. I mean, we did just get brand new furniture, but they’ve had old broken tables, old broken chairs. The textbooks are like halfway destroyed. It’s just not, not the new stuff that you see at the higher-level schools.

I ask if the students are equally aware of the F grade. Martha responds: “Yes, they are. They are bothered by the label. They know that they can do stuff. They just want to be given the opportunity and the resources to do it.”

Martha clearly associates the lack of resources relative to “higher-level schools” with the lower school grade. I ask her if students also associate the two. She says: “I don’t think they can correlate the two. I think they think because it’s in a poor neighborhood, they just assume that they don’t have those fancy things.”

Angela is a parent whose children receive state scholarships to attend a private school. She identifies as black. She works as a coordinator with struggling families and regularly advises them on school choice issues. Like Martha, she correlates low school grades with a lack of resources, including good teachers. She explains:

[Part of what makes an F school different from an A school is] the environment, the neighborhood. As I said before, I think honestly that certain schools, like the C schools, the F schools, they don’t have the support the other schools do. Like for instance, if A school kids were getting out of hand or stuff is going on, I think that

somebody would get involved and put programs in place for the students who were misbehaving or not doing what they're supposed to do. Either that or they would probably kick the child out of school because they only want a certain type of child in that school. And so once the kid keeps getting kicked out of certain schools, they're going to put them in the worst schools, and I think that the F schools that I know of, they put anyone there to teach. They put anyone there to teach and from what the kids have told me, the response that they get from the teachers is not that good.

I ask if she thinks teachers only work at F schools when they cannot find jobs elsewhere. She responds: "Maybe. I think so. I think so. Or the teachers who know they can get away with not doing, not teaching to their full potential. I think that's the ones who go to those schools."

None of the teachers at failing schools whom I speak with appear to want to "get away with...not teaching to their full potential." In fact, they all seem especially motivated to work with struggling students. However, both principals I speak with tell me that it can be difficult to attract staff and that sometimes they have to accept whoever the district sends them. One principal – Alex – in particular stresses that training new teachers is resource and time-intensive and that some of them leave as soon as they can transfer to another school.

Angela took her children out of the public school, but she tells me that in her work she encounters many families whose children cannot transfer. She says that while those children lack resources in school and at home, "the system" is designed to keep them disadvantaged. She says, "It's like, [if] you're poor, basically the plan is you're going to stay poor." She explains:

The kids have limited options too when their [parents] are going to work because the system that is put in place, it's like once you start working and you make above a certain amount – and you could be making very little money, 20,000 dollars a year taking care of two or three kids – so your food stamps is taken away, you're no longer

eligible for Medicaid, only your kids are. You can't afford medical insurance through your job, and so you don't have health insurance.

Now you're over income for food stamps, but at the same time you're unable to afford a lot of food or healthy foods to purchase for your family because you're trying to get work. So now you need reliable transportation. Your monthly payments are high and so you're thinking about monthly payments, you're thinking about your rent, because your rent is going to be based on your income. And even if your rent is not that high, you still have the car payments. Now you have car insurance. You have to pay for gas. You got to pay for food. All that type of stuff is adding up. And then you realize, "I can't do this."

So, it was easier for me if I just had food stamps and I wasn't working and so my rent is extremely low. I'm probably getting child support a little bit. So maybe my rent is 100 dollars or less. I'm getting food stamps. I have five kids. My food stamps is like 700 dollars and something, and so this is a lot easier. I'm not stressing.

Angela cannot say how, but she wants systemic reform that could enable families to provide more for children and demand more of schools.

Rhonda, a black teacher at a failing middle school, also believes that impoverished students face a rigged system. She thinks that poor schools feed poor neighborhoods and vice versa. She explains:

It's so unfair to this kid because it's like – where do they wind up? And for a lot of people if they come from a neighborhood like [this one], they come back. And why do you think the crime rate is so high? Because these kids are not graduating. So, when my education...the less education I have the more I'm going to...and there's studies for it, it's like the more crime I may possibly commit because my skill set is limited. So, you have them feeding back into a community that's already poverty stricken. No, if you want to build a community you need business owners. When your jobs are created that's when crime rates go down. But if you send these kids back into the community without a skill set, without a job, it's gonna further the economic divide between neighborhoods.²⁴

²⁴ I have omitted Rhonda's reference to her particular school's neighborhood.

Rhonda laughs as she says “it’s gonna further the economic divide.” I understand her to laugh because she believes the statement is too obvious for people to naively ignore.

Rhonda describes one way that the neighborhood influences the school. She tells me that the district should consider “rezoning” so that the school will be more diverse, because when all of the kids come from one neighborhood, social problems spill over into school and students lack exposure to alternatives. She explains:

What is happening currently is that a lot of the students that live around the neighborhood are going to be pushed more into the school and taken away from other schools. And the problems that we’re having currently is that these kids, they live in the same neighborhoods, they come to school together, so if there’s something that happens in the neighborhood, nine times out of ten they may possibly make it over to the school. And so when there’s no one, no other students that have a different walk of life in the school, you pretty much get more of the same thing.

Later in this chapter I try to explain why Rhonda thinks more students will be “taken away from other schools,” but whatever the reason, other teachers agree that this neighborhood isolation happens and causes problems.

Theresa works at an A school in the same district as Rhonda. She tells me about her experiences working with students from Rhonda’s school:

I was the lead teacher [at a dropout prevention program] for one year. I continued mentoring youth for years after that, until I segued into other areas of need. I’ve worked for Make-A-Wish [foundation] as a volunteer sub. I had a kid from [that school]. It is a school of absolute blight and despair, literally the poorest of the poor. It feeds a school district that is probably the single most abjectly poverty-stricken in [the city]. It’s 100 percent black.²⁵

²⁵ I have omitted Theresa’s references to a particular program as well as the name of a school.

Theresa's diction – "blight and despair" – is incredibly emphatic. Like Rhonda, she perceives a cyclical crisis in which the school and the neighborhood influence one another. She also points out – again emphatically – that the school is entirely black.

Student Segregation by Race

Several of my interviewees, like Theresa, stress that class segregation is intertwined with racial segregation. Some go so far as to use the terms "race" and "class" or their subsidiaries interchangeably. Mid-way through my interview with Angela, I point out that she seems to use the terms as synonyms. She says:

Well, yes, because the majority of wealthy people, in my opinion, are white. And because of everything that has gone on in the past, it's taking me, I mean other races, mostly black, a whole lot longer to get to a certain type of economic status. And it's kind of harder, and I think the reason why it's harder is because we don't really know what to do, which way to go, who to go to. A lot of the wealthy people, it's basically all money and so they pass it on to their kids, to their family members. They just basically allow it to stay in the same type of race.

Angela believes that white people keep money and resources away from "other races, mostly black" including in the schools. She explains:

For whatever reason the majority white schools have – I don't know – a better education system. They have better equipment. They have better teachers. Just the surrounding areas are better and so they're equipped, I think, more than other schools to be able to give the kids a better education. Whereas, the schools in the horrible neighborhoods or not-so-good neighborhoods, they're the ones that end up with not enough equipment or the teachers are not as dedicated.

In contrast to Angela, teachers at failing schools consistently tell me that they and their colleagues do feel very dedicated to helping struggling students, but they also share – as I

discuss in Chapters Three and Four – frustrations with poor equipment and resources.

Historically, Angela's suspicions have basis. At one point, state laws "established a dual tax system to assure White Floridians that their tax dollars would not underwrite schooling for African Americans" (Cobb-Roberts & Shircliffe, 2007, p. 21). As I discuss in Chapter Five, funding is now distributed equally; however, teachers point out that some schools can supplement their funding through fundraisers and parent donations. Additionally, some research suggests that the A+ reforms in Florida led to a re-allocation of funds away from predominantly black schools, due to the bonus money given to A schools (Dittmer, 2004).

Martha is one of the few white teachers at her F middle school. All of her students are black. She sometimes hears resistance from students because of her race. I ask her to describe it. She says:

They don't like it when I use certain language. For example, because I'm teaching civics and I'm teaching it in the historical context, I use words such as "black" or "negro" because that's the terminology that was used at the time. So when I show them pictures of whites-only and negro-only water fountains, when we're talking about civil rights, they're like, "Oh, that's racism. You're showing us..."

"No, I'm showing you actual history. For the good or for the bad, this is history but we need to learn from it, so we don't do these things again." But there was a little bit of resistance from parents too. Parents called me racist, said I was racist against their children. No.

Some of Martha's black co-workers grew up in or currently live in the same neighborhood as their students. Martha only recently started at the school and lives in a different neighborhood. I ask her if some student resistance comes from assumptions that she cannot understand their circumstances. She says:

Yes. That I don't understand what it's like to live with a single mom or a single dad. That I don't know what it's like to be, you know, the only parent with children. I was a single parent for many years. I have two children I had to raise myself. So, I quit high school. I had to put myself through college.

So, it's not that I don't understand because I do understand. I understand, but I chose to better myself and that's what they need to do as well. These kids need to learn that in order to get what you want, you have to get an education, so that you can get a job.

I ask Martha if she feels uncomfortable as a white authority figure in front of all black classes. She says, "no." She adds: "I have a black brother-in-law and a black sister-in-law. [I have] several mixed nieces and nephews and I'm very comfortable with it. That doesn't bother me at all."

Martha's presence as a white authority figure in her school represents a modicum of integration progress. In the early twentieth century, Florida laws forbid white teachers from teaching black students, and vice versa (Cobb-Roberts & Shircliffe, 2007). It is troubling, however, that Martha gives no credence to student concerns. An abundance of research suggests that students of color in particular benefit from racial and ethnic teacher-matching. Studies show that students of color see increases in reading and math performance (Egalite, Kisida, & Winters, 2015), that black teachers have higher expectations for black students than do white teachers (Gershenson, Holt, & Papageorge, 2016), and that black students earn harsher discipline assessments from white teachers (Bates & Glick, 2013; Downer et al., 2016; Morris 2016).

While Martha feels entirely comfortable with her students, two parents whose black children go to majority-white private schools tell me that their children find white teachers "racist." One of Angela's daughters attends a private school that is "income-based" and

“mostly for lower to middle income families.” Angela says that her school is “majority black.” However, her daughter’s private elementary school “was majority white because it was in a really great neighborhood.” She adds:

It would be great if [schools] were more integrated, but just based on economics, based just on the system and race and discrimination, all that type of stuff, I can understand why most of the private schools are majority white.

Two of Angela’s daughters attend a private high school. She says:

There are no black teachers at [their school]. None. None. None. None. I even asked the girls, “Hey, do you know if any black teachers were hired this year?” [They said,] “nope.” They always have something to say about it. They feel like the school is racist. They feel like the kids are racist. And so I think, honestly, even if they had a certain percentage of black teachers, it would make the transition [to private school] a whole lot easier for them, but being that there’s no black teachers, there’s black administrators but you could probably count on one of two hands how many black administrators they have. Both assistant principals are black, and who knows, maybe they’re doing it because of numbers that they need to have. I have no idea.

But the principal of course is white and I’m sure they will never have a black principal. That’s probably not going to be accepted. And some of the coaches, of course, are black for the basketball or football. There’s black coaches there, but other than that, majority white.

Angela suspects that the school may hire black administrators only “because of numbers that they need to have” and is certain that “they will never have a black principal.” She says the imbalance contributes to making her kids feel isolated.

Carla – Angela’s friend and a fellow parent – also took advantage of a state scholarship to send her children to private schools. She and her daughters are black and when one of her daughters started at a majority-white middle school in sixth grade, Carla says,

“She was so sheltered [before] that it just never really dawned on her people live separated.”

I ask if she would prefer that the school was more diverse. She says:

It probably would be [a better environment]. I think it probably would be. I can't honestly think of even them having any diverse staff members, other than the janitorial or the cafeteria center, grounds keeper. It's unfortunate, but that's how it looks. I can't think of any teachers that would work there [who] are Hispanic or of any other ethnicity. I just can't think of any.

Carla says that this lack of diversity makes her children feel ostracized. It should trouble us that black families feel access to a quality education entails losing access to teachers and other students of color.

The Private School Option

Carla says that, while she and her husband think the private school offers a better education than their local public school, their daughter did not want to go. I ask her why not. She answers:

I think she had reservations of fitting in. I think she had reservations of fitting in. I think she still has those reservations. You have people who say, “Yeah, I went to Paris this weekend.” And you know, we went to the mall.

Carla encourages her daughter to be more social, despite her class and race differences with classmates. She says that her daughter has “gone out a little bit more often” but that she has faced problems. Carla explains:

[We try] to encourage her that it's okay, but we've had some moments. I've had some moments at her school, where there was some racial inappropriate things myself and I was like, "Do you want me to address it?" And she's like, "No, no, no. Don't worry about it." I let it go.

While Carla's daughter does not want her to get involved, the "racial inappropriate things" do bother her. I ask if these things mainly happen with adults or other students. Carla initially says, "adults" but goes on to talk about other students:

There was actually a couple months ago, even from the students, and she was really hurt by it. Just certain comments, especially with the whole political scene and just the comments and just insensitive. I had to explain it to her. I said, "Children are only going to repeat what they hear. You're only going to repeat what you know, so don't hold it against that person." And she's like, "Well, she should know better!" Yeah, but they only repeat what they know.

Carla says the adults act racist, but also blames adults for influencing their children. She says that the "whole political scene" is making things worse.²⁶

Angela's daughters also tell her that their high school is racist. I ask in what ways.

She says:

I'm guessing because of the lack of black presence. I honestly think there's, even though I really like the school and the education they're giving my kids are really great, I don't think there's any – I would say – cultural awareness programs there that would be able to maybe assist not only the students but also the administrators and teachers in terms of being able to relate or understand children of another ethnic background.

I ask if the racism seems intentional. Angela says:

²⁶ My interview with Carla takes place just after the 2016 presidential election.

Oh, no. No. No. No, not at all. They haven't complained about a lot of teachers, but I guess you could say that some of the...I would say most the students, especially when the Trump campaign thing was going on was making little snide remarks and stuff. So, that's when they really started saying that the kids are racist and that type of thing. And then if they knew that a teacher was voting for Trump, they're like, "Ah, he or she's racist." They were saying little things before, but it wasn't that much until The Trump campaign. [I don't have to ask], they just bring it up. As soon as I pick them up, "Oh my gosh, Mommy, you should have heard what they said!"

Angela and Carla's kids go to different private high schools, but both say that teachers and students say racist things, usually unintentionally, and that things worsened during the presidential campaign. Both blame the situation on a lack of diversity. Angela tells me: "They have classes where there are no black classmates, where they're the only ones in the classes that is black."

Additionally, there is little evidence to suggest that private schools offer a superior education. As I discuss in Chapter Four, Angela and Carla are comforted by the "college-level" work that their children bring home, but private schools lack the accountability of school grades and state standardized tests. Recent articles in *the Orlando Sentinel* have criticized some private schools for hiring felon teachers and teachers who lack certification, college degrees, or – in some cases – even a high school diploma. *The Sentinel* also suggests that students at some private schools receive poor educations. The paper points to examples in textbooks used by many Christian schools, which "focus on white men, ignore women and sometimes insult people from Africa, Asia and Latin America." Some experts worry that such lessons may make students of color – who make up 60 percent of the state scholarship recipients – feel "punished" in history class. The books "downplay the horrors of slavery and the mistreatment of Native Americans" and teach that the Civil Rights Movement arose from

“power-hungry individuals [who] stirred up the people” while “most black and white southerners had long lived together in harmony.” Scientists worry that students learn “the Biblical Noah likely brought baby dinosaurs onto his ark” and that Charles Darwin’s teachings “were tied to Nazi Germany” (Martin & Postal, 2018; Postal, Kassab & Martin, 2018).

Internal Segregation

While inequalities and racial isolation are linked to segregation between schools, schools are also segregated internally. David teaches at an A school in an urban area of Florida, with mostly white and Latino students. He is young – in his fourth year of teaching – but is already ranked “highly-effective,” Florida’s highest rating. He tells me that – like these private schools – he has very few black students. While discussing alternatives to public schools, he says:

There are some private schools that are really high quality, that I hear great things about. And I think politicians and some in the media think that they’re the answer to failing schools. You know, if families aren’t happy they can always apply for vouchers. Well, not vouchers, we don’t have vouchers but we have scholarships. But, those schools are like my classes, almost all white. Failing schools are almost all black. I haven’t looked at the numbers but I think that’s the case. So, you’re asking black kids to go to a school where nobody does, where nobody looks like them.

Theresa – like David – teaches few students of color. They attend her school but do not take her advanced placement courses. I ask her why. She explains:

[They have a perception] that there are two [Bryan] High Schools. That you go to the good one or the bad one, and it's largely divided on racial and poverty lines. That's for the kids; that's not for me.²⁷

Theresa says that she suspects that students of color feel unwelcome in the “good” Bryan High. I can anecdotally attest to a similar phenomenon. While teaching in Florida public schools, I rarely had students of color in advanced classes. When I asked academically successful students of color why they did not enroll in honors and advanced placement courses, a majority told me that the white students in those classes did not want them there (see also: Gershenson et al., 2016).

A Growing Divide: The Worsening of Segregation in the 21st Century

The more experienced teachers I speak with all tell me that – since the passage of the A+ reforms in 1999, and the growing importance of school grades – segregation has grown worse. I ask Theresa – who has worked at both F and A schools – how segregation has changed. She responds:

The racial dynamic shifted [after school grades]. My first year at [a now F-graded school] was 1992. We had students that we euphemistically called “sons of the soil,” rednecks, poor white kids. We had a lot of Vietnamese students, a lot of Asian students, and then we had probably 40 to 50 percent African-American students, largely second-generation Haitian. By the time I left [that school] in 1996 and a new principal came in, the demographic was starting to slowly, subtly shift. It wasn't until I'd been gone from the school two or three years – my mother was still there, she taught there until 2002 – right now, the school is 97.3 percent African-American, again, largely of Haitian extraction.²⁸

²⁷ I have changed the name of Theresa's school to maintain anonymity while preserving the tone of her statement

²⁸ I have anonymized Theresa's school name. My data includes school membership by race for 2016, the year my interview with Theresa takes place. I divided the number of African-American students at the high school she references by the total

There are resources at the school. The school is struggling. The school is trying. I have friends who work there. The big shift, and I blame the letter grade system on part of this, is that right now, [the high school] is overwhelmingly populated by first-or-second-year teachers. No one lasts more than a year. They've offered 6,000 dollar pay raises. They've offered incentives. They've offered incentives like moving relocation. People can't stick there at that school because of the county's expectations to fulfill the letter grade. It's all about the letter grade.

Theresa implies that because of the pressure to raise the school's grade, experienced teachers leave and consequently, white students with the means to leave do so as well.

Jane, Theresa's co-worker who has also worked at both failing and A schools, agrees.

I ask her if segregation has changed since the 1990s. She says, "yes." She continues:

[School grades are] completely a factor. There's segregation and then there's the defacto segregation of neighborhoods and then there's the additional segregation that grades prompt. If a parent...it used to be you'd move into a neighborhood, in some cases based on ethnicity and color, and maybe there was a time where that was being de-emphasized, but now you take into consideration too the school grade. So, there's even more emphasis on it or those school grades afford transfers to students who are in that neighborhood who are savvy enough to go to a different school. Then you have a brain drain. Then you have it with the teachers too. I don't know, I feel like I'm treading on my own prejudices here.

I think segregation, it's like a no-brainer. I don't know how anybody can look out at society and not see that [our neighborhood], which is one of the wealthiest areas in [the city], is going to have an A and is always going to have an A. It's a majority-white school, although it may not be a vast majority, but then you look at the other population that's there, that their ethnicity is not as important as their economics. It just goes from there, that your D and F schools are going to be largely impoverished and people of color.

number of students. I find the 2016 percentage of African-American students as roughly 86.9 percent. Membership by race is not available prior to 2014.

Jane believes that students “who are savvy enough” or who have the means, will leave failing schools. She implies that the means and the savvy correspond to income and race, leaving “impoverished and people of color” at F schools.

She mentions “brain drain” – the idea that when academically successful people leave neighborhoods then it becomes harder for those neighborhoods to improve academically or economically. I point out that some believe that if a school fails “then the students should jump ship and go to charter schools or get state scholarships to private schools.” I ask if leaving the school or neighborhood is an appropriate answer. She responds:

I am not a fan of charter schools. I think charter schools not only enhance defacto segregation but they divert funding from the public schools. Is that a fair solution? No, because number one, there has to be the initiative and the support on the part of that family to send their kid to charter schools and they may not have that in place.

Some people do not have what it takes to do that. I don’t think that’s a solution at all. I think putting money back into those areas and building communities with services for healthcare, and food banks, and clothing, and childcare, and educating parents, providing adequate housing, jobs...those are the services that need to be emphasized, not charter schools. Charter schools are a bandaid on a problem that is only going to get bigger if we keep trying to solve it with schools.

Jane stresses that – even if the students who transfer to charter or private schools find success, many other struggling students are “left behind” at their poor public schools. Jane believes that, rather than help these struggling schools, the state government wants to see them close. She explains:

Part of the population [at failing schools] does leave to go to those charter schools. If that’s the legislature’s solution, then what’s their solution for the people left behind? Isn’t it “no child left behind”? Then we have these children who are the most in need

all in one place where nobody wants to go. If they want their solution to be to isolate the struggling population and initiate a sink or swim policy, then that's what they're doing.

Jane tells me that our conversation is "depressing" and that "the system will probably implode." I ask her what she means by "implode." She responds:

I think that we're moving more towards an elitist [system]. We're going back in time in a sense. The negative part of me thinks that, that we're moving towards education being an elitist acquisition.

Jane, Theresa, and Martha all agree on this. They have all worked at both failing and A-graded schools and they all worry that education is increasingly dividing the rich from poor and white from black.

Sarah – who has taught for 18 years in a small town – also wonders about the students "left behind" when their classmates transfer to charter schools. The high school where she teaches has a C grade and has faced problems in recent years as the town's population has grown smaller. Because Florida funds schools based partly on the number of students, less students means less funds. Sarah explains:

I don't think there's anything wrong with school choice. But here's what happened: [after students leave for other schools], our building is no different. In terms of impacting my students...we have almost a 50 percent free and reduced lunch rate. How many of those parents have the means to get their children to another school? They don't. So, the less students we have, [the less] we're paid. So if they're pulling all these other students to put them in other schools that offer more, and our schools are left bare-boned, to offer these kids who already struggle as a result of poverty, what do we do with them?

It's just going to get worse and worse. And there's, I've talked to some parents who are middle class, like you and I, and they have two working parents in the home. How do you get a kid to another school across town that now has more money than our public school? Which we still have to run all the building, I mean, the electricity, all those kinds of things, are not going to change. The prices are not going to go down.

Most of my study participants explicitly state that they are not attacking alternatives to public schools and parents I speak with favor charter or private options. David, Theresa, Sarah, and Jane all acknowledge that some such schools benefit students. The problem for these teachers is the practice of some public schools shepherding students to charters, where students may have less opportunities – in order to preserve high school grades. These schools are often lower-performing and lack diversity. They also worry about the students who do not transfer – whether by choice or not.

Jason works for a nonprofit that helps pair economically struggling students with state scholarships to private schools. His organization helped Carla and Rhonda's children. I ask him if programs like his might further worsen segregation. He tells me:

When Pinellis [county] achieved unitary status and was allowed to do [away] with the forced bussing, mandatory integration requirements and people were allowed to go wherever, that did increase racial segregation in the schools, but it did not lead to a worsening of test scores among the black students that remained behind. The overall school scores declined because the white kids were gone. The black students in the area still performed just as poorly today as they did 10, 15, 20 years ago. In fact, Pinellis has, if you're a low-income black person, the worst place you could live is South St. Pete. It is by far the worst in the state. If you were a black student in these schools they struggled with your academic achievement while they were integrated and now while they [are] essentially de-integrated.²⁹

²⁹ In our interview, Jason references reporting by the *St. Petersburg/Tampa Bay Times*. Specifically, he refers to the 2015 series: "Failure Factories" which documents stories of students at five struggling schools in the county. The series has

Jason sees his mission as helping black kids leave bad situations, rather than improving the overall quality of schools. Yet, he does note a trend of white kids leaving schools while those schools continue to fail. He continues:

Does school choice lead to more segregation? The answer to that is complicated. It depends. In some places it's been found to increase racial integration, especially in areas where essentially you have a white population in private school and you got the minority population in public schools, giving minorities access to private school scholarships, especially scholarship programs that are means-tested, may actually improve racial integration. Are you talking about individual schools, or the private school, or the public school, the overall district? There's a whole bunch of ways you can measure it.

Jason points out that in cases like those of Carla and Angela's children, black students are now attending schools with white, wealthier peers – which is a form of integration.

Ultimately, he tells me that this is “an extremely complicated issue” and “depends on local conditions and a bunch of other factors you can't control for.”

Excluding Students to Raise Scores

Another possible driver of segregation is clearly intentional: encouraging to leave students who risk decreasing a school's scores, especially its graduation rate. David teaches 12th grade English; so, he regularly talks with seniors about their graduation qualifications. David tells me that, while he does not believe it is widespread at his school, he has heard

prompted rebuttal articles, updates, and helped prompt a federal civil rights investigation. For the original series, (Failure Factories, 2015).

friends at other schools in the county describe using charter schools to “hide” students. He explains:

I don’t want to completely attack charter schools. I think they have their place. But there’s no regulation, there’s no necessary qualifications. So, I don’t necessarily encourage kids to go there. I know some schools do though, if a kid’s in danger of failing. Counselors will say: “You have too many credits to recover and it might be easier for you to do that at this alternative school.” So, they frame it as in the best interest of the student, and I’m sure some believe that, but it all depends on the charter. There’s charters in old grocery stores, in malls. But what I’m getting at is – there’s an incentive for the school because when those kids leave, the kids that might not graduate, it means that the school will likely have a higher graduation rate. It helps the school out, even if it doesn’t ultimately help the student.

David suggests that schools encourage struggling students to attend charters, not because the students will achieve more, but because they risk dragging down the public school’s graduation rate. As I discuss in Chapter Eight, these at-risk students are – according to my data – majority black and Hispanic. Between 2010 and 2014, the mean graduation rates for black and Hispanic students were 76.2 and 78 percent, respectively; while the mean white graduation rate was 83.77 percent.

David is one of three teachers who mentions the pressure for students to transfer to alternative schools. He is one of five who questions the quality of and requirements for charter schools. Theresa does both explicitly. She tells me:

You can start a charter school today. You can just walk up to the downtown office and say: “Here’s my budget; here’s my proposal,” and get money from the state. Most of our charter schools...there are a couple that are very good, like [a nearby school for gifted students] and [another] specifically for kids with extreme learning disabilities. Those make sense to me, that’s fine.³⁰

³⁰ (a) I have omitted Theresa’s references to specific schools by name; (b) Theresa is correct that anyone can open a charter school in Florida, pending approval of their application. If approved, applicants must complete a training program. The application and details on the program are available on the Florida Department of Education’s website.

Theresa goes on to say that most charter schools in her area “are working” but describes some that are “disasters” and one where the staff harmed autistic students. Her primary concern is the lack of a guarantee for students to succeed. She describes charters this way: “an open money season, an open market, anyone can start, and then they’re not held to the same standards as public schools.” She tells me that she recently read a “fascinating” study by ProPublica that I “need to read.”³¹

To follow up on my interview with Theresa, I seek out the ProPublica study. The piece – a collaboration with USA Today – seems to corroborate David and Theresa’s claims about student push-outs. The study focuses on Orange County Public Schools in central Florida, notably Olympia High School – an A-graded school ranked in U.S. News and World Report’s top 1,000 high schools. ProPublica compares it to nearby charter high school, Sunshine High – “an alternative school run by a for-profit company” that “stands a few doors down from a tobacco shop and a liquor store in a strip mall.” Unlike Olympia, which has Advanced Placement courses, clubs, and sports teams, Sunshine “offers no sports teams and few extra-curricular activities.” Students – mostly students of color – “sit for four hours a day in front of computers with little or no live teaching.” ProPublica suggests that students mostly goof off, listen to music, and cheat on tests. One student recounts being robbed twice in the parking lot. According to the report, similar alternative schools “are sometimes taught in crumbling buildings, school basements, trailers and strip malls.” This is eerily reminiscent of Cobb-Roberts and Shircliffe (2007)’s quotation of former school principal and Director of

³¹ ProPublica describes itself as: “an independent, nonprofit newsroom that produces investigative journalism with moral force.” In 2017, it – together with *The New York Daily News* – won the Pulitzer Prize for public service for an investigation of New York City Police Department.

Negro Education, J. Irving Scott, that in Jim Crow-era Florida “it was common for schools serving Black children to be housed in churches, lodge halls, and sawmill camps...[that] often lacked drinking water, sanitary toilets, desks, and adequate supplies” (p. 23).

ProPublica interviewed 32 students in Orlando and 15 of them “said that because of academics, they had been denied admission to regular public high schools or told they had to transfer from them to alternative programs.” Stephanie Langer, a special education attorney, claims that “schools sometimes push students and their parents to leave by telling them they won’t graduate otherwise,” a practice she calls “bullying.” ProPublica accuses Olympia of maintaining an A grade by “shipping its worst achievers to Sunshine” – one of many alternative schools that “become warehouses where regular schools stow poor performers to avoid being held accountable.” In exchange, charter schools receive millions in management fees from the state and in turn pay tens of thousands to lobbyists to “lower the accountability bar further for alternative education.”

According to ProPublica, schools like Olympia use charter schools to “hide dropouts and game the system.” Sunshine’s 2016 graduation rate was 3.5 percent and in 2015, its management firm – Accelerated Learning Solutions (ALS) – reported 1,038 students in the county withdrawing for adult education. Meanwhile, Olympia’s graduation rate was 90 percent and Orange County reported a mere 211 dropouts. This “loophole in state regulations” means that Sunshine “claims virtually no dropouts” even though “ALS cannot say where Sunshine students actually went – or if they even took GED classes at all” (Vogell & Fresques, 2017; see also: Dynarski, 2018).

Theresa tells me about a new charter school in her district that “was supposed to be their jewel.” According to Theresa, it has not worked out that way. She explains:

[The school] can’t keep teachers for more than a week. They’ve already fired their principal. The kids are literally running the building. Five-year-olds punching teachers. Kids running wild. Five security officers, four assistant principals, and a principal who locked himself in the office and cried all day. I don’t know why it’s not in the news. The district is hiding it. The district [thought] all you got to do is give kids breakfast, lunch, dinner, and a uniform, and their lives will be fixed. Failure. Utter failure.

Theresa says that the new charter school is “entirely black” and that its creation prompted the closing of a more integrated public school, where students had better opportunities.

Theresa’s example is not unique. An increasing number of Florida’s students of color are attending schools that lack a white population. School choice – whether it be a private school, charter school, or a different public school – is helping to reverse the already poor integration progress accomplished in the prior few decades. While some value freedom to choose over diversity, a majority of poor students lack such freedom. Public and charter schools that serve primarily students of color are more likely to be low-income and to earn D and F grades (Orfield & Ee, 2017).

The Associated Press finds that charter schools “are among the nation’s most segregated” and “are vastly over-represented among schools where minorities study in the most extreme racial isolation.” According to its report, in the 2014-2015 school year, over 1,000 charter schools nationwide had minority enrollment of 99 percent or more. While this may not trouble advocates of school choice, the Associated Press counters that “schools that enroll 99 percent minorities – both charters and traditional public schools – on average have

fewer students reaching state standards for proficiency in reading and math” (Moreno, 2017). This supports Theresa and David’s arguments that, while school choice and charters can benefit some students, we should feel wary of low standards and racial segregation.

Final Thoughts: Chapter Six

Issues of race and class infuse all of the chapters of my dissertation. The teachers and parents I speak with are quick to bring them up as explanatory variables. Some sound hopeful, some depressed, others angry. They all believe that A schools are primarily white and upper class, while F schools remain primarily populated by poor students of color. They also believe that the problem is growing worse.

These fears are confirmed by data. In a 2017 report for the LeRoy Collins Institute at Florida State University, Orfield and Ee (2017) show that the segregation of public schools has grown alarmingly over the past decade (see also: Cobb-Roberts & Shircliffe, 2007). In 1995, 29.6 percent of schools were at least 50 percent nonwhite. That number jumped to 44.2 percent by 2005, and 54.8 percent by 2015. In 1995, 10.6 percent of schools were at least 90 percent nonwhite; 2.1 percent were 99 to 100 percent nonwhite. In 2015, those numbers were 20.2 percent and 3.7 percent. In other words, the likelihood that nonwhite students attend school with nearly or exclusively no white students has doubled since 1995.

Between 1995 and 2015, white enrollment in public schools dropped from 58.7 percent to 40.4 percent, but the exposure of nonwhite students to whites dropped even more precipitously. The chance that a black student attends classes with white students fell from

40.1 percent to 26.3 percent; the same chance for Hispanic students fell from 34.5 percent to 27.5 percent. Students of color are therefore more likely than not learning alongside no white students, attending what Orfield and Ee call “apartheid schools.” These schools are much more likely to be low-income. In 2015, schools with at least a 90 percent nonwhite student body served 68 percent low-income students. The researchers call this “double segregation” (see also: Bates & Glick, 2013).

Orfield and Ee conclude that in the 2000s, after the adoption of school grades, “a very disproportionate share of schools with double segregation by race and poverty were branded with D’s or F’s” and “the state, under the leadership of Gov. Jeb Bush, blamed them, sanctioned them and encouraged the growth of charter and voucher schools” (p. 7). By ignoring segregation and adopting color-blind policies, “the goal of racial diversity was ignored” (p. 7). This process has driven many experienced teachers to leave failing schools, and teachers suggest that the students able to do so follow them.

Teachers emphasize to me that race and class are not arbitrary. As long as schools remain segregated, these teachers argue, their needs will differ. School missions should differ. Consequently, school grading metrics should differ. Teachers sound angry and exasperated in explaining this. They do not understand how anyone cannot see it. Jane refers to these realities as “Maslow’s Hierarchy” and she argues that without satisfying each level, education cannot be “a value” to anyone.

Alternatives to public schools are – in some cases – also problematic. The parents I speak with lament that the private schools where they send their children of color are almost entirely white, and that their children have difficulty fitting in, encountering instances of

intentional and unintentional racism. Media accounts suggest that private schools do not necessarily provide a quality education and they lack the oversight of school grades and standardized tests.

None of the teachers I speak with view charter schools as a solution. Evidence suggests that charters fail at a higher rate than traditional public schools, which is partly explained by the low threshold to get a charter approved. Some of these schools open in substandard conditions with little oversight. According to multiple researchers, majority-black charter schools – in Florida and nationwide – are often under-funded and fail in higher proportions than majority-white schools. Public school teachers complain that charters draw funding away from public schools, funding badly needed to maintain and improve buildings, technology, and student services. Charters may be contributing to worsening segregation when they replace more integrated schools or pull students of color into concentrated environments. Reading descriptions of underperforming, majority-black charters brings to mind disturbing images of Jim Crow-era segregated schools, and seems to make a mockery of even the offensive “separate but equal” doctrine.

Some teachers, and some in the media, fear that schools are intentionally pushing out at-risk students who they suspect might bring down their graduation rates. In order to maintain high school grades, it seems that these schools encourage students to transfer to charter and private schools – whether they will receive a better education there or not. According to reports, these schools often do not track whether students – after withdrawing – ever attend adult education, GED or alternative classes. As black and Hispanic students have lower graduation rates than their white peers, they are at the most risk from this practice of

“hiding drop-outs.” Moreover, they lose the most in this entire, disturbing cycle of re-segregation.

CHAPTER 7

TEACHER THOUGHTS ON THE PUBLIC STIGMA OF FAILING SCHOOL GRADES

I began teaching at a Florida high school in 2009. As I write in my introduction, the school had just dropped from an A to a C and I vividly remember the tone of our first staff meeting as depressed. The staff and teachers felt confused, dismayed, and embarrassed in the wake of the grade announcement. Their educational practices had not changed; their demographics and the state's formula had. They wanted the community to understand that. I think the principal tried to strike a balance between a stern diagnostic and an inspirational call to arms.

That meeting, perhaps more than anything else, inspired my dissertation research. I wondered: if we felt badly after receiving a C, after years of maintaining an A or B grade, then how do staff at schools that consistently fail feel? How do students feel? What impact do those feelings have on their policies and performances?

Upon beginning my research, I also wondered if the media worked as an additional source of shame. I suspected that reporters might buy into the grade system and use failing grades to admonish local schools. In a review of articles in Florida's largest newspaper, the *Tampa Bay Times*, between 1999 and 2014, I find this is not the case. Instead, the media seems to have worked as a microphone to amplify the voices of disaffected teachers. cursory reviews of the *Orlando Sentinel* and the *Miami Herald* reveal similar findings.

In the *Tampa Bay Times*, a fifth-grade teacher describes her school's D grade as "like the Scarlet Letter" and "kind of demoralizing." An associate superintendent of curriculum laments that schools "have to live with [grades] for a year." A president of a local teacher association calls grades, "the deliberate subjection of our children, our teachers, our support employees and our schools to humiliation and ridicule." Another teacher says that the FCAT test "was just to bat [F schools] over the head so they can fail again," adding, "You can't do this to people" (Haggerty, 1999).

In each of my interviews, I ask teachers and principals how they felt when the state released their school's latest grade. Some have never thought about the question, but after considering it their emotions begin to spill out. Others have an answer right away, as if the feelings are still raw. Others seem a bit taken aback by the question. They simply accept grades as a reality, unpleasant or not. I also ask teachers how students feel about grades. I receive mixed answers: some think students are quite aware of the school's grade, while others do not. In this chapter, I explore these answers as well as the possible benefits or harms of a failing grade's stigma.

First, I review what teachers say about their feelings and the media and community portrayals of their schools. Second, I delve into how teachers think school grades affect students – directly and indirectly. I find that while teachers consistently express feelings of embarrassment and hurt at receiving failing grades, their thoughts on how grades affect students vary widely.

“A Punch in the Gut”: Educator Responses to the Reception of a Failing Grade

Alex, who recently departed as the principal of an F school, captures much of what teachers at his school and others tell me. I ask how educators react to an F. He says: “It’s disheartening. I mean, it hurts.” He continues:

It hurt me as a principal. I was actually on vacation when the school grades came out and they sent the e-mail with all the, the database and, you know, you look. And it kind of ruined the rest of my trip. Because you take it personal. And, what I saw last year was teachers pouring their heart and soul into kids. Teachers really doing the best that they could with what they had to work on, you know, reading, to work on math or civics or science and do everything that they knew to do. Even resource people plugging into classrooms. People taking up classrooms and teaching, co-teaching side by side with teachers, and doing everything that they could for the success of the kids.

Alex feels hurt that the grade does not capture the efforts and passion of his teachers. He also feels frustrated by the metric. He says:

We came up short of a D. It’s a punch in the gut. Because, especially last year, it was eight points for us. We were eight points away, that’s less than one percent, but eight points away from that D threshold. Now, a D isn’t anything that you want either, but it’s better than an F.

I ask Alex about the reaction of teachers.

So, when the teachers came back initially, for pre-planning this year, I started with that. I said, “Well, we need to talk about our school grade.” Because I knew that needed to be something front and center. “Let’s talk about it. Let’s see, like identify what happened. Let’s look at where our points totals fell.” But then we flipped it and looked at what we can do differently this year so that we can continue to grow. I could tell just by looking at the room, when the teachers were in there, just like anguish.

Maintaining Morale: Establishing Plans of Action

Alex acknowledges teachers’ anguish but goes on to say that he “flipped it” to try and build morale. This combination of anguish and morale is representative of what teachers at his school – and other failing schools – tell me. However, a few teachers feel like the stigma of an F grade makes maintaining that morale difficult. For example, Sandra, a young teacher at Alex’s school, tells me:

I think [the] school grade is just so interesting. Like, they just change how they’re scored. In Texas, they just use: “needs improvement,” which I think is better because having an F, I feel like sends a message like it’s horrible. Like it’s so bad. Like no one should be sent here. Whereas, “needs improvement” means you could improve. There’s a chance for you. That’s why in our schools you can’t give a kid zero, that’s what I feel like F is.

Sandra points out that her school has a policy that teachers cannot award a student a zero for work they complete. She thinks a policy that treats children one way and their schools another is hypocritical.

Sandra’s colleague Lakisha, who has been teaching for 16 years, agrees. She tells me:

I personally don’t like that stigma [of teaching at an F school], because I personally feel bad, because I’m a teacher at the school and I’ve seen students say that [we’re an F] and I’ve heard people make those comments. And it’s definitely not the teacher’s [fault], it’s just a combination of things.

I ask Lakisha what she means by feeling “bad.”

I feel bad, because I'm a teacher at the school and I'm coming from a school that was a letter grade upper-B. Have I been in an F school before? I have, but I have not been in an F school that was continuously an F for a number of years and had not moved up. That can be a little bit degrading in some aspects. So, that's my concern.

Lakisha suggests that it is "degrading" to teach at a school that persistently fails.

David – who teaches at an A school – says that the feelings Lakisha describes will prevent him from ever working at a failing school. He explains:

I remember when I first got my license and I was applying for jobs. And I have a friend who is a teacher and she would ask me what schools I'd applied to. If I told her a failing school, she'd shut me down right away. She'd say, "You don't want to work there." My mom said something similar after she saw a school's name in the newspaper. And I began to feel like – there's just no way I'd work in those places. It wasn't about the job being hard, it was about the embarrassment of it. Like, if I went to one of those schools, what would that say about me? You know, about my ability to get a job?

David has never worked in a failing school, but he feels like teaching at one would send a signal to his friends and family that he was not a good enough teacher to get hired elsewhere.

Under Pressure: The Unique Stress of "Core" Content Teachers

The A and B school teachers I speak with, like David, speak more negatively about failing grades, perhaps indicative of why they have chosen to teach at non-failing schools. Jane tells me that when she worked at a C school, grades had not yet become as significant as they are now. She says that as a result she felt "pretty inoculated from" the grade at that time and more motivated by "what would best serve our population." However, Jane says that now, given the greater significance of grades, "the nuances and impacts of an actual school

grade and in particular upon the teacher and the administration” can “trickle down into the psyche of the student body [and] it’s certainly pejorative.”

Later in this chapter, I focus on how grades might “trickle down into the psyche of the student body.” For some teachers, the impact is more immediate. Alex says:

Especially the ELA teachers and the math teachers [felt anguish] because they know that so much of the burden falls in their areas. Now, that’s not to say that social studies teachers don’t teach reading and that kind of thing, but those teachers in particular – ELA and reading teachers or math teachers – you know, it is what it is. And you go through it, and you almost hear grunts and like, “Ugh. Man...” And they look at it and break it down by grade level and it’s like, “How did that happen?”

Theresa – Jane’s colleague at a B school – describes a similar pressure, albeit from a very different principal. After her school dropped from an A to a B, she says:

[The first staff meeting was] funeral in tone. Our principal is very conscious of his perception to his bosses, which is the most terrible way I can put it, and he started by saying, “We’re a B, guess y’all got to step it up.” The abuses that he has heaped upon the Algebra One instructors is unconscionable. The other problem this year, is that the ninth and tenth grade team is so abused for ELA, because his logic is: if they can bring their scores up significantly, it would temper the math problem. They’re not allowed to teach full works of literature at all. Their supervising administrator, who has no background at all in ELA – none – has sat in on meetings with them and said, “Why are you doing that? They just need test prep.” Two of them are resigning and we’re losing another one in February. They can’t take their English classes.

Theresa describes an internal policy pressure as well as a lack of respect from administration.

Like Alex, she points out that some departments are more affected than others by grades.

“Even When It’s Positive, the Negative Is There Somewhere”: Teacher Reactions to Negative Press Coverage and Other Community Reactions

Teacher quotes in this chapter suggest that teachers see failing grades as degrading, stigmatizing, discouraging, but in some cases, motivating. Some feel an intangible impact, while others face more measurable administrative pressure. All of them say that their feelings are complicated and amplified by press and the community. For example, Rhonda – who teaches at Alex’s former F-graded school – says that she feels “overshadowed” professionally by the public perception of her school. She explains:

Me and some of my colleagues were at a coaching seminar with other coaches in the district, and we told them that we came from Thompson Middle School. And they had this look on their faces, as if we did not know as much as they did. But when we started talking their perception changed. So, I think just the name of the school deceives them. It doesn’t matter who you are, if you come from a particular school then there are certain perceptions about you like you’re not a good teacher.³²

Rhonda says that these perceptions are far from accurate. She continues:

In reality, people that teach at schools like Thompson generally are probably better prepared than some other schools, because you have such a microscope on you as a teacher, and you’re forced to do certain things that other teachers at schools are more...that probably have higher school grades do not.

Rhonda wants to work at a school with struggling students, but she did not expect the degree of professional misconception in her district. Other teachers, like David, are so aware of the

³² I have substituted Rhonda’s school name with the pseudonym “Thompson Middle School” in order to anonymize her while still conveying the name recognition her school has in the district. For consistency, I repeat this substitution in other interview excerpts.

stigma that they avoid working at failing schools. David tells me that his father does not think teaching is a prestigious enough career and “the only pride he gets is that I at least teach at a good school, a high-graded school.”

Other teachers also tell me that they struggle to defend their schools to family, friends and the community. Rhonda says that:

People just have a perception of the school because of where it’s located and then the school grade. Now, being one of the only middle schools that has a F, people hear the name and they just assume certain things about Thompson Middle School. They say that the kids are bad. But I’ve been here eight years, and granted there may be some challenging students, but that’s at any school.

Rhonda’s colleague Lakisha, tells me that the school’s reputation made her “hesitant” to work there. She explains:

When I did move to [the city], they would ask me, “What school are you gonna be working in?” And I would say Thompson Middle School. And they were like, “Thompson? Are you serious? You want to be at that school?” So, that was a little nerve-racking and it made me hesitant. But once I got to the school and I realized they’re just like a lot of other children that I’ve experienced. They just need coddling. They need love and once they realize that you really, truly care about them as an individual, then they begin to work for you.

Rhonda is glad that she came to the school despite warnings from family and friends, but she says that the media’s description of her school as “in crisis” is accurate. She says:

When you have single digit [test scores] at a school, I heard not too long ago when we had 640 students taking the FSA and only 40 passed the reading part of the FSA. That’s a crisis. But why does that crisis exist? I can’t say it’s all students. I think it’s a part of instruction, as far as we are given to them to instruct them. But it’s a number of things. But the school definitely is in a crisis. It’s a rough area of [the county]. The students definitely come from low socio-economic backgrounds.

While Rhonda agrees that her school is in a crisis, she thinks people ignore why the crisis exists. I ask her if the F grade or the media's descriptions accomplish anything good. She says, "no." Then adds:

Even before I came to the school, I found myself having to explain and say, "you know, it can't be that bad. You know, there are things that can be done." I find myself doing that all the time, because I'm a teacher at that school and I don't want...Even at my grandmother's funeral my family were like, "Well, what school did you get a job at?" And I told them, I said, "Thompson Middle," and they were like, "Are you serious?" You know, it was just such a bad thing and I'm like, "It's not, it's rough but it's on the rise. It's not that bad."

Rhonda thinks that her family and the media should pay more attention to the positive things that teachers can do. She thinks people need to know that the school is "on the rise" and "not that bad."

Several other teachers at Rhonda's school make similar statements. Joy, for example, tells me that she sees her school's name in the newspaper a lot. I ask her if coverage is ever positive. She says:

When we won [a science competition] and when we did, I think it was a community service and it got featured in the thing, but just two out of all of the eight or nine other things I saw, only two of them were positive.

Joy tells me that this negative press is hard to read. She says:

Sometimes even when they say something positive like, "These kids are doing such and such, although they're coming from a low-income." Like, you don't have to put that in there. Even when it's positive the negative is there somewhere, you know?

Rhonda shares this frustration. She thinks the F grade obscures positive aspects of the school that would attract academically successful students. When I ask what the media should report on, she says:

The great things that we do have to offer. I will say, they could definitely say the traditions that I've seen so far that are present at the school. They do have a great sports program. They have things, [like] a great Baccalaureate program. They can speak to the positive aspects of the school and support that part of it. We have teachers who have been teaching there for a number of years, many, many years, who are truly dedicated. So, there are great aspects of the school and it is one of the oldest schools in the district. I think from a historical perspective, it's a well-respected school. It just needs to turn around.

Teachers fear that the school's grade and press coverage will scare away students who could help turn it around. They worry that the F grade makes it look doomed rather than in need of support.

Possible Success: Do Failing Grades Inspire Success?

There are some signs of Rhonda's school turning around. I begin this chapter discussing Alex's meeting with Rhonda and her colleagues, at which he reports the "anguish" of teachers at receiving an F grade. However, he describes his attempt to "flip" it and focus on how to improve. Alex says:

We went into it with that mindset [of how to grow] and came up with a goal as a faculty. It wasn't something that I put together and said, "This is what we're going to do." It was as a faculty. And, it changed the mood of the room. And this was our first day of pre-planning, but it just changed the mood in the room, because it went from, "Oh my goodness, I can't believe this happened," to, "Okay, but this is what we're going to do to solve it." And it served as a morale booster in a lot of ways, for the teachers to say, "Okay, it's not a good situation we're in, but we can get out of it and we're going to do it together."

Alex emphasizes that the positive spin and plan of action was not his effort alone, but a collaboration with staff.

Four months later, the county transferred Alex to a different school. In the media, it praised Alex's work at Rhonda's school, despite the F grade. Teachers also praise Alex and tell me that they have no idea why the transfer occurred. Even Alex's successor – Jennifer – tells me that she does not know the reasons. Multiple teachers tell me that such changes in staff disrupt students and make it more difficult for them to forge meaningful relationships with administration. At the end of that school year, the school grade rose from an F to a C – a grade it has maintained since.

At the end of that school year, the county also launched a media campaign to showcase its efforts to turnaround the school. It offered significant bonuses to attract qualified teachers to apply there, resulting in a flood of applicants.³³ It is unclear what caused the rise in grades: Alex's turnaround plan with teachers, his transfer and replacement, bonus money for teachers, new applicants, or the county's increased scrutiny, which included policy interventions. Likely, a combination of these things have had an effect. This series of changes prompts two questions: first, did the F grade help or hinder the improvement? Second, have students benefited or is the increase in scores an example of the curriculum perversion I discuss in my literature review?

³³ In order to protect the identity of Alex and his co-workers, I have withheld details of the county's program, such as the amount of bonus money and the details of its media campaign. I have also refrained from identifying the exact years that these events took place. While this obscures important aspects of the story, it is necessary to maintain the school and staff's anonymity.

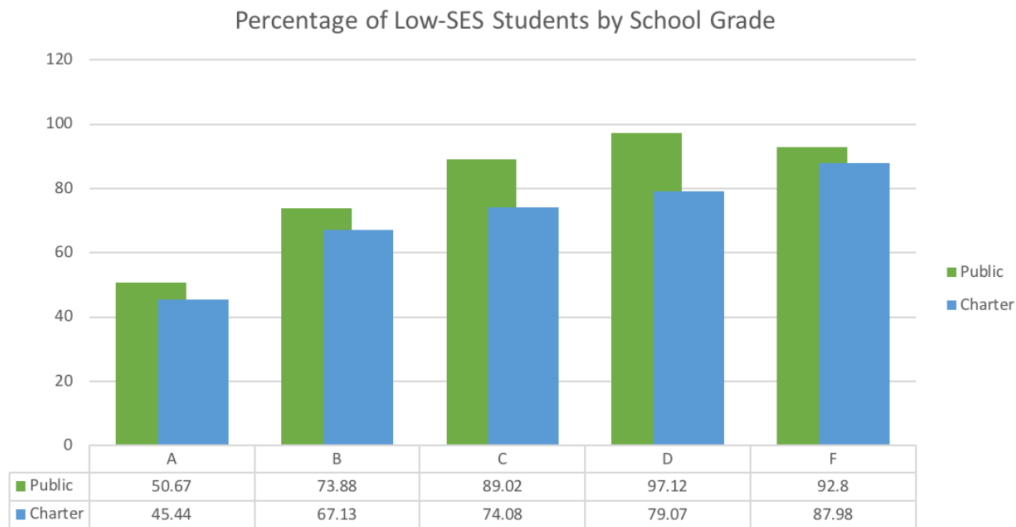
Did the F grade help? If Alex and Jennifer are correct in their statements – quoted in Chapter Five – that an F grade makes attracting staff more difficult, then it is possible that, without grades, the county would not have needed to offer a large bonus to improve the school’s faculty. The superintendent of the school’s county said as much to the press, explaining that school grades make improving schools more difficult because an F repels teachers and parents.³⁴ It is also possible that the grade is ultimately irrelevant. Although researchers like Bowen and Trivitt (2013) have concluded that the stigma of a failing grade inspires reforms, Alex’s school earned D and F grades for nearly a decade, only “improving” after significant county intervention.

As I discuss in Chapter Eight, based on school grades between 2010 and 2014, roughly 50 percent of schools that earn a D or F grade earn another failing grade the following year. Roughly 40 percent earn another failing grade two years later. So, roughly half of schools do not appear to benefit from the stigma attached to a failing grade. Moreover, as I discuss in Chapter Six, school grades seem like a proxy for socioeconomic status (SES). A 2019 article published by the Florida Education Association, titled “The Reality of School Grades” and published by the Florida Education Association argues that this continues to be true (“The reality of school grades,” 2019). The authors describe grades as “having schools ranked by socioeconomic status of their students” and provide the following graphic to illustrate as much.

³⁴ I refrain from quoting the superintendent verbatim or identifying the newspaper, again to preserve the anonymity of participants.

Figure 1

Percentage of Low-Socio-Economic Students by School Grade



From: “The Reality of School Grades,” Frontline, 2019, *Florida Education Association*, <https://www.feaweb.org>

As the bar chart illustrates, grades for both traditional public and charter schools align closely with low-socioeconomic status.

I conclude that, despite limited success stories, the shame of failing grades is not worth it. At best, a failing grade draws attention to a school, attention that should clearly be drawn by the school’s challenging economic conditions. At worst, a failing grade causes unnecessary pain and humiliation, making it more difficult to attract staff and affluent families who might help the school improve.

Secondly I ask, did Thompson’s turnaround actually benefit teachers and students? Theresa argues that her school’s decline from an A to a B has had an adverse effect. She says

that the ninth and tenth grade Language Arts teachers are being “abused” by her principal, that they “can’t take their English classes” and “two of them are resigning and we’re losing another one in February.” Theresa says that this reaction is partly due to a new supervisor who has no background in their subject and says that instead of teaching students literature, she says, “They just need test prep.” This sounds similar to Sandra’s statement – which I quote in Chapter Five – that her administrators say, “We’re not about mastery, we’re about exposure.” It is also reminiscent of Goldhaber and Hannaway (2004)’s findings – which I discuss in my literature review – that schools narrow curriculum and follow scripted plans to increase scores. This is particularly problematic in schools serving disadvantaged students, which – following NCLB (2001) – more significantly increased math and language arts instruction while limiting time for other subjects (Tienken & Zhao, 2013). While I was unable to speak with any of Thompson’s teachers after their grade rise, I suspect that teachers might argue that the improvement is superficial.

Teacher Perceptions of Student Reactions to Failing Grades

Most teachers agree that, while students are aware of their schools’ grades, they do not necessarily connect them with their own personal performances. When I ask teachers if students are aware of the school grade, Lakisha tells me that: “Everyone I come in contact with are completely aware of it.” Rhonda says that all “the students are quite aware” and that “you’ll hear them talk to teach other about it in the hallway.” Jennifer, their principal, says: “I think all of them are [aware of the grade].” Only their colleague Sandra tells me: “I don’t

know if they know they're an F," but – she adds – they "know that it's a bad school." In this chapter, I mean to examine, through the eyes of teachers, how student knowledge of failing grades makes them feel and how it may impact their academic outcomes.

Teachers have a variety of opinions on this topic. Some think that students are completely unaffected by grades, while others think they are affected, but indirectly – by the atmosphere of the school. Some admit to not knowing. A slight majority believe that students feel discouraged.

Jennifer falls into the first category. She tells me that, while all students are aware of the school's grade, none feel discouraged by it. I ask: then what is their reaction to attending an F school? She responds:

I think that some of them come from elementary schools that may have had low grades, so it's just a fact of – oh, that's what they say and they move on. I don't think that the kids are internalizing that, "I'm at an F school." I don't think the kids are doing that at all.

I ask if she thinks they feel any added inspiration or pressure to perform well. She says, "no."

Jennifer's predecessor – Alex – disagrees. He tells me that students do care about the school's reputation and that, while some blame the school, others blame the students. He explains:

I think the kids are very much aware [of the grade]. Much like the parents – probably where the kids get it from – it's almost like a scapegoat sometimes, like, "Oh, well we're an F school." But, we did an interesting thing this summer, with my mentor. The kids that were in summer school – kids that are struggling in school that failed courses – we actually pulled some mini focus groups with those particular kids and we asked them questions about the school.

One of them was, something along the lines of, “How do we feel about the reputation that we have at this school?” And surprisingly – and these are kind of rougher kids – surprisingly the kids didn’t feel good about the reputation as a whole at this school. And when we probed a little bit more and asked them, “Well, why do you think we have that kind of reputation?” A lot of them indicated that it was the kids. They said that, “There are some kids at this school who don’t care. Some kids at the school who don’t respect the building, who don’t respect the teachers, who don’t respect the rules.” You know, those kinds of things.

And they identified that those kids actually make it more challenging for the school to move forward, because the teachers can’t teach, because the kids are being a classroom disruption or the teachers can’t teach because they’re having to get parents on the phone or do different kinds of things like that. So, I think that the kids are aware of where we are as a school, but they don’t know what it looks like to be a successful school.

Teachers who, like Jennifer, do not think kids are affected by the grade, do say that kids use non-grade indicators like behavior to measure the school. Alex’s experience shows students explicitly linking the school’s reputation with such indicators. I quote him at length because his is the clearest description I hear of students being self-reflective.

Other teachers have heard students discuss their schools in ways that mirror Alex’s story. Sandra, who teaches at Alex’s school, tells me that students use the school’s ranking to tease one another. She says:

They don’t say “F school,” but maybe because they just don’t know the terminology, but they do know that this school is low-ranked and they do say, “That’s Thompson kids.” They’ll say that to each other like, “That’s why you bad kids from Thompson, hah, hah, hah.

Rhonda describes this as acceptable sibling rivalry, but thinks that kids feel defensive if anyone outside of the school speaks badly about it. She explains:

Amongst each other, it’s kinda like siblings. You have siblings, you can bother those siblings, but when it’s someone from the outside that’s bothering your siblings it’s a different story. And so for our students, I believe that amongst each other it may be a

different feeling because they're in it together, but if there's someone on the outside looking in that's trying to say something negatively about the school, then it's perceived in a different way.

I ask Rhonda if the students take the grade self-critically.

Some of the students do see it as a reflection of themselves because they're here at this school, and then you have some students who will almost – in a sense – speak of the school as if it's not a part of their being, or they're not the reason [for the grade]. Like they don't help make up the school.

Collectively, my interviews suggest that students at Thompson Middle School are conscious of the school's reputation, if not its letter grade. They ordinarily act as if this reputation has no impact on them, or discuss it jokingly. However, when pressed or when confronted with people from outside of the school, they admit that the reputation does bother them and point to other students as the cause.

Lakisha articulates this clearly. She says:

[The F grade] makes them feel bad, because they'll say, they've heard other people say before, "Well, you're a student at Thompson," or "Thompson's a low-performance school." It is a stigma attached to being a student at Thompson and I think they somewhat feel a little rejected because of it.

While, Mareka believes the grade makes the students feel badly, she suggests that the "stigma" is larger than the grade. The students feel "rejected" by their association with the school.

I ask Rhonda about Sandra's suggestion that schools should be scored with descriptive phrases like "needs improvement," rather than grades. In response, she tells me:

There was this Ted Talk where the teacher said she was grading a student's paper and this one kid only got two right out of like 20. So, instead of her writing minus 18 on the paper, she wrote plus two. And the kids said, "Well isn't this scale?" And she said, "Well, you got two right that means you know something." So, if that's the way

that we go, I think it would change the mentality of not just the teachers, but also the students, where it's like we know something. But when you see a F, it's like you're failing, like there is almost, in a sense, no hope.

It crushes you as a student when you see a F on your report card or it crushes you when you look at your credit score as an adult, and it's like, "Man, I really jacked that up." So it makes a difference, the way that you word it, just for people's confidence because a lot of times it's like, it's not in the way I appear to be, but it's like – who am I? And if I'm telling you consistently that you're failing, that's what [you are] going to see, and that's what I'm going to think about myself.

Rhonda seems to take issue with the entire system of grading – both schools and individual students. She says that her observations may or may not be unique to the environment of a failing school, or of her particular school.

Some teachers believe that school grades may not influence students at a micro-level, but that they do impact a school's environment, which indirectly affects teachers and students. Jane talks about this. She says that at her current, B school, students "don't really care" about the grade. She says: "I think intellectually they have a sense that how they perform affects the school grade but I don't think they care so much about that." However, she implies that students at her previous, failing school were more affected. She explains:

I distinctly remember the kids [there] not understanding why they could do the work in the classroom but they couldn't do the work and didn't recognize the work on the exams, because we were teaching exceptional ed. kids and we were trying to meet them where they were. We couldn't get them to jump three or four grades and be prepared for a standardized test but we could get them to maybe demonstrate a year's growth in the classroom. They definitely felt frustration there. I've definitely seen kids affected very much by their grades, but not so much – at least directly – by the school grade. It does determine the climate so there's an ineffable affect.

Jane captures something that characterizes most of my interviews: an equivocation of school grades, standardized tests, and teacher performance. They seem to use the words interchangeably. If I ask a question about the school grade, for example, many answer by

discussing standardized tests or their evaluation instead. I conclude that teachers see these – not just as interconnected – but as inseparable components of a single system. In Jane’s view, even if kids do not care directly about the school grade, they are indirectly affected by the policy environment and their teachers’ behavior.

Standardized Tests as Formative and Summative Assessments

In my literature review, I discuss Claude Steele (2010)’s concept of stereotype threat (see also: Carr & Steele, 2009; Lawrence, Marks, & Jackson, 2010; Osborne, 2001; Shih & Pittinsky, 1999; Steele & Aronson, 1995; Van Loo, Boucher, Rydell, & Rydell, 2013; Wout, Danso, Jackson, & Spencer, 2008). While Steele explores stereotypes of race and gender, I wanted to ask teachers about school membership. Do students feel so afraid of confirming the stereotype of their school’s students – as challenged or failing, as “Thompson kids” – that they suffer additional anxiety and worsened performance? Some teachers believe they do; however, most teachers I speak with feel unsure. They have never considered the question before. I hear more evidence of four trends: (1) students may feel an out-sized pressure from teachers to perform; (2) many students feel exhausted by an abundance of different tests; (3) some struggling students would rather not try than accept the stigma of having tried and failed; and (4) teachers and students at failing schools sometimes feel a weariness about testing and stop pushing for success they see as unattainable. I include discussion of these themes here because teachers clearly connect them to the issues of grades and stereotypes.

Teacher Perceptions of Student Responses to Academic Pressure

Some teachers tell me that students naturally feel anxiety to perform well because of testing's high stakes for them personally, if not for their schools. Their scores will eventually determine whether they graduate high school, but even at the middle school level, schools use tests to place students in remedial or advanced classes. Martha explains:

They have test anxiety to begin with because it's ingrained in them, that if they don't pass these tests, they could be held back. It could prevent them from getting certain classes in high school. They're going to stay in intensive math. They're going to stay in intensive reading. So, it gives them test anxiety to take any of these standardized tests.

Martha sees daily evidence of anxiety. She says that students even "have test anxiety when I give them a five-question test." She is not sure whether anxiety makes students – especially students of color – under-perform. She adds that, "this is my first testing experience with a predominantly black school, so I'm curious to see how [students perform]."

All of the teachers I speak with generally agree with Martha – that students naturally feel anxiety with any test – but teachers vary in how much extra pressure they think a failing school grade creates. A majority of teachers and principals in my study – nine of fourteen – believes that while students are aware of school grades, the degree to which each student connects her own performance to her school's depends on the student's age, personality, and teacher. When I taught high school in Florida, our administration encouraged us to explicitly remind students that their performance on the FCAT affected the school and staff. However, of my participants, only Theresa shares this experience.

Theresa says that the extent to which students connect their performance with the school grade, “depends on the instructor,” but she adds:

I also know that instructors are at a point of such frustration that they probably say things in classrooms that are not appropriate out of panic. We have a new teacher [who] they threw into ninth grade ESL class, because why not? She has the children digital only. They’re online all day. She says stuff to them in classes that [our administrator] called on her about. “You guys have to do well on this test, or I’ll lose my job. You have to do well and you don’t historically do well.” That message goes to the kids, and ultimately, that should not be a student’s problem. A student should not have to care about the school letter grade. A student should not have to care about a teacher’s tenure or job security. A student should care about getting to their goal, whatever that goal may be.

Theresa implies that her new colleague feels overwhelmed, is untrained in working with ESL students, and has to rely on “digital only” to teach. As a result, she passes the pressure of her job onto her students.

Even veteran teachers feel that pressure from administration hampers their teaching abilities. Sarah has taught for 18 years and currently works at a C school. She tells me:

There’s not a lot of pressure on me because my scores are good but there’s a lot of pressure on a lot of other teachers. I’ll tell you what it does do to me. For example, we lost a couple of planning periods because they want us to analyze data, tell them what we’re doing with it. So now, because the school grade is a C, and we have a new superintendent who wants all the schools to be As and Bs, the administrators have to report back to her.

Teachers tell me that a lack of planning time and increased paperwork both contribute to their stress and consequently, make it more difficult for them to teach effectively.

Theresa worries most about how this affects students with individual education plans. She explains:

We have more and more kids diagnosed as being on the spectrum. We have more and more kids being diagnosed correctly or incorrectly with ADHD, and we're telling them, "The way that you think is wrong, and you may affect whether your teacher has a job next year." They're being told this by administrators, by guidance counselors. It's part of an overall, top-down, demoralizing process. To put it to my accounting, [our superintendent] is responsible to the governor, who puts pressure on her to deliver results. She has [three assistants] under her. She puts pressure on them to deliver. They have [one person] under them and all he does is visit high schools to shriek at teachers.³⁵

Theresa describes pressure being passed down from the governor, to the superintendent, to assistants, to teachers, and finally to students. She fears that students, especially those already stereotyped as thinking "wrong," face too much of this cumulative pressure.

Some teachers at a failing middle school make the opposite observation: they think teachers need to more explicitly connect the school grade with test scores. Sandra says:

I don't think teachers emphasize that we're an F enough to the kids, and I feel like – I don't know – I don't know if that's a good thing or a bad thing. They just don't. The kids know it's a bad school, like we had two fights today, they know that things aren't good here. I just don't know if they know they're an F. I don't think it affects them whereas, if it was me, and I went to an F school, I would be like, "Mom, take me out."

Sandra thinks that students pay more attention to non-grade indicators, like fights, to gauge school quality, but she implies that if they were more aware of the grade, they might work harder to transfer schools.

Sandra's colleague Rhonda believes students need to understand that their work affects their school and teachers. She explains:

I can't say that they are aware of the fact that their performance affects the school grade, kind of seeing it as, "If I don't perform well the school doesn't perform well."

³⁵ I have removed Theresa's references to particular people.

They're not aware of [that]. And I think it's a matter of teachers making them aware of that, and not in a derogatory way, but just letting them know, "Hey, I can only succeed if you're successful."

Joy – another teacher at Sandra's school – agrees. She stresses that if students do not understand the connection, they will never take responsibility. She says:

They are aware [of the school grade] but then I think they're not, they don't realize how much they're affected. I don't think they realize it. They know, but they think it's everyone else's fault. But they don't realize it's their actual test scores. Even after we harp on them about it...some kids care but for the most part [they are] nonchalant.

Joy feels like her efforts are somewhat futile. Even when teachers pressure students to perform well, the students "think it's everyone else's fault" or feel "nonchalant."

Joy's principal – Jennifer – thinks that this nonchalance is due to the students' age-group. She agrees with her teachers that her middle school students do not care, but tells me that "high schoolers do." She continues:

I think they do [feel pressure] and I think it's different kinds of kids. Like, you have the kids who are academic. They feel it because they believe that [the school grade] has something to do with where they can get into college. They don't want to go to an F school or a D school because they think they won't get into college, even though we know that's not true, that's in their minds. Kids who play sports, they don't want to go to a school that's not seen as academically upstanding. They want all of them. They want to go to someplace that's great so that they can go greater.

Jennifer cannot say whether this pressure affects the way students perform on tests, but she thinks it affects all students: academic students and athletic students both see an F school as a drag on their college applications.

Alex – Jennifer’s predecessor – thinks that middle schoolers do feel pressure, but indirectly. He tells me that sometimes educators pressure students too much and forget to focus on basic things like test preparation. Alex explains:

On the school’s side, we focus so much on curriculum and just push and push and push and push and push the kids, and push and push and push the teachers, that there are some things that can sometimes get lost in it. Even basic testing strategies. You know, we’re busy teaching standards and moving through points and, “This is how you analyze and assess this piece of literature and this is how you write and do this,” but when the kids actually sit for a test, it’s overwhelming and it becomes almost like a do or die.

Alex thinks that pressure on students cannot be “push and push and push,” but must include “basic testing strategies.” Tests cannot be so high-stakes as to feel “do or die” or tests become overwhelming.

Exhaustion as a Contributor to Student Failure

Teachers also believe that the constant pressure and high stakes that Alex talks about eventually lead to exhaustion, and exhausted students perform poorly. This is another way in which school grades might affect students indirectly. When teachers feel pressured to provide more and more test preparation, students become burned out. I ask Lakisha what factors contribute to student stress and harm their performances.

They’ve said they’re tired of testing. They don’t like to hear the word “test.” They feel like there’s no need. And even, they’re like a-motivated, some of them. When they get to a test and they already kinda defeat themselves before they start, because they look at the reading tests and say, “Oh my gosh, this is too much reading.” And

then they don't read. So, necessarily it's just the lack of putting forth the effort. The aptitude is there.

Lakisha thinks students feel tired from all of the testing and practice. Even when students have the aptitude to perform well, they "don't read" and do not put "forth the effort."

Both of Lakisha's principals I speak with echo this belief. Alex tells me that students simply "Christmas-tree" tests – or bubble in random answers, but if he sits down with a student in his office, the student can answer the questions correctly. Lakisha's other principal, Jennifer, calls this a lack of "stamina." I ask her if students are "afraid of failing."

No [I don't think they're afraid]. I think that the stamina is not there for the kinds of questions that we ask and the things that we ask the kids to do. A lot of times when we're doing FSA-type questions, what I see is that the kids are able to answer the first part well, but if there's a second part, they haven't read the second part. They want to be able to skim. They want to be able to look for and regurgitate [answers]. And that's not what's on the test.

Like Alex, Jennifer believes that students at her failing school can answer the test questions, but they want a shortcut to make the test take less time.

Some teachers describe this as "laziness" and say that it is characteristic of students at failing schools. Theresa disagrees. She sees this same phenomenon with her advanced placement students at her B high school. She says:

There is such a resentment about the testing culture that I hear kids in the hallway, I hear kids in my class say, "I just have to go bubble in one answer and then I can walk out." I think that the students are more savvy than we give them credit for, and I think they know it's politics. My students, who are over-tested to death by the time they get to me, they've taken the ACT twice, the SAT twice, countless EOCs – end of course

exams – and multiple AP tests. We have a kid graduating this year who will take eighteen AP exams. They're tested out. They're done.

Instead of laziness, Theresa sees “savvy” and thinks students view tests as “politics.” She continues:

By the time they hit ninth grade, they've done nothing but take tests, and now it's hitting. This started at the beginning of my teaching career in increments, and it goes slowly more normalized, but now 16- and 17-year-olds today, and maybe nationwide, it may not be just Florida, but Florida seems to be leading the trend with testing. 16- and 17-year-olds have taken hundreds of tests, and they've given more class time to test preparation or test execution than actual content. They resent it.

Theresa adds that students who fail tests, and have to re-take them, get trapped in a cycle of failure. The more they fail, the more they have to test and the more they test, the more their exhaustion grows. She adds:

Even that kid that you think hates you and is kind of argumentative will say, “Why do I have to go sit through a U.S. history EOC again? This is crap. I want to be here. I want to be in class. I want to learn.

Sarah also highlights this problem of students constantly torn between class and testing – and possibly falling behind in both. She explains:

I don't know if [students] feel so much teacher frustrations, but I think that their own frustration is testing. They're pulled out of the classroom a lot. I mean, the months of March and April, you can plan very little because they're pulling the kids out all the time. So, I feel really bad for them sometimes. I really do. Especially for those kids who struggle

It may be that such a student is not feeling anxiety because of a stereotype or increased pressure or even high-stakes. Some students may simply feel over-tested and willing to fail in order to finish the test more quickly. Some ambitiously want to return to class and keep learning. They may lack the foresight to realize that failure will only result in more testing. Researchers distinguish between extrinsic and intrinsic motivations. Research has found that extrinsic motivators – such as testing – are associated with anxiety and stress, reduced creativity, inflexible thinking, and an increased likelihood to drop out of school (Alispahić, 2013; Harlen & Crick, 2003; Jones et al., 2007; Schunk, Meece, & Pintrich, 2013).

Several teachers talk about the consequences of student myopia. They tell me that sometimes academically gifted students do not try on tests and as a result are placed in intensive or remedial reading classes. This frustrates students and further disenchant them with school. Even students who might benefit from an extra reading course face burn-out. Rhonda explains that – for her – dance class helps “release some of the endorphins.”

I’m going to have to have that release [of dancing], but we have some students that have no electives at all. And not to say that they need to just play, but it’s like – as a kid – if I’m struggling, if I’m a struggling reader and all my classes are academic, and there is no outlet throughout the day, that’s a problem.

Unfortunately, teachers tell me, an elimination of electives in favor of additional academic classes is part of the turn-around plan for students at many failing schools.

“They’d Rather Not Try”: Student Attempts to Maintain Agency

The most common observation across my interviews is that students would rather not try – on a standardized test or in the classroom – than try and fail. Educators believe that not trying allows a student to view himself as rebellious rather than “dumb.” Alex says that, as a principal, he sometimes pulls disruptive students out of class. These kids “would rather do anything else than math” because they feel anxiety about their abilities. He explains:

[If I] pull them into an office and give them a passage to read, they can do it but with difficulty. Or, “Try these math problems of varying complexity.” And they can do maybe the easier ones, but they can’t do the progressively difficult ones. My thought on it is that, “If I don’t do it, you can’t say that I’m dumb. I just, I chose not to do it. But if I try, and I fail, I’m dumb.”

So it almost creates some anxiety with the kid where they’re like, “Well, I don’t want to try it because I can’t do it and if I can’t do it at that level, then I don’t want to do it at all because I’m not good at it.”

Alex explains that on a test, students start by skipping one hard problem, then “that turns into skipping four or five or six” until the student has skipped most of the test.

Teachers at Alex’s school agree. Joy tells me that students “don’t seem stressed [because] they’re so nonchalant about everything,” but she suspects they use nonchalance as a cover for fear. She says:

I don’t want them to feel defeated on the test. I don’t want them to feel defeated before they can even give a try. And I know they do that often with anything reading. Like, if I put this in front of them and say, “Hey, read this and sign it.” They’ll just put their signature because they feel defeated by reading. They just say, “This is too much to read. Why do I have to read?”

I ask Joy if students skip reading because they are bored or lazy, or because they want to avoid failure. She says, “They want to avoid that failure.” She explains:

Because a lot of them dodge quiz days. They'll dodge quiz days and they'll dodge assessment days just so they don't have to worry about their name being on the board. They'd rather see "N/A" next to their name than an actual score just to keep from having to be the failing student. They just [think] "I'd rather just not do it before I found out that I fail."

Joy's supervisor requires her to post the students' grades on the classroom wall. She tells me that this works for some students who feel competitive, but it discourages others who fear the embarrassment of having their failing grades publicized. Some students go so far as to skip test days altogether.

Sandra shares similar experiences. She tells me that students skip assignments or bubble in random answers and then say, "I didn't try on that, that's why I did bad on it." Miss Charles tries to counter-act the students' fears of failure by emphasizing the positive. She explains:

[Students were taking] an i-Ready test, so if the students did well on the test, it [became] a longer test. So, some of my kids that are, that read slower, I was like, "You know, you must be doing awesome on this because you're taking a longer amount of time," and that kind of boosts their morale. I don't know, you're not really allowed to do that, but I was just kind of encouraging them and I feel like these kids really do need that going into standardized testing.

Several teachers share similar stories with me, stories of students who need encouragement to assuage their anxiety over reading slowly or struggling with a particular question.

Not every teacher at Alex's school agrees. Martha tells me that students "are not afraid of failure. There are some students that are just lazy and they don't want to try because they just don't want to." Similarly, Lakisha says that while students do get anxiety, "they

don't necessarily put forth their best effort" and "sometimes the children are not necessarily as motivated."

Stereotype Threat or Self-Fulfilling Prophecies: The Impacts of Low-Expectations on Student Performance

Contrary to Steele's theory that students fear confirming negative stereotypes, several educators tell me that students and teachers at failing schools can begin to accept those stereotypes as true. Lakisha says: "[Testing] brings out a sense of inferiority within [students] and they feel like, 'Well, you already think that I can't do it, so it's no big deal if I don't perform or do well.'" Steele says that students feel so much anxiety about proving stereotypes true, that they suffer academically. Teachers tell me that some students do not feel anxiety because they have internalized the stereotypes. Alex says:

I think that [a failing grade] makes it okay to be less than what the expectation is. Because, when you're labelled as a failure, then failure becomes almost a norm that exists at the school. Or it's, "okay, well, we're an F." If I were going to a different school, and that was a B school, there are certain expectations that exist. "Oh, we're a B. We have to maintain." But, I think that when you look at a D or an F school type of situation, it makes it easier for a kid to internalize that and say, "Oh, it's okay. We're allowed to act this way. We don't have to do our homework. This is what we do here."

Lakisha echoes Alex's comments on both academic performance and behavior. She says:

I have children who, it's expected for them to not perform. They continue to misbehave, because it's okay. They feel like it goes hand-in-hand. "Who cares? I don't really care about school and so I'll just misbehave, because I'm not doing well in school as it is anyway." 'Cause I have a student who came to me from another reading class and he is a higher-performing student, but when he took the FSA, he didn't put forth the effort and he clearly scored a level one. But then, you get a double-block reading class and he's upset, because he feels like, "I shouldn't have a double block reading. I should have a single block." Well, you don't have that

because you didn't put forth the effort. Granted, he's very smart. He can do the work, but that's just a repetitive pattern of the mindset. You know, "It's expected for me to do a level one."

Even though Lakisha's student is "higher-performing," he only scored a level-one – the lowest level – on the state test. As a result, he was assigned an extra reading class. Lakisha suggests that being in a class for struggling students simply reinforces the expectation that said student is not capable.

These educators believe that this same mentality affects teachers. Sandra thinks that the reputation and environment of the school can reinforce teachers' negative conceptions of students. She says: "Like kids talking during tests, our CRT – our testing coordinator – she literally said last year, 'Oh, if a kid talks during a test that's fine because it's Thompson and you know how they are.'" Alex also complains about the low expectations. He explains:

[When I started at Thompson], I was a 27-year-old, first-year administrator. In the summertime before school started, I called the AP teachers in just to talk with them. I just wanted, "Hey, what happened? Here's your scores. Let's talk about it." And I actually had one teacher come in and the excuse was, "Well, that's good for our school." They had like a four percent pass rate or something. Which was one kid. "Well, that's good for our school." And it almost just becomes okay.

[They said,] "You're new. You're young. You don't understand. This is good for our school. We had one kid, last year no one passed! So we actually did really well." And needless to say, that teacher didn't teach AP the next year, because they didn't believe in the kids. And that's – I think – what happens. You kind of fall into complacency, where it just becomes okay.

Theresa tells me that this mistake – underestimating students – can have a racial dynamic.

She describes a recent faculty meeting where her principal said: "I'm not worried about our white kids, our white kids are doing fine, but our children of color, we've got to step it up."

She calls this "tone-deaf" and says:

Race is real. The kids are very aware of it. The kids feel othered. I think that sometimes when you have people in leadership who don't understand what children carry around with them in their emotional backpack, you can't just throw another test at them. When you say – and everything is by demographic – and you say, “Well, you know, our black kids, our black kids are bringing down our test grade.” What kind of perception does that give these children?

Theresa argues that the “perception” could be outward: teachers do not believe in me; or inward: I do not believe in myself.

Final Thoughts: Chapter Seven

Teachers report a complex array of emotions associated with school grades. Both the media and other researchers have referred to the association with failing schools as a “stigma” (Bowen & Trivitt, 2013; Chakrabarti, 2013). Teachers confirm this association. They describe receiving a failing grade as a “punch in the gut” that seems to ignore all of their efforts and successes. While failure inspires some teachers to work harder, other teachers leave or refuse to work at failing schools in the first place. This is especially true for teachers who teach “core” subjects like math and language arts; these teachers receive a disproportionate pressure from administrations.

Educators at failing schools see their schools represented in the media often – almost exclusively in negative terms. They worry that such coverage scares away talented teachers and students. They want the media to paint a more three-dimensional, realistic portrait of their schools. Such an image would acknowledge that school populations are not monolithic and include students at different levels and with different academic and extra-curricular

specialties. Some positive coverage might mitigate the degree to which teachers must defend their schools and employment choices to family and friends. If coverage is to be negative, teachers want some attention paid to the causes of failure, in order to provide the public with context.

There are instances, including at Alex's school, where the F grade seems to have contributed to success. I question whether this is actually true. In this case, the county significantly intervened, including offering bonus money to attract qualified teachers to the persistently failing school. These policies may have succeeded, but the grade seems to have been unnecessary, especially when considering the overlap between grades and socioeconomic status. Instead of waiting for a school to fail for a decade, teachers imply that counties should intervene in schools of low-socioeconomic status or high-segregation. It is also questionable whether interventions – even if they raise the grade – are genuinely beneficial to teachers and students. As teachers make abundantly clear, they do not believe standardized test scores correlate to actual classroom success.

Teachers have different beliefs about student reactions to grades. Some think that students are not aware of the school's grade. Others think that students are highly aware. This difference may correspond – at least in part – with a student's age (Jones et al., 2007). For students who are aware, it remains unclear how they feel affected – if at all. While teachers are eager to discuss standardized testing, they have not thought about it in terms of stereotype threat, as I initially suspected they might.

According to their teachers, students who perform poorly may be affected by testing

in four ways. As pressure “trickles down” – from the state, to the superintendent, to administrations, to teachers, and finally to students – it may increase already-existing anxiety, for both teachers and students. As teachers feel more anxious, they may teach less effectively, or even say inappropriate things. Teachers disagree about the value of pressure: while some think that students should not feel burdened by their school’s grade or teacher evaluations, others think that students need to understand those connections more.

Some believe that students feel exhausted from an over-abundance of tests. As they take more and more, they might lose stamina, or might come to view tests as unimportant. Some students grow fatigued quickly, especially when tests require long reading passages. These students tend to “Christmas-tree” answers in order to get out of the test as quickly as possible. Some do this without foresight of the consequences for their schools, teachers, or their enrollment in remedial classes. Teachers believe that this is true of a majority of students – that they do not try. A couple of teachers blame laziness, but most tell me that apathy is a way for students to maintain agency. If they try and fail – they feel dumb. If they refuse to try – they feel rebellious, and in control.

Finally, teachers believe that low expectations – especially at the school level – form a kind of self-fulfilling prophecy. Teachers at failing schools can come to expect less of their students. Inappropriate behavior during tests is accepted, as are low scores. Likewise, students can impose negative stereotypes on themselves or others.

Some of my participants discuss more than one of these theories, suggesting that they are not mutually exclusive. Some students might fit into one category or another, while

others might feel affected in multiple ways. Unfortunately, student failure is incredibly difficult to diagnose. Jason – who works at a non-profit dedicated to matching students with state scholarships – tells me this:

I would be very interested to see data on this, like what student's sense of self-worth is at school and how do you figure out if that is occurring because of the F grade and not because of the other conditions in the school that lead to the F grade. The F rated schools tend to have more distractions and more violent crimes occurring in them. Their academic achievement is obviously lower. There's higher concentrations of poverty. Would that be causing depress or lower self-confidence or [would] the F grade? I don't know.

Lakisha thinks that, ultimately, the causes do not matter. She says that sometimes, in a failing school, teachers and students feel “degraded.” She tells me that the only solution she sees is:

You have to develop relationships [with students]. There's this quote: “Children care how much you know when they know how much you care.” And I truly believe that once they realize that you care about them as an individual, they'll begin to work for you, or they'll try and they'll sit there and read the whole entire story when it's time to take the test. And they won't sit there and scan. They'll sit there and really put for the effort, because they know that you care and you work hard to help them. But if they don't feel that from you, then they're not gonna really want to help or work hard for you. It's about relationships, building relationships with your students.

I ask her if those relationships help assuage the feelings of degradation that comes with an F grade.

Correct. Correct. We have to have more of those type of teachers within this school and that's why I felt that I would stay at this school. This is a great opportunity for me, because here I am in a position to do so much to help.

At the end of my interview with Jane, she strikes a similar note. She says that the conversation is depressing and concludes: “I don't know that there is anything realistically hopeful in the current trajectory of education other than just one individual impacting another in some small way.”

CHAPTER 8

TESTING THE IMPACTS OF RECEIVING A FAILING SCHOOL GRADE ON NON-TEST OUTCOMES

Much data suggests that Florida's school grading policies have contributed to a marked improvement in the state's academic performance over the past decade. Moreover, at a micro level, the reception of a failing grade seems to correspond with an improvement in performance for most schools. As I cover at length in the introduction of this dissertation, Florida policymakers – and former Governor Jeb Bush in particular – have touted these improvements as a policy success.

Alarmingly, as the last four chapters make clear, teachers I speak with, as well as journalists and editorialists, paint a contrasting picture. They argue that school grades are divisive, invoke feelings of shame and stigma, drive away many effective instructors, sow

distrust between parents and schools, inaccurately portray schools as monolithic, and fail to adequately inform the public of a school's quality. Teachers argue that many of the subsequent policies – most notably scripted curriculum and unrealistic pressures – do further damage. Additionally, some teachers suggest that failing grades give students an excuse to act out, skip class, and disengage from the very tests Florida uses to gauge quality.

In this quantitative chapter, I seek to explore these contradictory visions. The primary research question is: What effects do the reception of failing grades have on academic outcomes other than standardized test scores? I include both student-level outcomes such as attendance rates, drop-out rates, and graduation rates as well as the educator-level outcome of teacher separation.

Descriptive Statistics

For the purposes of this dissertation, I count both a D and an F as “failing” grades. I do this because both prompt certain policy responses and both seem to carry a negative stigma among teachers and in the press.

Table 4

Breakdown of School Grades by Year

Grade	2010	2011	2012	2013	2014	Total
A	1,521	1,618	1,465	1,008	1,166	6,778
B	669	673	754	819	610	3,525
C	548	523	569	803	886	3,329
D	145	134	217	362	370	1,228
F	41	24	33	101	190	389
I	0	0	0	0	31	31
UG	331	283	217	162	2	995
Total	3,255	3,255	3,255	3,255	3,255	16,275

As shown in Table 4, the numbers of A and B grades declined slightly from 2010 to 2014, while the numbers of C, D, and F grades notably increased.³⁶ These numbers contradict the narrative of success written by the state. The years prior to 2009 demonstrate the opposite trend, with a gradual improvement in grades, however – as I explain earlier in the chapter – 2009 introduced a new grading formula. Since 2014, and as of 2018, the number of D schools has remained relatively unchanged (195), while the number of F schools has declined (35).

Of the five years included, 2012 to 2013 marks the most drastic change. The state introduced new grading policies that year, which schools expected to negatively impact their scores. The state also began a “safety net” policy to prevent schools from dropping more than one letter grade in a year. These changes are unfortunate for analytical purposes. They confuse the meaning of grades both for policy and public perception. While my literature and

³⁶ While schools missing grades or marks entirely in one or more years were excluded, schools with I (incomplete) or UG (ungraded) marks were included in cases where said schools received one or more grades during the same time period (e.g. grades in 2010 and 2011, but UG or I in 2012).

media reviews suggest that most parents were unaware of – or did not understand – the formula change, most teachers were almost certainly informed. The impact this might have on results is unclear.

Changes in other education indicators over the years reflect a more varied picture.

Table 5 shows the changes in the means and standard deviations of the average daily attendance rate, the total drop-out rate, the white graduation rate, the black graduation rate, the Hispanic graduation rate, the Asian graduation rate, and the teacher separation rate.

Table 5

Variables of Interest – Means, Standard Deviations of Schools (N=3255), in Percent, by Year

Year	Average Daily Attendance	Total Drop-Out Rate	White Graduation Rate	Black Graduation Rate	Hispanic Graduation Rate	Asian Graduation Rate	Teacher Separation Rate
2010	94.37 (2.50)	1.41 (1.73)	88.82 (8.97)	82.33 (12.18)	80.69 (13.42)	91.31 (9.12)	5.61 (6.16)
2011	94.59 (2.66)	1.42 (2.32)	79.02 (13.46)	70.63 (15.21)	72.24 (14.96)	86.72 (12.60)	6.21 (6.58)
2012	94.80 (2.56)	1.34 (2.95)	82.08 (11.71)	74.17 (15.55)	76.50 (13.90)	89.88 (10.51)	6.63 (7.13)
2013	94.59 (2.58)	1.20 (1.91)	83.41 (11.59)	76.22 (14.73)	79.59 (12.67)	90.06 (9.52)	7.62 (7.68)
2014	94.55 (2.85)	1.15 (1.98)	85.35 (10.20)	77.76 (14.23)	80.59 (14.42)	91.03 (8.96)	7.96 (7.88)

Attendance rates remained relatively static through the five-year period – the mean hovering around 94.5 percent with a standard deviation between 2.5 and 2.84. Drop-out rates decreased slightly but increased in variability, with a change in means of roughly 1.4 to 1.15 percent and an increase in standard deviation from 1.7 to 1.98 (see Table 6 for these rates

broken down by race). Graduation rates for all four of the largest racial categories dropped between 2010 and 2011, before rebounding slightly, then dropping again in 2014. Black students experienced the largest drop and most fluctuation, ending 2014 with the lowest graduation rate of the four groups. Teacher separation increased steadily – consistent with my interview data – from 5.6 percent in 2010 to 7.96 percent in 2014.

Table 6

Drop-Out Rates – Means, Standard Deviations, in Percent, by Year, by Race

Year	White Drop-Out Rate	Black Drop-Out Rate	Hispanic Drop-Out Rate	Asian Drop-Out Rate
2010	1.43 (2.07)	1.40 (1.85)	1.62 (2.43)	0.79 (1.94)
2011	1.43 (2.33)	1.40 (2.37)	1.42 (1.99)	0.51 (1.33)
2012	1.18 (1.78)	1.46 (3.27)	1.50 (3.90)	0.41 (1.26)
2013	1.21 (2.10)	1.14 (2.20)	1.12 (1.78)	0.38 (1.45)
2014	1.10 (1.91)	1.20 (2.46)	1.27 (2.28)	0.31 (1.12)

While school grades likely changed – at least in part – due to changes in grading formulas and testing policies, these other measures should represent consistent measurement methods. Perhaps these represent a more reliable observation of changes in student outcomes.

In sum, over the five-year period between 2010 and 2014, the number of A schools increased but so did the number of D and F schools. Mean drop-out rates decreased slightly

and mean daily attendance rates held steady. Unfortunately, mean graduation rates declined for all four of the largest racial categories and mean teacher separation rates increased.

Overall, among all years, my various dependent variables show expectedly high degrees of skew. While average daily attendance rates range from a low of 44.88 percent to a high of 100, they are clustered between 89 and 100 percent – as is expected and ideal. Consequently, the data displays a leftward skew, as shown in Appendix A, Figure 1. In order to conform the data to a more normal distribution, I created a complementary variable – absence rate – skewed to the right and then logged it (logabsence).

Drop-out rates – again ideally – skew in the opposite direction, with most concentrated in the zero to 10 percent range as shown in Figure 2 (Appendix A). In order to approach normal distribution, I also logged the drop-out rate variable (logdrop). I did the same with the teacher separation rate variable (logteachersep) since its distribution also skews left, with most falling under 20 percent, as shown in Figure 3 (Appendix A).

Unfortunately, I do not have a total graduation rate that combines racial subgroups. White, black, Hispanic, and Asian graduation rates all skew left. I created an average graduation rate variable (grmean) in order to illustrate this, displayed in Figure 4 (Appendix A). I did not include this average variable in my analysis because it conflates quite disparate rates. For example, the mean white graduation rate (83.7) is significantly higher than the mean black graduation rate (76.2). Consequently, I focus on each racial sub-group in my analysis. In order to approach normal distribution, I created a complementary failure rate

variable for each, then I logged each of the resulting variables (logwhitefailure, logblackfailure, logHispfailure, logAsianfailure).

While none of these variable transformations entirely corrects for skew, it is improved, as shown in Figures 5-11 (Appendix A).

Bivariate Analysis

Among my concerns is whether a failing grade makes a school more or less likely to fail again. While the data shows that the number of D and F schools increased between 2010 and 2014, it is not initially clear whether the same schools repeatedly failed. In order to determine this, as well as the statistical significance of the relationship between a failing grade and a failing grade in the previous year, or the previous two years, I performed a chi-squared test, the results of which appear in Table 7. Each of the 3,255 schools in my study sample is observed for five years, producing 16,275 school-year grade assignments. There are 1,253 such school-years in which a school received a D or F grade, in any particular year between 2009-2013.³⁷ These could represent 1,253 different schools, each receiving a D or F grade in only one year, or a smaller number of schools receiving a D or F grade more than once over the five years.

³⁷ Note that there are 3,255 schools in the study sample with outcomes observed between 2010 and 2014. In order to test whether the chance of receiving a D or F grade in any of those years was affected by the grade the school received in the previous year, 2009 grade data were added. For the 3,255 schools which received a grade in 2010, year t-1 is 2009. For 2011, year t-1 is 2010, and so forth.

Table 7

Chi-Squared Test of the Association between a Failing Grade in Year t-1 And Year t

Fail Grade in Year t	Fail Grade in Year t-1		Total
	No	Yes	
No	14,020 95.65 %	638 50.92 %	14,658 90.06 %
Yes	1,002 6.67 %	615 49.08 %	1,617 9.94 %
Total	15,022 100 %	1,253 100 %	16,275 100 %
Pearson chi2(1) = 2.3e+03			Pr = 0.0

We can see that among the 1,253 situations in which a school received a D or F grade in some year between 2009-2013, 615, or roughly 49 percent, of said schools also received failing grades in the following year. This relationship does appear statistically significant ($p \leq 0.0001$), meaning I can reject the null-hypothesis that these variables are independent of one another. If there was no association between D and F grades received in successive years, then about 10 percent of schools which received a D or F grade in any year would be expected to receive a grade lower than C in the following year. These results suggest that if a school received a D or F grade in one year, there is nearly a 50 percent chance that it also received grade lower than C in the following year.

I ran the same test to determine the relationship between a failing grade and another failing grade two years later. In this case, Year t-2 and Year t ranged from 2008-2012 and Year 2010-2014, respectively. There were 961 situations where a school received a D or F grade in some year and also received a grade lower than C two years later. Results of a test

of independence between a school grade received in a given year and two years later are shown in Table 8.

Table 8

Chi-Squared Test of the Association between a Failing Grade in Year t-2 and Year t

Fail Grade in Year t	Fail Grade in Year t-2		Total
	No	Yes	
No	14,078 91.93 %	580 60.35 %	14,658 90.06 %
Yes	1,236 8.07 %	381 39.65 %	1,617 9.94 %
Total	15,314 100 %	961 100 %	16,275 100 %
Pearson chi2(1) = 1.0e+03 Pr = 0.0			

We can see that, among the 961 situations in which a school received a D or F grade in some year between 2008-2012, in 381 of them, or roughly 40 percent of all such situations, said schools also received failing grades two years later. This result is also statistically significant ($p \leq 0.0001$) suggesting that I can reject the null hypothesis and conclude that these variables are not independent of each other. While it is positive that the percentage of schools that repeat failure within three years is smaller than the percentage failing two years in a row, 49 and 40 percent still suggest an alarming pattern of failure. If 40 percent of schools persistently fail then the state's argument that all schools can be spurred into improvement via grades seems doubtful.

In order to further explore the variables in the dataset and their relationships, as well as finalize a fixed-effects regression model, I ran a correlation test (results shown in Table 9).

Table 9

Correlation Matrix (Coefficient, R-Value, Number of Observations)

Variable	Fail Grade	Fail-Prior	Fail-2 Years Prior	ADA Rate	Total Drop-Out Rate	White GR	Black GR	Hispanic GR	Asian GR	Teacher Separation Rate
Fail Grade	1.0 16,275									
Fail-Prior	0.38 0.0* 16,275	1.0 16,275								
Fail-2 Years Prior	0.25 0.0* 16,275	0.34 0.0* 16,275	1.0 16,275							
ADA Rate	-0.07 0.0* 15,196	-0.09 0.0* 15,196	-0.10 0.0* 15,196	1.0 15,196						
Total Drop-Out Rate	0.12 0.0* 2,320	0.13 0.0* 2,320	0.13 0.0* 2,320	-0.39 0.0* 2,282	1.0 2,320					
White GR	-0.16 0.0* 1,916	-0.13 0.0* 1,916	-0.24 0.0* 1,916	0.33 0.0* 1,901	-0.54 0.0* 1,911	1.0 1,916				
Black GR	-0.13 0.0* 1,750	-0.08 0.0* 1,750	-0.18 0.0* 1,750	0.41 0.0* 1,734	-0.50 0.0* 1,744	0.70 0.0* 1,663	1.0 1,750			
Hispanic GR	-0.19 0.0* 1,695	-0.15 0.0* 1,695	-0.25 0.0* 1,695	0.37 0.0* 1,681	-0.55 0.0* 1,692	0.75 0.0* 1,633	0.73 0.0* 1,568	1.0 1,695		
Asian GR	-0.09 0.02* 791	-0.10 0.01* 791	-0.16 0.0* 791	0.19 0.0* 784	-0.30 0.0* 789	0.48 0.0* 790	0.43 0.0* 774	0.51 0.0* 781	1.0 791	
Teacher Separation Rate	0.10 0.0* 15,183	0.07 0.0* 15,183	0.09 0.0* 15,183	-0.02 0.02* 14,828	0.01 0.55 2,294	-0.10 0.0* 1,902	-0.01 0.60 1,740	-0.05 0.06 1,688	0.01 0.77 791	1.0 15,183

*r-value significant at the P<0.05 level

The correlation matrix shows significant and positive correlations between the three failing grade variables, student drop-out rates and teacher separation rates. It shows negative correlations between the failing grade variables and attendance and graduation rates. As

expected, based on my literature review and qualitative data, this suggests that the reception of a failing grade is associated with predicted increases in drop-outs and teacher separations and predicted decreases in attendance and graduation rates.

Not only do the coefficients show the relationships I expected, there is a lagged effect, meaning that the results worsen between the fail and fail-prior year results. I suspected that – given the late publication of grades in summer and fall – the effects of a failing grade might be delayed. The correlation coefficients between drop-out rates and the fail-grade variables, for example, rise slightly from 0.12 to 0.13. All four graduation rate coefficients fall between fail and fail-prior, before rising with two-years prior. This suggests that drop out rates may increase and graduation rates may decrease more significantly in the year after a school fails. If, following a failing grade, schools successfully improve policies and performance, then I do not believe we would see this lagged effect, instead coefficients would increase or decrease in a desirable direction.

Table 10

Correlation Matrix Including Alternative, Logged Variables

Variable	Fail	Fail-Prior	Fail-2 Years Prior	Log Absence	Log Drop-Out	Log White FR	Log Black FR	Log Hispanic FR	Log Asian FR	Log Teacher Separation
Fail	1.0 16,275									
Fail-Prior	0.38 0.0* 16,275	1.0 16,275								
Fail-2 Years Prior	0.25 0.0* 16,275	0.34 0.0* 16,275	1.0 16,275							
Log Absence	0.06 0.0* 15,100	0.08 0.0* 15,100	0.08 0.0* 15,100	1.0 15,100						
Log Drop-Out	0.16 0.0* 2,052	0.20 0.0* 2,052	0.19 0.0* 2,052	0.08 0.0* 2,020	1.0 2,052					
Log White FR	0.14 0.0* 1,854	0.11 0.0* 1,854	0.22 0.0* 1,854	0.24 0.0* 1,835	0.48 0.0* 1,688	1.0 1,854				
Log Black FR	0.14 0.0* 1,690	0.09 0.0* 1,690	0.19 0.0* 1,690	0.19 0.0* 1,672	0.41 0.0* 1,555	0.65 0.0* 1,589	1.0 1,690			
Log Hispanic FR	0.17 0.0* 1,616	0.15 0.0* 1,616	0.24 0.0* 1,616	0.19 0.0* 1,601	0.47 0.0* 1,504	0.68 0.0* 1,540	0.66 0.0* 1,478	1.0 1,616		
Log Asian FR	0.07 0.10 587	0.11 0.01* 587	0.12 0.0* 587	0.15 0.0* 581	0.32 0.0* 531	0.49 0.0* 582	0.41 0.0* 575	0.44 0.0* 573	1.0 587	
Log Teacher Separation	0.11 0.0* 12,293	0.10 0.0* 12,293	0.09 0.0* 12,293	0.03 0.1* 12,072	-0.02 0.39 1,795	0.07 0.0* 1,662	0.07 0.0* 1,536	0.06 0.02* 1,469	0.06 0.15 541	1.0 12,293

*r-value significant at the P<0.05 level

Table 10 shows a new correlation matrix following transformations in the attendance rate, drop-out rate, and graduation rate variables. The transformations have not led to any extreme changes. The coefficient signs have flipped – as expected – for all of the flipped

variables. Otherwise, we see similar significance and predicted associations, with one interesting exception. Now, it appears that graduation rates (flipped to failure rates) have a stronger correlation with the fail-two years prior variable than the fail-prior variable. In other words, while the first matrix suggests that graduation rates might worsen and then improve in the years following a failing grade, this matrix suggests that less students fail in the year after a failing grade, but then failure rates increase sharply the following year.

Multivariate Results

I ran seven sets of fixed-effects regression models with seven different dependent variables: teacher separation rates, attendance rates, drop out rates, white graduation rates, black graduation rates, Hispanic graduation rates, and Asian graduation rates.³⁸ As previously discussed, I flipped the attendance and graduation rate variables – to become absence and failure rates – in order to correct for skew. They, along with the other variables, were logged in order to conform to more natural distributions.³⁹ I utilized Stata’s cluster option in order to address autocorrelation among the model errors due to repeated observations.

In each model I include year-dummy variables, the three fail-grade dummy variables, as well as the total number of teachers. The year dummy variables are intended to account for common secular factors that could affect the school outcomes of all schools in the state. I

³⁸ Florida did not publish a total graduation rate during the years 2010-2014. In each of my models, I include white, black, Hispanic, and Asian graduation rates as possible explanatory variables or as dependent variables. These four were chosen based on their population share in Florida schools.

³⁹ See Appendix F for a table of sample means, standard deviations, minimum and maximum measurements for all variables, included log-transformed variables.

include the total number of teachers as a rough indicator of the size of a school. Since Florida's class sizes are limited by constitutional amendment, this should be a uniform estimate of size. I also report results of each regression after adding several interaction terms. In order to test whether failing in consecutive years has any compounding effect, I interact the fail-grade variable and the fail-prior variable (Fail-Interaction), as well as all three failure variables (Fail-3 Years). I also questioned whether failing in the 2012-2013 and 2013-2014 academic years affected schools differently than in prior years, due to the grading formula changes and increased number of failures in those years. Thus, I interact the fail-grade variable with a dummy variable representing the 2012-2013 and 2013-2014 school years (Fail-2013-2014), as well as the fail-prior variable with a dummy variable representing 2013-2014 (Fail-Prior-2014).

This section includes results of all seven regressions, with brief analyses, followed by a section of comprehensive analysis and discussion.

Teacher Separation Rates

I suspected that – of all of my dependent variables – teacher separation rates might be most affected by the reception of failing school grades. My literature review suggests that teachers seek to avoid, or leave, stigmatized and underperforming schools, and my qualitative findings largely support that suggestion, with principals lamenting staff turnover. This regression tests the null hypothesis that teacher separation rates remain the same

regardless of grade reception, against the alternative hypothesis that teacher separation rates change after the reception of a failing grade.

The results of four fixed-effects regression models are shown in Table 11.

Table 11

Results of Fixed-Effects Regression Models with Dependent Variable – Log Teacher Separation

	Model 1			Model 2			Model 3			Model 4		
Variable	CE	T	P> T	CE	T	P> T	CE	T	P> T	CE	T	P> T
Log Teacher Separation (Dependent)												
Year 11	.047	2.83	0.005	.038	2.26	0.024	.037	2.24	0.025	.037	2.22	0.026
Year 12	.113	6.72	0.0	.108	6.42	0.0	.108	6.41	0.0	.109	6.46	0.0
Year 13	.215	12.63	0.0	.207	12.14	0.0	.207	12.09	0.0	.201	11.48	0.0
Year 14	.287	16.65	0.0	.281	16.16	0.0	.280	16.12	0.0	.276	15.22	0.0
Fail-Grade	.025	1.05	0.295	.027	1.13	0.258	.015	0.54	0.587	-.029	-0.76	0.445
Fail-Prior	.051	2.07	0.039	.043	1.73	0.084	.024	0.77	0.441	.022	0.63	0.527
Fail-2 Years Prior	.006	0.23	0.816	.011	0.42	0.678	.010	0.34	0.733	.004	0.15	0.88
Total Teachers				.010	6.52	0.0	.010	6.52	0.0	.010	6.53	0.0
Fail-Interaction*							.047	0.90	0.367	.039	0.75	0.455
Fail-3 Years*							-0.0	-0.01	0.996	-.001	-0.02	0.984
Fail-2013-2014*										.070	1.67	0.096
Fail-Prior-2014*										-.012	-0.26	0.794
Constant	1.719	146.68	0.0	1.124	12.23	0.0	1.125	12.23	0.0	1.126	12.23	0.0
R-Square	.046			.056			.056			.056		

*Interaction terms: Fail-Interaction = (Fail-Grade*Fail-Prior); Fail-3 Years = (Fail-Grade*Fail-Prior*Fail-2 Years Prior); Fail-2013-2014 = (Fail-Grade*Dummy for 2013 or 2014); Fail-Prior-2014 = (Fail-2 Years Prior*2014)

All four models include the four year-fixed effects dummy variables and three fail-grade dummy variables (as well as dummy variables for schools implicit in a fixed-effects model).

Model One includes only the year and failure variables. Model Two additionally controls for the total number of teachers, the best available measure I have for the size of the school. While Model Two decreases the significance of the fail-prior variable, it marginally improves the overall R-square value. Model Three and Model Four include the interaction terms – meant to test the impact of repeated fail grades and failure under the post-2013 grading formula. These do not appear to add any value to the models; consequently, I will focus my analysis on Model Two.

The dummies for year-fixed effects (2011-2014) show significance.⁴⁰ Each has a P-value at or close to zero, with T-statistics above the 95 percent confidence level. The coefficients increase from 0.038 for 2011 to 0.281 in 2014.

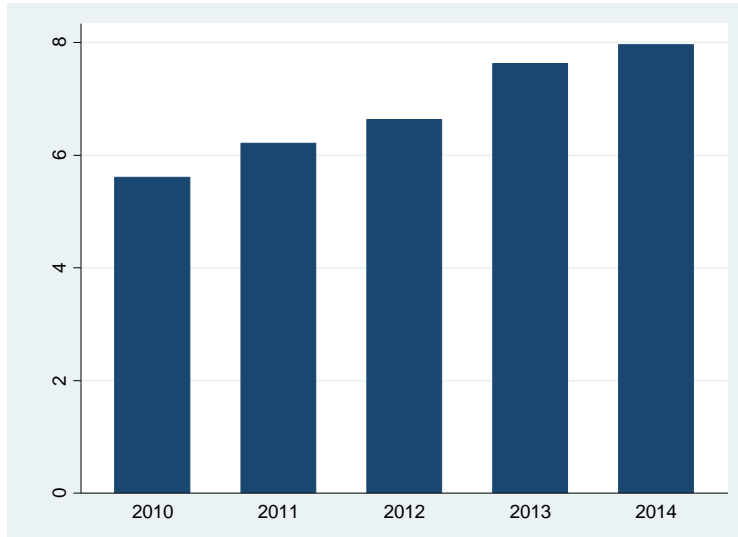
This is a notable change. In order to interpret the predicted effects, I took the positive exponentials of each coefficient.⁴¹ We can say that the year 2011 is associated with a predicted 3.87 percent increase in teacher separation relative to the omitted year 2010, while 2012 is associated with a predicted 11.41 percent increase, 2013 with a predicted 23 percent increase, and 2014 with a predicted 32.45 percent increase. I confirmed this trend by graphing the separation rates by year. Figure 2 demonstrates a steady increase in rates each year between 2010 and 2014. This suggests that a year-to-year change other than grades is having a significant effect. I cannot say whether these changes are related more to factors endogenous or exogenous to education.

⁴⁰ I omitted 2010 as a reference year, thus coefficients are interpreted as relative to 2010.

⁴¹ Marginal effects = $100 \times [(\exp^{\text{coefficient}}) - 1]$

Figure 2

Mean Teacher Separation Rate by Year



While the three failure variables are not significant ($p < 0.05$), the fail-prior variable comes close with a p-value of 0.084 and a coefficient of 0.043. This is a conservative estimation. If we instead use a less conservative one-tail test, the p-value would be 0.042 and thus significant. This provides modest evidence that in the year after a school receives a failing grade, we might expect a roughly four percent increase in teacher separation rates.

These results are surprising. How to reconcile them with the statements of teachers and principals that failing schools have trouble attracting and retaining teachers? Of course, the regression lacks possible explanatory variables. Income, race, discipline, and administration support – which also come up in interviews – may have more to do with teacher separation than a school's grade.

Other possible explanations include: first, while a failing grade might repel some teachers, it attracts others. Several teachers I speak with tell me that they intentionally came to a struggling school. Second, the regression does not include the experience levels of teachers. I considered whether teacher experience level is a confounding variable. Are new teachers more likely to leave and to work at failing schools? While I do not have data on the overall experience level of each school's teachers, I do have the combined years of experience of all separated teachers per school, per year.⁴²

Fixed-effects regression models use within-school variation to estimate the effects of receiving a failing grade. That is, they control for factors internal to schools that we assume are time-invariant. We can also analyze between-school variation. While doing so can tell us much about trends across schools, we cannot isolate any one factor as potentially causal. In order to try and determine whether departures and experience levels coincided with school grades, I graphed mean teacher separation rates and experience levels of all schools in my sample by school grade. Figure 3 shows that F schools have the highest mean separation rate (10.81 percent), followed by D schools (8.41 percent), C schools (7.5 percent), B schools (5.88 percent), and A schools (5.88 percent). This confirms that teachers are more likely to depart F schools than their counterparts in schools assigned higher grades.

⁴² Florida does not publish teacher experience levels by school. It does release teacher salary levels by county, but not by school.

Figure 3

Mean Teacher Separation Rates by School Grade

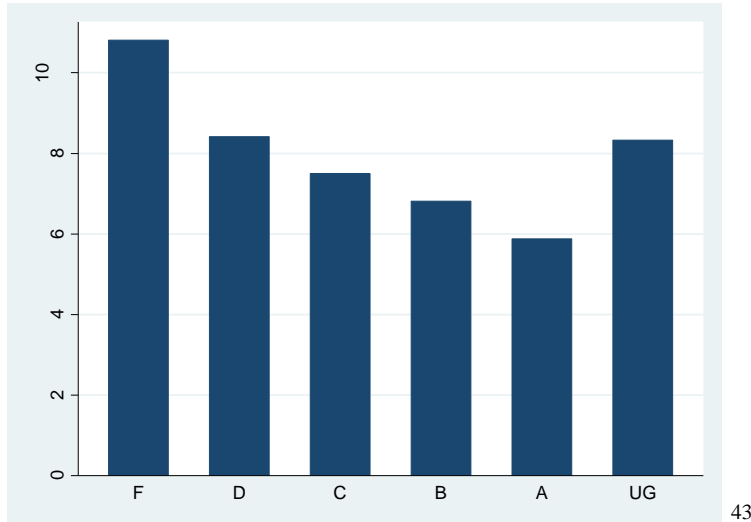
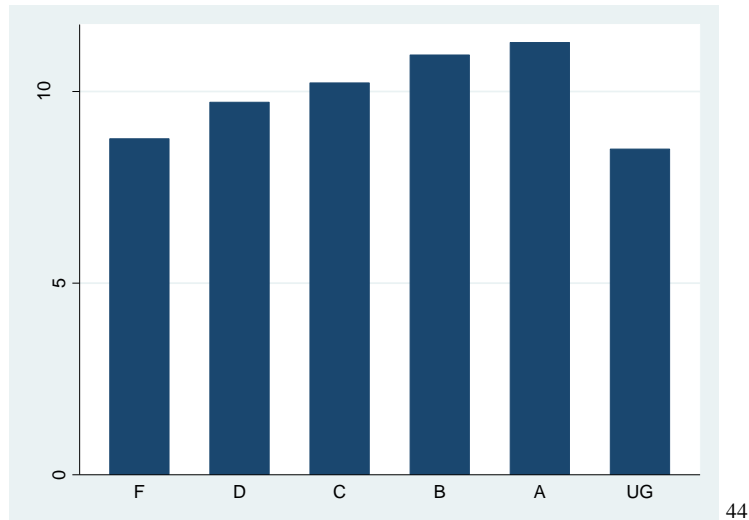


Figure 14 shows the mean-combined years of experience for separated teachers by school grade. Unfortunately, the variable does not allow us to know the experience of each departing teacher, but it may provide some useful insight. Teachers departing F schools have the least mean-combined experience (8.77 years), followed by D schools (9.72 years), C schools (10.23 years), B schools (11.28 years), and A schools (11.28 years). One possible explanation is that A schools simply hire more experienced teachers, and less new ones, and thus departing teachers show more cumulative years. A second possible explanation is that both A and F schools hire new teachers, but those teachers stay at A schools for longer into their careers, while teachers leave F schools quickly.

⁴³ Differences found to be significant ($p < 0.05$) using a Scheffe multiple comparisons test; for complete test results, see Appendix B.

Figure 4

Mean Combined-Years-Experience for Separated Teachers, by School Grade



The data in Figures 3 and 4 demonstrates that – on average – a higher percentage of teachers depart F schools and that those departing teachers have less cumulative experience than other schools’ departures. This could be due to initial hiring practices or to teacher decisions. While the regression provides moderate support ($p=0.08$) that a failing grade, or policy responses to the grade, motivates additional teachers to leave, I cannot say why more teachers depart failing schools than non-failing schools. In interviews, teachers suggest that administrative support, behavioral issues, poverty, and county and state intervention all play a role, while changing grade formulas, shifting expectations of teachers, and other policy upheavals can motivate people to leave the profession altogether. These factors may help

⁴⁴ Differences found to be significant ($p<0.05$) using a Scheffe multiple comparisons test; for complete test results, see Appendix C.

explain why the regression results show an expected increase in the likelihood of teacher departures each year between 2010 and 2014.

Drop-Out Rates

My qualitative data and literature review suggest two possible impacts of a failing school grade on drop-out rates. First, teachers describe a failing school grade as one cause of hostility from students and their parents. We might imagine a student thinking: “This school is terrible so I might as well drop out. I won’t learn anything anyway.” Second, teachers describe a practice of encouraging students to transfer to charter schools. While a transfer would not count as a drop out, students might feel discouraged by a guidance counselor telling them they cannot meet graduation requirements. Such a student might think: “If it’s impossible for me to graduate here, I’ll just drop out. I don’t want to change schools.” If either of these scenarios is true, we should see an increase in drop outs following the reception of a school grade. This regression tests the null hypothesis drop-out rates will remain similar before and after the reception of a failing grade, against the alternative hypothesis that drop-out rates change following the school’s reception of a failing grade.

The results of four fixed-effects regressions are shown in Table 12.

Table 12

Results of Fixed-Effects Regression Models with Dependent Variable – Log Drop-Out Rate

	Model 1			Model 2			Model 3			Model 4		
Variable	CE	T	P> T 	CE	T	P> T 	CE	T	P> T 	CE	T	P> T
Log Drop-Out Rate (Dependent)												
Year 11	-.065	-1.49	0.136	-.068	-1.54	0.123	-.066	-1.47	0.142	-.070	-1.57	0.117
Year 12	-.123	-2.35	0.019	-.130	-2.45	0.015	-.128	-2.40	0.017	-.132	-2.49	0.013
Year 13	-.186	-3.42	0.001	-.192	-3.50	0.001	-.190	-3.47	0.001	-.187	-3.30	0.001
Year 14	-.310	-5.66	0.0	-.317	-5.76	0.0	-.316	-5.76	0.0	-.326	-5.78	0.0
Fail-Grade	-.014	-0.19	0.851	-.013	-0.18	0.861	-.009	-0.09	0.929	.069	0.52	0.602
Fail-Prior	.057	0.89	0.373	.055	0.86	0.390	.063	0.83	0.408	.051	0.66	0.509
Fail-2 Years Prior	.035	0.59	0.555	.032	0.55	0.584	.008	0.13	0.899	.008	0.12	0.902
Total Teachers				-.002	-0.66	0.511	-.002	-0.66	0.509	-.001	-0.63	0.530
Fail-Interaction							-.179	-1.32	0.187	-.257	-1.54	0.124
Fail-3 Years							.273	1.53	0.126	.255	1.23	0.221
Fail-2013-2014										-.156	-1.02	0.309
Fail-Prior-2014										.304	1.37	0.173
Constant	-.111	-2.72	0.007	.031	0.14	0.887	.033	0.15	0.880	.029	0.13	0.896
R-Square	0.037			0.038			0.039			0.041		

Once again, Model One includes only the year-fixed-effects dummy variables and the failure dummy variables, while Model Two controls for total number of teachers and Models Three and Four add in the interaction terms. As in the previous section, the interaction terms do not appear to add any value to the model. In fact, except for the year dummies – which have relatively stable p-values and coefficients in all four models – none of the other variables appear significant.

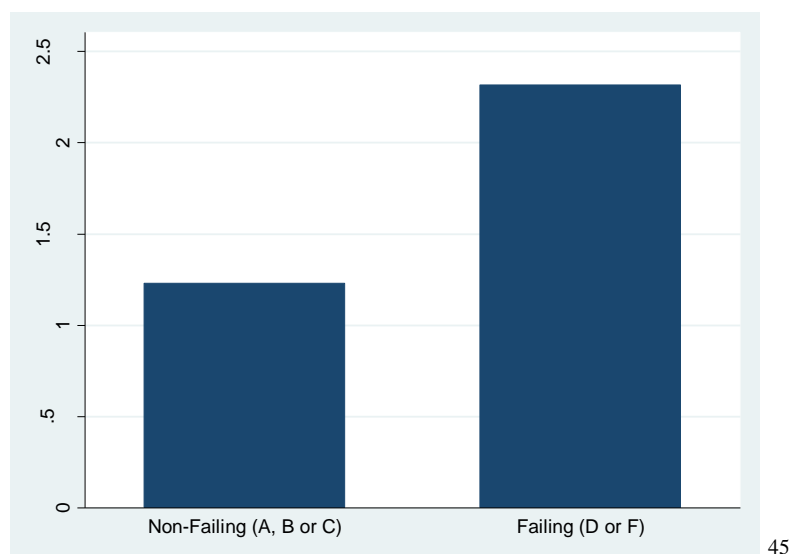
The results do not provide sufficient evidence to reject the null hypothesis that fail grades have no effect on school drop-out rates. The fail, fail-prior, and fail-2 years prior variables are not significant. It is especially surprising that there is no significant association between the drop-out rate and a current failing grade. Given that graduation rates factor into school grades we might expect a drop in graduating students; however, the coefficient for the fail-grade variable is negative, actually suggesting a decrease in drop-out rates during a failing year.

One possible explanation is that drop-out rates are somehow incongruous with graduation rates. The Florida Department of Education suggests as much in its 2016-2017 report on drop-out rates. The Department explains that it “used to publish a single-year drop-out rate” but has discontinued the practice because the score “often caused confusion because it was not comparable to the graduation rate.” The report continues: “[The drop-out rate] represented the percentage of high school students who dropped out in any one year and was not an inverse of the graduation rate” and “a drop-out is defined as a student who withdraws from school for any of several reasons without transferring to another school, home education program, or adult education program” (Florida Department of Education, 2018). This disconnect may help explain why there appears to be no significant relationship between drop-out rates and current grades; although, in my correlation matrix, the various failure rates (the inverse of graduation rates) and the drop-out rate show positive correlations.

As in the previous section, we can turn to a between-schools analysis to try and further explore relative drop-out rates. Figure 5 displays the drop-out rates of failing schools as compared to non-failing schools.

Figure 5

Relative Mean Drop-Out Rates

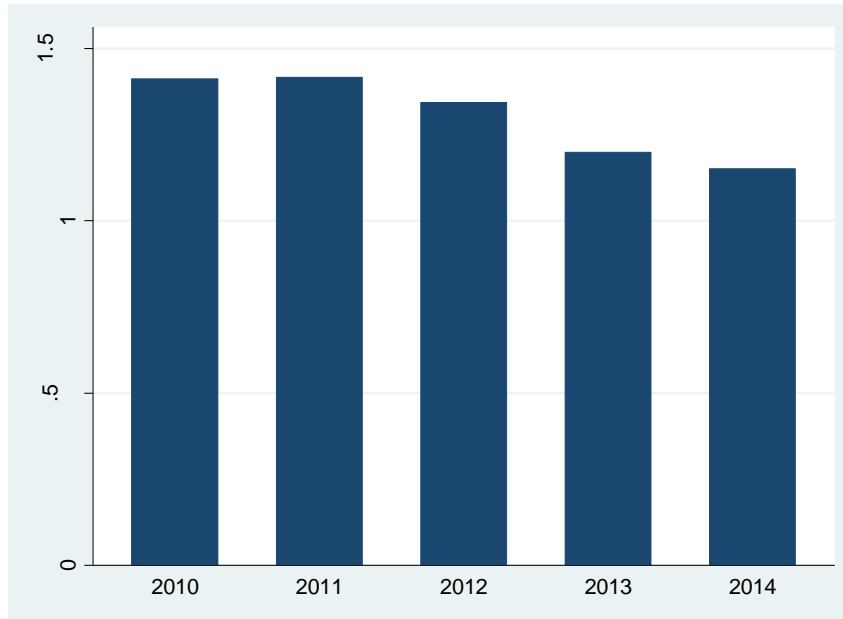


The results show that failing schools have – on average – drop-out rates near 2.3 percent, while non-failing schools have about half that, near 1.2 percent. It is a positive sign that both of these mean rates are very low. In another positive sign, the coefficients for each year-dummy variable are increasingly negative, suggesting that – relative to 2010 – drop-out rates increasingly fell in 2011, 2012, 2013 and 2014. As shown in Figure 6, this was the case. Mean drop-out rates for all schools fell from 1.4 percent in 2010 to 1.15 percent in 2014, with the biggest drop happening between 2012 and 2013.

⁴⁵ Differences found to be significant ($p < 0.05$) using a Scheffe multiple comparisons test; for complete test results, see Appendix D.

Figure 6

Mean Drop-Out Rates for All Schools by Year (2010-2014)



While this decrease in rates is undeniably a positive sign, there remains a couple of caveats. First, as I discuss extensively in Chapter Five, some teachers and media reports suggest that schools are increasingly “hiding” drop-outs by encouraging students to transfer to GED programs or charter schools, without following up. If true, this may mean that the – admittedly small – drop in rates can be partly attributed to students transferring to programs from which a drop-out would not be measured.

Secondly, we should acknowledge that the decrease in drop-out rates is not standard across all racial subgroups of students. Figures 7 and 8 illustrate the break-down in rates across the four largest racial categories. Overall, between 2010 and 2014, Asian students show the lowest drop-out rate by far, while black and Hispanic students drop-out at marginally higher rates than white students. When broken down by year (Figure 8), we see

that Asian students experienced a steady decrease in rates between 2010 and 2014. White, black and Hispanic student drop-out rates also saw a net-decrease from 2010 to 2014, but in the intervening years they fluctuated, with black and Hispanic students notably increasing their 2014 rates in comparison with the prior year.

Figure 7

Mean Drop-Out Rates (2010-2014) by Racial Subgroup

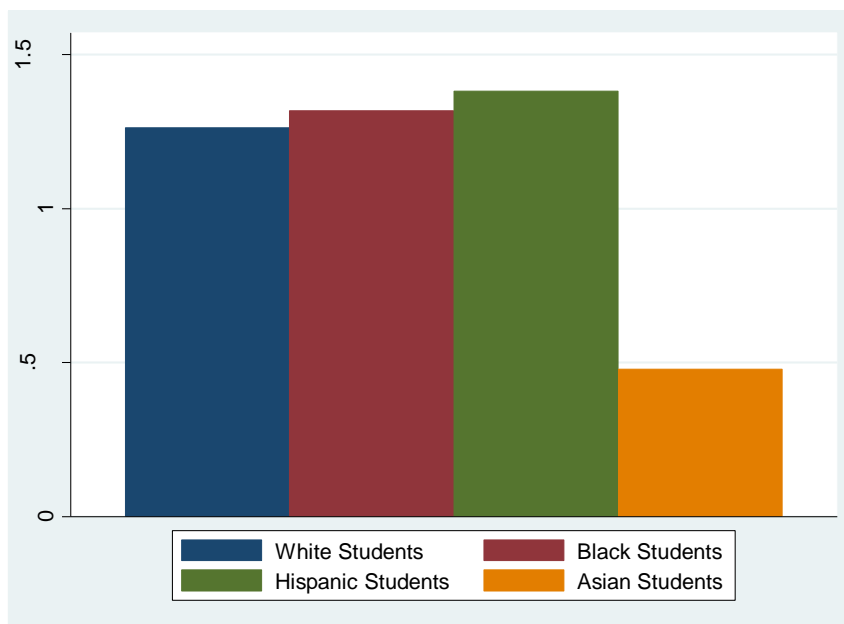
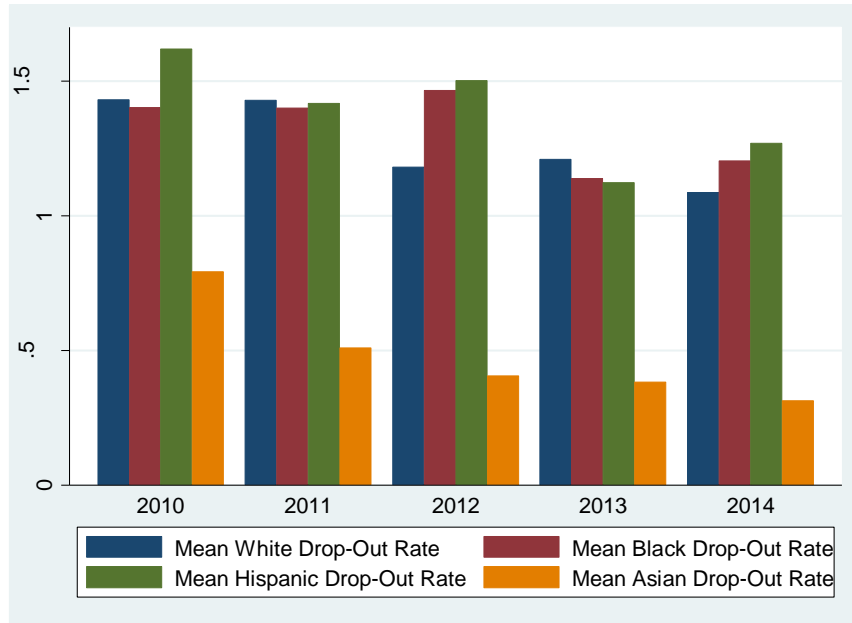


Figure 8

Mean Drop-Out Rates by Racial Subgroup, by Year



Despite these caveats, the news on drop-out rates for the sample years is positive.

Florida has a relatively low number of students drop-out of school and saw that number decrease further between 2010 and 2014. I should note that a lower number of drop-outs does not necessarily coincide with a higher number of graduates. I discuss graduation rates later in this chapter.

Absence Rates

While my literature review and qualitative data suggest that absence rates (the inverse of attendance rates) have a significant impact on student learning, teachers do not suggest it as a dependent variable. That is, they doubt whether the school grade plays any role in

students deciding to attend or skip school. I wanted to test a model in case a hostile or apathetic relationship with the school, a fear of testing, or diffused discouragement may prompt students to stay home. This regression tests the null hypothesis that students are similarly likely to attend school regardless of school grade reception, against the alternative hypothesis that student absence rates change following the school's reception of a failing grade.

The results of five fixed-effects regression models are shown in Table 13.

Table 13

Results of Fixed-Effects Regression Models with Dependent Variable – Log Absence Rate

	Model 1			Model 2			Model 3			Model 4		
Variable	CE	T	P> T	CE	T	P> T	CE	T	P> T	CE	T	P> T
Log Absence Rate (Dependent)												
Year 11	-.045	-8.32	0.0	-.043	-7.88	0.0	-.043	-7.89	0.0	-.043	-7.93	0.0
Year 12	-.089	-15.18	0.0	-.086	-15.48	0.0	-.086	-15.47	0.0	-.085	-15.38	0.0
Year 13	-.053	-6.98	0.0	-.050	-6.96	0.0	-.050	-6.97	0.0	-.053	-7.15	0.0
Year 14	-.045	-6.25	0.0	-.044	-6.38	0.0	-.044	-6.36	0.0	-.044	-6.25	0.0
Fail-Grade	.012	1.20	0.232	.013	1.35	0.177	.018	1.54	0.123	-.004	-0.32	0.747
Fail-Prior	.020	1.90	0.058	.013	1.35	0.177	.020	1.57	0.116	.024	1.69	0.090
Fail-2 Years Prior	-.011	-0.78	0.437	-.017	-1.27	0.205	-.009	-0.62	0.533	-.011	-0.73	0.465
Total Teachers				.001	3.07	0.002	.001	3.08	0.002	.001	3.10	0.002
Fail-Interaction							-.007	-0.33	0.745	-.010	-0.48	0.633
Fail-3 Years							-.038	-1.22	0.221	-.036	-1.16	0.248
Fail-2013-2014										.038	2.22	0.026
Fail-Prior-2014										-.023	-1.16	0.244
Constant	1.652	388.38	0.0	1.579	61.58	0.0	1.579	0.0	0.0	1.580	61.63	0.0
R-Square	.016			0.018			0.018			0.019		

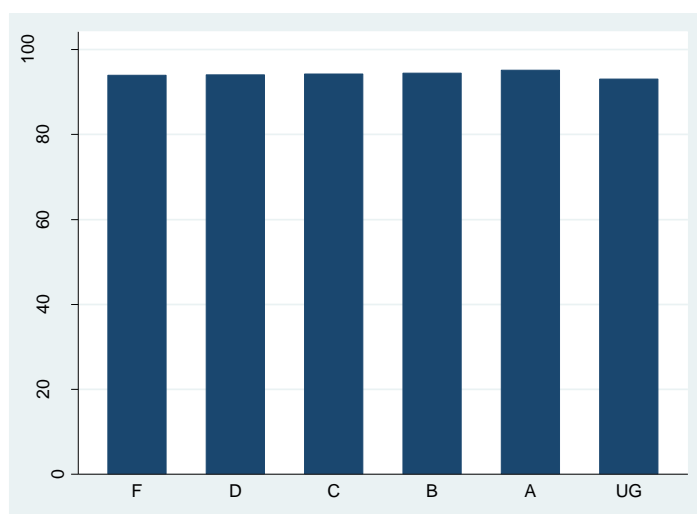
The results seem to confirm teachers' ambivalence towards the alternative hypothesis. The only variables that consistently appear significant are the year-fixed-effects dummies, but the coefficients are weak and relatively consistent, suggesting that attendance rates remained static between 2010 and 2014. A check of the data confirms this, with the mean attendance

rate for all schools hovering around 95 percent in each school year. The total teachers variable is significant in the last three models, but the coefficient suggests only a negligible 0.1 percent impact.

The model does not provide sufficient evidence to reject the null hypothesis. It is unclear – but ripe for future research – what school-level factors do affect attendance. It may be that students are similarly likely to skip schools regardless of a school’s grade. Figure 9 suggests as much, showing very little difference in mean attendance rates at each level of school. F schools – on average – have an average daily attendance of roughly 94 percent, while A schools have a slightly higher average of 95 percent. Evidence suggests that school grades do coincide with demographic factors not included in the model, such as racial makeup and relative poverty. It may be that the one percent difference in attendance has more to do with these factors and is thus relatively unreactive to the failing grade itself.

Figure 9:

Mean Average Daily Attendance Rates by School Grade



Failure Rates

Failure rates – the inverse of graduation rates – are unique among my dependent variables in that they actually factor into school grades. Unfortunately, for my study years, the Florida Department of Education only provides graduation rates broken down by race. It is unclear whether these graduation rates of racial sub-groups are equivalent to the overall graduation rates incorporated into the grading rubric. According to the 2017-2018 Guide to Calculating School Grades, the graduation rate is “based on an adjusted cohort of ninth grade students, and measures whether the students graduate within four years” (Florida Department of Education, 2017). This “adjusted” rate is not available in Florida’s archives, so I ran four sets of regressions with logged failure rates of each of the four largest racial subgroups: white, black, Hispanic, and Asian. The aim of these regressions is to test the null hypothesis that failure rates are expected to change at similar rates regardless of school grade reception, against the alternative hypothesis that failure rates are more likely to change after the reception of a failing grade.

Table 14

Results of Fixed-Effects Regression Models with Dependent Variable – Log White Failure Rate

	Model 1			Model 2			Model 3			Model 4		
Variable	CE	T	P> T	CE	T	P> T	CE	T	P> T	CE	T	P> T
Log White Failure Rate (Dependent)												
Year 11	.743	20.51	0.0	.746	20.29	0.0	.746	20.29	0.0	.751	20.38	0.0
Year 12	.583	15.13	0.0	.582	14.83	0.0	.584	14.91	0.0	.591	15.07	0.0
Year 13	.489	12.44	0.0	.488	12.30	0.0	.490	12.39	0.0	.503	12.65	0.0
Year 14	.422	10.35	0.0	.422	10.33	0.0	.186	10.44	0.0	.442	10.60	0.0
Fail-Grade	.149	2.81	0.005	.153	2.84	0.005	.010	2.85	0.005	.245	2.19	0.002
Fail-Prior	-.003	-0.06	0.952	-.001	-0.03	0.978	.043	0.19	0.848	.025	0.46	0.646
Fail-2 Years Prior	.041	1.20	0.232	.041	1.18	0.238	-.001	1.20	0.233	.051	1.42	0.156
Total Teachers				-.001	-1.07	0.283	-.064	-1.04	0.299	-.001	-1.11	0.269
Fail-Interaction								-0.53	0.597	-.048	-0.40	0.692
Fail-3 Years								-0.17	0.868	-.047	-0.40	0.691
Fail-2013-2014										-.167	-1.94	0.053
Fail-Prior-2014										-.106	-0.88	0.381
Constant	2.116	64.79	0.0	2.257	16.83	0.0	2.251	16.71	0.0	2.249	16.68	0.0
R-Square				0.336			0.336			0.338		

Table 15

Results of Fixed-Effects Regression Models with Dependent Variable – Log Black Failure Rate

	Model 1			Model 2			Model 3			Model 4		
Variable	CE	T	P> T	CE	T	P> T	CE	T	P> T	CE	T	P> T
Log Black Failure Rate (Dependent)												
Year 11	.572	16.21	0.0	.575	16.24	0.0	.575	16.16	0.0	.578	16.04	0.0
Year 12	.412	10.67	0.0	.413	10.51	0.0	.411	10.41	0.0	.415	10.33	0.0
Year 13	.348	9.01	0.0	.348	8.86	0.0	.345	8.72	0.0	.354	8.56	0.0
Year 14	.309	7.71	0.0	.309	7.57	0.0	.305	7.44	0.0	.317	7.15	0.0
Fail-Grade	.131	3.01	0.003	.129	2.92	0.004	.072	1.18	0.238	.105	1.49	0.137
Fail-Prior	.042	1.12	0.265	.044	1.16	0.245	.023	0.53	0.594	.031	0.69	0.49
Fail-2 Years Prior	.076	2.36	0.019	.076	2.34	0.02	.071	2.06	0.04	.075	2.15	0.032
Total Teachers				-.001	-0.45	0.653	-.001	-0.52	0.604	-.001	-0.56	0.573
Fail-Interaction							.098	1.23	0.219	.114	1.37	0.171
Fail-3 Years							.031	0.36	0.718	.002	0.02	0.981
Fail-2013-2014										-.088	-1.15	0.251
Fail-Prior-2014										-.057	-0.56	0.579
Constant	2.667	85.67	0.0	2.712	21.53	0.0	2.726	21.45	0.0	2.724	21.43	0.0
R-Square	0.221			0.221			0.222			0.223		

Table 16

Results of Fixed-Effects Regression Models with Dependent Variable – Log Hispanic Failure Rate

	Model 1			Model 2			Model 3			Model 4		
Variable	CE	T	P> T	CE	T	P> T	CE	T	P> T	CE	T	P> T
Log Hispanic Failure Rate (Dependent)												
Year 11	.425	13.36	0.0	.429	13.31	0.0	.430	13.30	0.0	.432	13.33	0.0
Year 12	.248	6.64	0.0	.253	6.73	0.0	.255	6.74	0.0	.258	6.69	0.0
Year 13	.120	3.13	0.002	.125	3.25	0.001	.127	3.31	0.001	.130	3.22	0.001
Year 14	.101	2.63	0.009	.107	2.75	0.006	.109	2.81	0.005	.116	2.80	0.005
Fail-Grade	.158	3.20	0.001	.161	3.26	0.001	.192	3.11	0.002	.187	2.52	0.012
Fail-Prior	.011	0.26	0.798	.014	0.31	0.757	.028	0.55	0.582	.033	0.64	0.525
Fail-2 Years Prior	.071	2.07	0.039	.072	2.12	0.035	.068	1.93	0.055	.071	1.94	0.053
Total Teachers				.001	0.51	0.610	.001	0.52	0.601	.001	0.51	0.613
Fail-Interaction							-.113	-1.11	0.269	-.093	-0.83	0.409
Fail-3 Years							.075	0.59	0.557	.072	0.54	0.589
Fail-2013-2014										.004	0.03	0.972
Fail-Prior-2014										-.090	-0.59	0.555
Constant	2.753	91.63	0.0	2.679	19.73	0.0	2.675	19.65	0.0	2.674	19.64	0.0
R-Square	0.163			0.165			0.165			0.165		

Table 17

Results of Fixed-Effects Regression Models with Dependent Variable – Log Asian Failure Rate

	Model 1			Model 2			Model 3			Model 4		
Variable	CE	T	P> T	CE	T	P> T	CE	T	P> T	CE	T	P> T
Log Asian Failure Rate (Dependent)												
Year 11	.296	3.90	0.0	.291	3.81	0.0	.308	4.14	0.0	.302	4.05	0.0
Year 12	.225	3.06	0.002	.229	3.10	0.002	.252	3.50	0.001	.249	3.46	0.001
Year 13	.158	2.17	0.031	.160	2.20	0.029	.177	2.53	0.012	.167	2.33	0.021
Year 14	.098	1.37	0.172	.100	1.40	0.162	.126	1.78	0.077	.109	1.49	0.137
Fail-Grade	.047	0.32	0.746	.045	0.31	0.756	.145	1.09	0.279	.019	0.12	0.905
Fail-Prior	.197	1.73	0.084	.189	1.65	0.101	.250	2.20	0.029	.239	2.08	0.039
Fail-2 Years Prior	-.035	-0.34	0.733	-.038	-0.36	0.716	-.066	-0.64	0.526	-.072	-0.70	0.487
Total Teachers				.002	0.60	0.548	.002	0.64	0.526	.002	0.61	0.540
Fail-Interaction							-.615	-2.11	0.036	-.830	-3.62	0.0
Fail-3 Years							.747	1.78	0.077	1.088	2.84	0.005
Fail-2013-2014										.209	0.98	0.329
Fail-Prior-2014										.496	1.23	0.220
Constant	2.216	42.69	0.0	1.981	4.98	0.0	1.961	5.18	0.0	1.979	5.23	0.0
R-Square	.049			0.050			0.060			0.064		

The results of these regressions are shown in tables 14-17. In each regression, except for that on Asian failure rates, the fail-grade variable is significant. This makes sense, as graduation rates factor into school grades. This explains why – in each case – the fail-grade variable carries a positive coefficient, suggesting that in schools that fail, we should expect to see higher failure rates relative to schools that do not fail. In these regressions – on white, black and Hispanic failure rates – the fail-prior variable is not significant, but curiously, in the case of the black and Hispanic regressions, the fail-2 years prior variable is. Model 2 in these regressions suggests that a failing grade two years prior is associated with an expected

7.9 percent increase in the black failure rate, and a 7.5 percent increase in the Hispanic failure rate.

It is unclear why the models would suggest such a prediction two years after a failing grade but not one year after. We know that each school district has an office or department to intervene in failing schools. Teachers report mixed feelings about these interventions. Perhaps interventions serve to decrease the number of failures, but not in any sustainable way. Once the “rescue team” departs, failure rates increase again. In Chapter Six, I quote one teacher – Theresa – who says that in the year after a failing grade, the district is helpful, but in the second year it begins to unfairly fire teachers based on student scores. It is possible that policy options like staff layoffs actually discourage students.

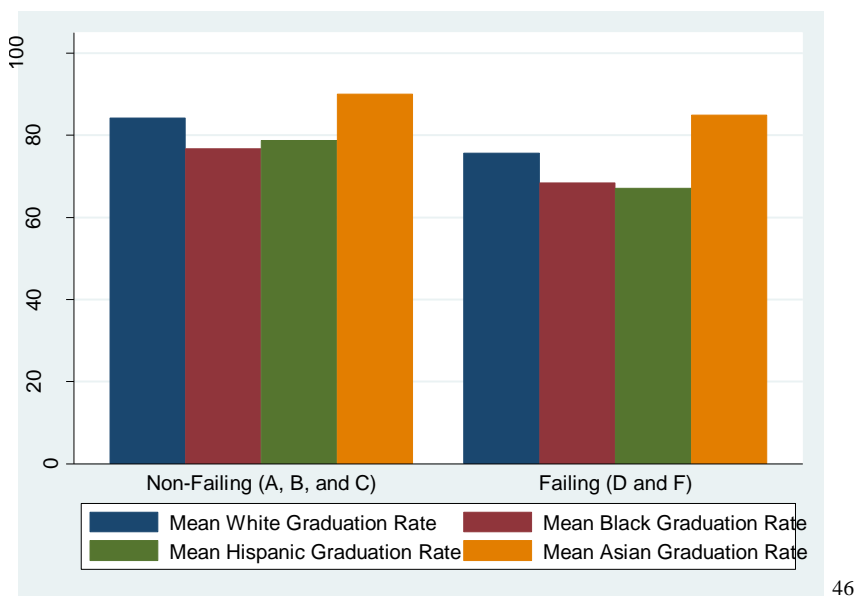
The results of the regression on Asian failure rates stand out as unique. In Model 4, the interaction terms for failing two and three years in a row are both significant. According to the results, we should expect a decrease of 56.4 percent in Asian failure rates in schools that fail two years in a row, but an increase of 196.8 percent in schools that fail three years in a row. These seem like unrealistically extreme expectations. In reviewing the data, it is clear that some schools have such a small number of Asian students that a change in one student’s performance could skew school-wide results. I speculate that these small numbers exaggerate the results of this last regression. This does, however, provide some support to the theory that policy interventions initially work but lack sustainable positive impacts.

It is clear that failing schools experience higher failure rates relative to non-failing schools. To further explore this, I graphed the mean graduation rates for each racial subgroup

in failing compared to non-failing schools (see Figure 10). As the regressions suggest, black and Hispanic students experience a larger gap in graduation rates between failing and non-failing schools, relative to their white and Asian peers. Although I lack demographic data, anecdotally I suggest that this is partly because students of color make up the majority of the student population at failing schools.

Figure 10

Graduation Rates by Race in Failing and Non-Failing Schools



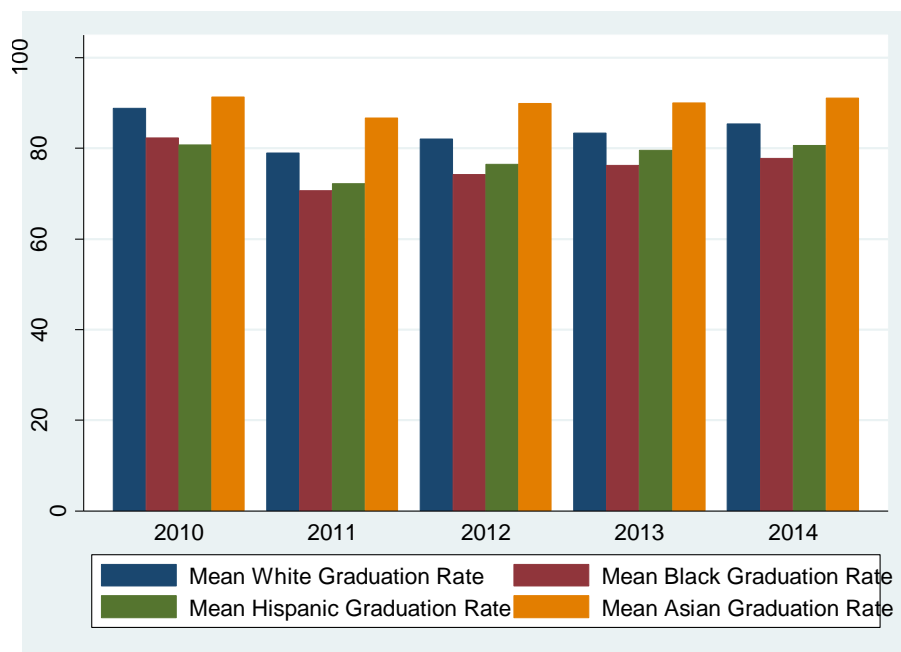
In each regression, the year-fixed-effects dummies are also significant. In all four groups, 2011 and 2012 show higher coefficients – relative to 2010 – than do 2013 and 2014. This suggests that failure rates increased post-2010, but less egregiously as the years passed. Figure 11 confirms this trend. I graphed the mean graduation rates for each racial subgroup

⁴⁶ I tested differences for all four subgroups by grade level using a Scheffe multiple comparisons test; while not all comparisons were significant, those between failing (D and F) and non-failing (A, B, C) schools were ($p < 0.05$); for complete test results, see Appendix E.

by year. Rates fell most significantly in 2011, then steadily increased from there. This fall affected black and Hispanic students more than their white and Asian peers.

Figure 11

Mean Graduation Rates by Race, by Year



While rates for all years and races are relatively high, they are not as high as the low – single digit – drop-out rates would suggest. As I discuss earlier in this chapter, the two are not necessarily complements, but the incongruity may lend credence to the worry that schools hide drop-outs by coding them as transfers to GED programs and charter schools.

Discussion of Quantitative Findings

Overall, the results of these quantitative analyses are more ambivalent than the strong statements made by some teachers in interviews. In the cases of teacher separation rates, attendance rates, and drop-out rates, there is not sufficient evidence to reject the null hypotheses and conclude that the reception of a failing grade is associated with significant changes – positive or negative. In regards to failure rates, the inverse of graduation rates, there is evidence to reject the null hypothesis and conclude that not only are decreases in graduation rates reflected in failing grades – as the state intends – but two years after a failing grade, the failure rates of black and Hispanic students continue to increase.

It is unclear why this would happen. I speculate that the interventions of counties and the states in failing schools may initially help students of color to stay in school and graduate. Perhaps, however, these interventions lack sustainable impacts and so – two years after a failing grade – graduation rates decline once more. Teachers offer some qualitative evidence to suggest as much. Future research should further probe this phenomenon, especially post the 2015 implementation of a new standardized test required for graduation.

There are other worrying trends regarding graduation and failure rates. While graduation rates are relatively high, there is a notable achievement gap between white and Asian students and their black and Hispanic peers. The failure rates of the latter groups appear to be either negatively impacted by a failing school grade, or to be unresponsive to county and state interventions. As expected, for all subgroups there is also a gap between students at failing schools and non-failing schools. If one of the aims of school grades is to

drive failing schools to improve in this and other measures, then grades seem to have been ineffective between 2010 and 2014. In fact, graduation rates actually fell in 2011 and otherwise remained stagnant in the five-year sample.

Surprisingly, my regression results do not show a significant correlation between the reception of a failing grade and teacher separation rates. This seems at odds with my qualitative data suggesting that failing schools have a hard time recruiting and retaining teachers. I suspect that unavailable explanatory variables may have more to do with retention than the grade itself. Perhaps the stigma of an F that I feared warded off teachers is less of a driver than factors of location, income, race, discipline, and administrative support, among others. It is also possible, as some of my teacher participants attest, that while a failing grade repels some teachers it attracts others who feel a calling to work with struggling students.

There are concerning trends in teacher separation rates. Between 2010 and 2014, teacher separation rates increased from a mean of 5.6 to 8 percent. The data does not reveal why this occurred, but teachers suggest in interviews that policy upheavals, increased accountability, and decreased respect for teachers play a role. Whatever the reasons, failing schools clearly lose more teachers. The mean teacher separation rate of F schools is nearly double that of A schools (roughly 11 versus six percent). Also worrying, the teachers who leave failing schools have less cumulative experience than their peers at higher-achieving schools. Interviews suggest that failing schools either hire less experienced teachers or cannot maintain staff long enough for them to develop more experience. Either case is problematic. Principals lament devoting energy and resources to train new staff, only to see them depart after a year or two. Experienced staff could potentially be more effective on their

own, as well as help to assist and train new teachers. I argue that for these reasons and others, it is impossible to evaluate the quality of schools without considering staff experience, training, retention and morale.

My regression results also do not support rejecting the null hypothesis regarding drop-out rates. This suggests that students are not significantly inspired to drop-out or stay in school following the reception of a failing grade. While this is unsurprising, based on what teachers describe as student apathy about the school grade, it is curious that an increase in drop-outs is not associated with the initial reception of the failing grade. School grades partly rely on graduation rates, but drop-out rates seem to have no significant effect. This implies that drop-out rates are not an inverse of graduation rates, which the state Department of Education confirms. I argue that this is somewhat misleading, especially since teachers and the media suggest that some schools hide drop-outs and failures by encouraging students at risk of not graduating to attend adult GED programs or charter schools. Policymakers should be conscious of these potentially “hidden” drop-outs.

The data does show that failing schools have average drop-out rates nearly double those of non-failing schools (2.3 versus 1.2 percent). While this divide is concerning, it is positive that both rates are – assuming their accuracy – quite low. Moreover, drop-out rates increasingly fell during the sample years. Mean rates decreased from 1.4 percent in 2010 to 1.15 percent in 2014 – a small change but in the desirable direction.

Attendance rates do not seem to be significantly impacted by the reception of a failing school grade. Thankfully, this suggests that failing grades do not deter students from

attending school, but it also suggests that interventionist policies designed to increase attendance in failing schools are not working. Future research should explore what school-level factors do affect attendance. Students attend F and A schools – on average – at roughly the same rates, so it may be that factors such as race and poverty have greater impacts on attendance than do school grades.

Once again, these findings detail a mixed and ambivalent relationship between school grades and these other variables. To some degree, these findings mirror my teacher interviews: the number of failing schools and teachers leaving schools increased in these years; failing grades do not appear to increase graduation rates, particularly for students of color; attendance does not appear to be impacted by school grades. In other ways, these findings are less damning than what teachers suggest: drop-out rates appear to be improving; the reception of a failing grade is not itself strongly associated with teacher turnover; drop-out rates are consistently low and graduation rates high.

I suspect that grades are more of an indicator than a causal factor in these outcomes. Teachers undoubtedly feel unfairly judged by grades, partly because they say grades are so dependent on demographics, relative poverty, geography, and history – variables absent from my models. Future research with more comprehensive datasets should explore these variables. Such exploration requires the state of Florida to publish – in an easily accessible way – this data. To do so, policymakers must admit that schools struggle on unequal footing. If they did not, then we would not witness 615 instances of schools failing at least two years in a row over a five year time period.

CHAPTER 9

CONCLUSIONS AND POLICY IMPLICATIONS: ADVOCATING HOLISTIC MEASURES OF SCHOOL QUALITY

At the outset of this research project, I wanted to explore the unintended and unforeseen consequences of Florida's school grades. As a former Florida teacher, I found it difficult to reconcile Florida's self-publicized story of school turn-arounds and state-wide success, with the frustration and anger I had heard from teachers. I wondered if a failing grade might in some way motivate further failure by shaming teachers and students. In my qualitative interviews, I sought to uncover how teachers feel about the reception of school grades – particularly failing ones – and what effects they observe for students. In my quantitative analysis, I sought to determine whether the reception of a failing grade was associated with any expected changes in factors other than test scores, such as attendance, graduation, drop-out, and teacher separation rates.

My quantitative analysis does not provide sufficient evidence that the reception of a failing grade is itself associated with expected changes in the aforementioned dependent variables, except for modest evidence that it prompts more teachers to leave. However, descriptive statistics show that there are significant differences in mean teacher separation, graduation, and drop-out rates between failing and non-failing schools. Not only did teachers increasingly leave failing schools between 2010 and 2014, they increasingly left all schools. Five years later, the state began the 2019 school year with roughly 3,500 teaching vacancies, with about 40 percent of teachers leaving within their first five years (Maxwell, 2019; Carver-Thomas & Darling-Hammond, 2017). Failing grades or not, something is wrong; and while the data does not demonstrate a correlation between receiving a failing grade and declines, it also does not demonstrate correlation with improvement.

In my qualitative interviews, teachers suggest that a lot is wrong with the school grading system. While most support the idea of accountability, they think school grades are simultaneously simplistic and confusing, reducing complex school environments to a single grade, based on a formula they suspect few understand. According to teachers, school grades ignore student growth as well as the unique challenges and strengths of schools, including socioeconomic and racial divisions. Despite some turnaround successes, there persists a strong correlation between failing grades and high numbers of students of color and low-income students. A failing grade seems to make attracting highly-qualified and experienced teachers difficult, while shaming those that do work with struggling students.

In my literature review, I outline three theoretical frameworks that shaped my early thinking on school grades. They revolve around intrinsic versus extrinsic motivation,

stereotype threat, and institutional shame. Regarding the first, research has shown intrinsic motivation to effectively improve teaching and learning, while extrinsic motivations – such as testing and grades – have negative effects (Alispahić, 2013; Harlen & Crick, 2003; Jones, Jones, & Hargrove, 2007; Schunk, Meece, & Pintrich, 2013). My research supports this conclusion. Teachers and principals report that failing grades make attracting high quality teachers difficult; they can also dissuade families with means from enrolling their children. While there are success stories of schools “turning around” after multiple failures, it is unclear whether their turnarounds result from the negative stigma of a grade, or from policy interventions by their counties and the state. It is also unclear whether a rising school grade effectively measures learning. Teachers disagree on whether students are aware of school grades – surely this depends on the school, classroom, and student – but they all agree that the testing that fuels grades is stressful and de-motivating.

I find no evidence that failing grades motivate teachers or students to work harder – an insinuation that frustrates and offends teachers who feel accused of laziness. They argue that they work as hard, if not harder, than teachers at A schools, often without the same financial rewards. Some counties have offices of school improvement or “achievement zones” meant to lift up struggling schools. While teachers disagree about the value of interventions, it is arguable that such pragmatic policy agents achieve much more than the motivational pressure of achieving an elusive A grade.

Stereotype threat comes from a person’s fear of confirming a negative stereotype; Steele (2010) and others find that when activated by an external test, that fear can interrupt academic performance (Carr & Steele, 2009; Lawrence, Marks, & Jackson, 2010; Osborne,

2001; Steele & Aronson, 1995; Wout, Danso, Jackson, & Spencer, 2008). I ask all of my educator participants whether they have observed such a phenomenon. No one had ever thought about student anxiety in terms of confirming stereotypes before, but related issues arose. Teachers tell me that testing is demotivating for students. Some, who fear they will perform poorly, would rather not try and maintain their agency, then try and fail. Others, feeling exhausted from multiple standardized tests, will rush through them rather than honestly answer questions. Still others believe that they will fail and their belief becomes a self-fulfilling prophecy.

I suspected that a failing school grade might contribute to stereotype threat or these other, related issues. Teachers disagree. Some think that students are aware of the school's grade and connect it with their own abilities, while others do not think students are aware or care about the grade. Teachers do suspect that the stress that comes with administrative pressure trickles down to affect students, something that other researchers have observed (Jones et al., 2007). In order to find out whether school grades add a new layer of stereotype threat, future research will need to focus more directly on students. For now, whether stereotype threat is an appropriate frame for understanding student failure or not, we ought to conclude that students of color, those challenged socio-economically, and those in struggling schools do not monolithically fail because they do not work enough. Larger numbers of these students fail and attend F rated schools for a plethora of complex social and institutional reasons that must be explored and addressed.

Finally, I build on Creed et al. (2014)'s concept of an institutional "shame nexus" – a combination of internal and external shame that reinforces shared values. I argue that this

concept applies to schools, where shared ideas of success and common symbols – such as grades – facilitate shaming. Despite sometimes bad press coverage and the public stigma of an F, teachers do not express any shame of themselves or their schools. Teachers tell me they feel proud of their schools and coworkers and that failing grades inaccurately hide their true accomplishments. Some teachers intentionally work at a failing school because they feel that their talents can best serve struggling students.

Still, principals tell me that the reception of a failing grade feels devastating and that it can make it harder to attract quality teachers. Some teachers at A schools tell me that they, or their colleagues, would not work at an F school partly because of the associated stigma, or what friends and family would say. Jacobsen, Snyder, and Saultz (2014) find that parents perceive a greater disparity between school grades than with other types of descriptors, suggesting that a D or an F grade amplifies community pressure on a school – which could be desirable if the school has the means to improve.

In this final section, I focus on some policy proposals that might improve the rating system and more accurately signal how a school needs to improve. These may seem like band-aids that fail to address the root problems. Unfortunately, some root problems may differ from school to school. More than any one state, or nation-wide diagnosis and solution, what schools need is flexibility in evaluation and policy. For example, I do not attempt to figure out why teachers leave schools, but I do believe that teacher retention needs to be part of a school's evaluation. Each school needs to recognize teacher departures as a problem, diagnose it, and build a solution into reform plans. Subsequently, the state ought to support school plans and incorporate progress into evaluations.

Other root problems are simply beyond the scope of my research. Some, for example, argue that it is impossible to improve public schooling without taking on standardized testing. My own feelings, as a teacher of eight years, are that standardized testing has trained students to look for the fastest route to the “right answer” rather than ask critical questions of themselves and others. There is no shortage of researchers, writers, and politicians who agree. Moreover, all of the teachers I interview articulate strong feelings on standardized testing, with some divergence of views. I acknowledge the centrality of standardized testing to the questions of education reform; yet I also acknowledge that throughout the modern education reform trajectory of No Child Left Behind, Race to the Top, and Every Student Succeeds, testing remains the most prominent constant. It is not going away. We should explore ways to mitigate the impacts of testing on curriculum, since the elimination of non-core classes, including the arts, disproportionately impacts disadvantaged students (Tienken & Zhao, 2013).

There is also a strong argument that nothing is hurting Florida’s students more than segregation, and that no true education solutions can ignore the achievement gap between white and Asian students and their black and Hispanic peers. It is extremely worrying that the failing school mostly prominently featured in this dissertation is described by multiple participants as “100 percent black.” While I do not have racial make-up data for the schools in my sample, the research of others shows that students of color attend schools with low scores, high drop-out rates, and insufficient college preparation; their schools tend to have less experienced instructors and less advanced peers (Orfield, 2013). Segregation – in Florida and across the nation – is growing worse, and school choice is contributing to the trend

(Orfield, 2013; Orfield & Ee, 2017; Strauss, 2013). As Orfield (2013) argues: “We cannot afford to continue discarding good communities with good schools one after another and investing money in replicating them 25 miles away at enormous environmental costs simply because we have been unable to develop mutually beneficial ways of living together” (p. 54). He suggests that school choice might contribute to integration efforts if states invest heavily in diverse schools and provide high quality options in integrated communities. While this issue deserves a comprehensive policy package of its own, I hope that a reformed, more holistic school evaluation system could help eliminate the stigma attached to schools of color, stemming the tide of white flight and attracting investment.

The Every Student Succeeds Act (ESSA, 2015) should have provided Florida with an opportunity to embrace constructive change. The law requires that school evaluation systems include a non-academic measure of well-being. It also requires that schools create “comprehensive improvement plans” in schools with underperforming student subgroups (Klein, 2016). Instead of diversifying its rubrics, Florida resisted the mandates and its ESSA plan was the last to be approved, even under the loose standards of Education Secretary Betsy DeVos’s Department of Education. Florida objected to federal requests that it incorporate the scores of English Language Learners into its school accountability. The state also wanted to avoid identifying the performance of special education students and racial subgroups separately. It eventually compromised by creating a “federal index” system separate from school grades, which includes the federal requirements to which it objected (Feinberg, 2018). This attitude and ultimate agreement is a mistake, a missed opportunity to diversify the factors included in school grading and to draw attention to the unique needs of

student subgroups (Hopkins, Thompson, Linqunti, Hakuta, & August, 2013; N. D. Jones, Buzick, & Turkan, 2013).

Others, including civil rights groups, have been effectively arguing these points since Florida's initial ESSA plan. The Tampa Bay Times, for example, posits that – without including the performance of English Language Learners, students in poverty, and racial subgroups – “it’s too easy, to borrow from the phrasing of the old federal law, to leave some students behind.” It calls out the state for circumventing ESSA by reporting such data to the federal government but not including it in school grades. As the paper states, “That means a school could achieve a good letter grade by Florida’s reckoning but still do poorly on the federal scorecard” and “it’s easy to guess which one parents and politicians will focus on” (Tampa Bay Times Editorial Board, 2019).

While I do not emphasize individual-level problems and policies – like merit pay for teachers – Florida could make progress with teachers by dropping programs and plans perceived by many educators as silly – if not discriminatory – time-wasters. For example, the Florida Education Association has filed a complaint with the federal government over the state’s “Best and Brightest” program, which rewards teachers with bonuses based partly on their own high school SAT or ACT scores. The Education Association argues that the program discriminates against older teachers who lack access to their test records “as well as teachers of color who historically do worse on standardized tests” (Brenneman 2016). While it is admirable to try and attract well-educated teachers to the field, many teachers – myself included – find the idea of judging adult professionals by their high school scores to be offensive and absurd.

Florida counties are rife with creative and effective education reform programs. Some are investing in struggling schools in innovative ways, or setting up “Achievement” or “Transformation” zones to help assist them. A sense of hope and optimism imbues these efforts. How do we ease the negative pressure or shame that teachers at failing schools describe feeling? I hope that making school grades accurately reflect the positive growth of schools is a start, as is integrating teachers into the evaluation process and encouraging internal-reform plans. Finally, by making grades more holistic and diverse, less pressure falls on a single test – for both teachers and students.

Grades Ought to Incorporate Additional Measures of Student Growth

The Every Student Succeeds Act (ESSA 2015) encourages states to consider factors additional to standardized test scores in judging school quality. One factor ubiquitous to all 50 states, and long a staple of Florida’s policies, is student growth. In other words, Florida is interested in – and ought to be – the degree to which student performance improves from year to year. The growth of the lowest 25 percent, and overall student growth, in math and English performance are included in the school grading formula and value-added model (VAM) data is included in each teacher’s evaluation. Clearly, the teachers and principals in my study favor this inclusion of growth, but they take issue with the way Florida chooses to measure it.

Measuring the value-added of any one teacher is difficult and controversial, not least because – as one principal tells me – teachers do not understand the scores they receive.

Value-added models are not the only means of evaluating teachers. For example, student-growth models (SGM), differ in that they measure growth based on classroom median, rather than mean, scores. There are pros and cons associated with each, but both are problematic insofar as they suffer from missing data (e.g. students new to a district with no comparable prior scores), non-random assignment (e.g. students arrive in particular schools and teacher classrooms based on ability, parent preferences, neighborhood, and other factors), and covariance (i.e. confounding variables are not independent) (Ballou, Sanders, & Wright, 2004; Kurtz, 2018; Wright, 2018).

The primary concern of my research is not value-added for individual teachers, but rather how to measure learning gains on a school-wide basis. The Florida Department of Education requires students to increase a level on the English and math sections of the Florida Standards Assessment from one year to the next, or to maintain an already high level. Unfortunately, as teachers point out and I discuss in Chapters Three and Four, in addition to suffering some of the same flaws as value-added models, these gains do not always reflect student progress. Students may begin a year significantly far behind or increase their scores but fall shy of entering the next level bracket. Teachers want recognition for the sometimes-significant gains their students do make.

One way to think about scores is in terms of proficiency versus growth. Systems based on proficiency, or status, look for students to achieve a target or “proficient” score, while systems based on growth look for students to improve their scores (Ladd & Lauen, 2010). While Florida grades incorporate both standards, its growth measures consider change from one year to the next and movement from one level of score to another, rather than marginal

movement within a year. This is a flawed consideration for several reasons, among them: students forget knowledge over the summer, emphasis and expertise differs between teachers, and – most notably – comparing one test to the next year’s is not a direct comparison, but rather a pairing of two completely different tests. Moreover, students take no comparable tests in grades 11 and 12, and in grades three through 10, only consistently test in two subjects: math and language arts. A high school student may take seven classes for four years, for a total of 28 classes; yet the school is only judged based on two years of scores for two classes, four of the 28.

By requiring students to move from one bracket to another rather than tracking their contiguous growth, Florida is making its growth model resemble a proficiency model. As many researchers point out, proficiency models consistently underestimate the progress made by schools serving disadvantaged students (Ladd & Lauren, 2017). In 2018, the Northwest Evaluation Association (NWEA) conducted a national study of 1,500 randomly selected schools. It finds that “many schools with low achievement were producing average or above-average growth.” Specifically, “this was true for 60% of schools in which more than 90% of students were eligible for free and reduced-price lunch” but “student growth in schools with few disadvantaged students varied widely” (Minnich, 2017). Chris Minnich, writing for the Learning Policy Institute, cites an example of Moos Elementary School in Chicago, which was “identified as among the 5% lowest performing schools” in Illinois, which – like Florida – “defines growth as year-over-year improvement in summative test performance.” Minnich continues: “Under the Chicago Public Schools School Quality Rating Policy, however, Moos

Elementary earned a top rating” (2017). This contradiction stems from Chicago’s use of a within-year growth model.

NWEA provides a good example of such a model, which would incorporate formative assessments throughout a school year to track student growth. Formative assessments differ from summative assessments in that they can guide a teacher’s practice in the classroom and inform student learning (Taras, 2005). Teachers already use mock and practice tests to help prepare students for the Florida Standards Assessment at the end of the year, but these practices are not standardized across teachers nor is student progress tracked by the state. They ought to be. This proposal of staggered, standardized and accounted formative assessments runs the risk of increasing the number of “tests” that students take, something that teachers discuss as problematic and the state has tried to reduce. In Chapter Seven, I discuss teacher fears that students are “exhausted” from testing and that in some cases they act disruptively to avoid tests or randomly answer questions to finish early.

The formative assessments that I am advocating would arguably not add to the number of tests students take, but instead replace school or teacher-generated tests already being administered. While teachers might balk at the loss of autonomy, they clearly want the state to measure the growth happening in their classrooms. The goal for teachers would be to incorporate standardized formative or “benchmark” exams into the classrooms, using the resulting data to guide instruction (Blaich & Wise, 2010). If students see these tests organically integrated into their daily learning and understand the goal of tracking their learning bit-by-bit, they may welcome them. It may decrease some of the anxiety associated with one high-stakes, unfamiliar exam at the end of the year (Akkaraju, Atamturktur,

Broughton, & Frazer, 2019). The state could thus track student growth within each year, under each teacher. It could use school-wide, mean growth rates to grade schools and it would no longer need potentially flawed value-added, fixed effects models to predict each teacher's contribution to student scores.

Even better, formative assessments could be made – in the words of Lakisha – “a little more authentic” if they were developed by teachers and involved activities other than bubbling answers. Schneider (2017) argues that during standardized tests, students are “effectively learning nothing. But when working on a research project, or a science lab, they are actively building their skills” (p.131). He advocates putting the products of such activities into a portfolio that would offer far more information to parents and teachers than a pair of standardized scores in math and English. He also believes that “such snapshots [of assessment] can also be empowering to students – showing them clear evidence of their growth” (p. 131).

In either case, teachers would not have to wait until the following fall to get their – in the words of Alex – “secret science” scores back, they could track and use the data throughout the year. This is essential, not only to make school grades more accurately reflect student progress, but to make testing a prognostic, useful tool. Boykin and Noguera (2011) argue that, key to closing achievement gaps, is providing teachers “with the results of assessments at the beginning of the school year so that information can be used to guide instruction and learning interventions” rather than “typically administered in the spring” with results “not released to schools until the end of the school year and students have been

assigned to new teachers” (p. 173). Districts across the country already administer such “benchmark” tests intermittently throughout the school year.

Grades Ought to Be Holistic Measures of School Quality

While teachers want scores to account for within-year growth on standardized tests, they also want school grades that capture more holistic measures of school strengths and weaknesses, as well as student well-being. Some bring up diverse achievements as strengths in their own right, things that might attract students; examples include advanced and honors programs, academic clubs, sports teams, and unique scholarship opportunities. Others, in particular the parents I speak with, see holistic measures as ways to ensure effective policy interventions; for example, without taking into account school psychological services and family supports, such services may never improve. Finally, many teachers see holistic observations as important context for school grades. These teachers tend to believe that the mission of schools differ according to the populations they serve.

Research supports these teachers and parents. Noguera and Wing (2006), building off multiple research and case studies, argue that “high-performing/high-poverty schools” across the country succeed by tending to the holistic well-being of their school communities. This involves “ongoing, site-based professional development” for teachers, “safe, attractive” campuses “equipped with appropriate technology,” diverse curricula including “the arts, sports, and extended learning opportunities after school,” “social support systems” to meet the “health, nutritional, and emotional needs” of students, and partnerships with parents and

communities (p. 178). Such schools also need targeted policies depending on their communities. High numbers of English-Language Learners (ELLs) may require teacher training in language acquisition; students with poor nutrition or health care may require partnering with social and health services (Noguera & Wing, 2006). This is not to say that schools in Florida do not employ these tactics, but that such measures and the issues they address are not considered by the state. State policymakers have argued that the main driver of poor school performance is low expectations (Borman & Dorn, 2007; Boykin & Noguera, 2011), yet evidence suggests that it is primarily socio-economic factors that affect educational outcomes like early childhood literacy (Haughbrook, Hart, Schatschneider, & Taylor, 2017; Porfeli, Chuang, Audette, McColl, & Algozzine, 2009; Welner & Carter, 2013).

We also know that schools – in Florida and across the country – are increasingly segregated by race and class and that segregation coincides with unique challenges (Goldstein, 2014; Orfield & Ee, 2017). Teacher retention is one such challenge, which I highlight in Chapter Eight. Economist Kirabo Jackson, using data on schools in Charlotte, North Carolina, finds that between 2002 and 2005, Charlotte schools grew more segregated and resulting schools of color lost significant numbers of highly-qualified, effective teachers – both white and black (Goldstein, 2014). Goldstein (2014) concludes that “while discrimination certainly accounts for some of the movement...part of the explanation is that high-poverty, minority-majority schools are more likely to experience administrative turnover and inept management” (p. 180).

Boykin and Noguera (2011) discuss the health challenges of poor and minority schools:

Children who can't see well can't read as well as those who can, and lower-class children, on average, have poorer vision than middle-class children. Lower-income children have a higher incidence of lead poisoning, poorer nutrition, and higher rates of iron deficiency anemia, which results in impaired cognitive ability. They have greater exposure to environmental toxins, air pollution, and smoke, and therefore greater incidence of asthma. Lower-class children have less adequate pediatric care, resulting in more frequent absences from school (p. 62).

Similarly, Stretesky and Lynch (2002) find that in the 1980s and 1990s, Hillsborough County, Florida schools located near environmental hazards, such as abandoned toxic waste sites, were disproportionately black and Hispanic. I contend that the state of Florida – in ascertaining the quality of schools – cannot ignore racial composition and any associated inequities.

Malcolm Harris (2017) persuasively argues that schools are adept at meeting goals laid out for them. The law, he says, “is effective at incentivizing school administrations to meet their yearly progress goals, but that’s not the same as incentivizing them to improve instruction or learning.” He continues: “The state accepts these assessment reports as valid representations of educational quality, so administrators and teachers – and most of all, students – have to generate the right reports” (p. 115). The way they go about producing the right reports, according to Harris and others, is to limit arts and music instruction in favor of test preparation. My point here is not to argue that standardized testing and test preparation are inappropriate goals, but that they are not sufficient goals. The law ought to expand the definition of quality education, capturing the different roles that schools play. In so doing,

perhaps it could capture context for school grades, encourage appropriate policy interventions, and allow schools to publicize what they do well.

I see several ways in which Florida and other states could accomplish this. As I discuss in my literature review, the form of ratings has been shown to affect the nature of parents' perceptions (Jacobsen, Snyder, & Saultz, 2014), so we know this is not an arbitrary issue for parents or for teachers. One option, as some teachers I interview suggest and as some other states have done, is to replace the A to F scale with descriptive indicators such as "needs improvement" or "highly effective." This is how Florida evaluates teachers. While such a rating is more qualitative, it runs the risk of replacing an arguably objective and easily understood metric. While everyone has received a letter grade at some point in their education, parents might have very different interpretations of what "needs improvement" means. Such descriptors also fail to add any nuance or holistic quality to the ratings system. Numbered rankings – which some states use – seem to suffer the same flaws.

Another option is to provide different grades for different performance arenas. Jon Valant, of the Brookings Institute, proposes such a system. He suggests that "a school might receive an 'A' for academic growth but a 'C' for extracurricular activities," with some indication of which domains the state values more (Valant, 2017). This would be an upgrade over the current system of one letter grade per school; however, it suffers some of the same flaws. First, I worry that the brevity of newspapers and the media would mean that only one grade – the one based on test scores – would be reported. Second, while a variety of grades would paint a more nuanced picture for parents, the formula for arriving at grades would be

even more esoteric, with as many different calculations as there were scores. This would do little to satisfy teachers' desires that parents understand grades.

As an alternative model, I suggest a state-produced school profile similar to those published by organizations like the nonprofit GreatSchools. Searching for a school on the GreatSchools website will display a wealth of information including reviews from students and parents and a star rating out of ten. In addition to the single star rating, visitors will see sub-ratings in the categories of academics, equity, and environment. Additional sections discuss college readiness, test scores, advanced STEM (science, technology, engineering and math) courses and demographic information on race, low-income students, and students with disabilities.

Despite this diversity of information, GreatSchools profiles suffer from some of the same flaws as school grades. School star ratings are based primarily on standardized tests and features of the website work to hide lower-scoring – and typically more diverse – schools from the map (McKay, 2018). Other online school ratings tools are also insufficient and potentially harmful insofar as they perpetuate stereotypes or gauge quality based on the wrong metrics. Schneider (2017) writes of such organizations:

There is a clear desire among the public for information about school quality. But for-profit third parties have met demand in a manner that often borders on the irresponsible. Their methodologies can be misleading. They can exacerbate stereotypes. And their presentation of results unwaveringly pits schools against each other – as if there are only a handful of decent schools out there, and as if there is a one-size-fits-all school for all young people and all values (p. 67).

Schneider identifies a “clear desire among the public” for school profiles, which is all the more reason for states to provide them. States should find ways to improve upon these online

models. The responsibility would be great, as such a profile published by the state may carry greater authority and weight with parents and schools. Arguably, for school choice advocates, this would increase parents' agency by providing them with more information. Parents could choose schools based on their own priorities and values.

In addition to maximizing individual choice, more information empowers communities to push for systemic change. Schneider (2017) explains:

More robust and carefully designed information systems might also help parents and members of the public engage more fully as key stakeholders in the schools. Parents, for instance, often know about the experiences of their own children in school; but they rarely know how those experiences relate to those of others. This is fundamentally disempowering because it means that any advocacy tends to be viewed as personal in nature rather than as a campaign on behalf of the whole community. Personal advocacy, of course, is fine; but it does little to create systemic change (p. 92).

More information may also enable community members who want to contribute to schools, but are not sure how, to become active.

Most importantly, in my view, publishing this information would increase pressure and accountability on the state and its schools. Policymakers and educators should have to acknowledge that factors as varied as integration, poverty, course options, and college entrance impact school quality. With this said, a state must hold itself and schools accountable for this enriched data. Under Mayor Bill de Blasio, New York City developed a "School Quality Snapshot" composed of categories including effective school leadership and strong family-community ties. Unfortunately, New York does not hold schools accountable for the scorecard. Instead, it relies solely on Annual Measurable Objectives (AMOs) based primarily on standardized tests (Schneider, 2017).

A school profile ought to also include information useful to teachers. Fullan and Quinn (2016), analyzing education reforms in Ontario, Canada, advocate schools “adopting and applying indicators of organizational health as a context for individual teacher and leader performance, such as staff retention rates, leadership turnover rates, teacher absenteeism levels, numbers of crisis related incidents, and so on, in addition to outcome indicators of student performance and well-being” (p. 120). Not only might this information be valuable to teachers looking for employment and parents looking for schools, but it should provide context for teacher evaluations. Each teacher’s “value-added” should not be considered independent of factors like crisis related incidents that might interrupt learning.

As I discuss in Chapter Eight, on quantitative findings, Florida already produces much of this data. Unfortunately, each variable is housed in a separate Excel spreadsheet, by year. Confusingly, across years and spreadsheets, school names are often entered differently (e.g. George-Brian High School versus G.B. High School). Some variables are available for some years, but not others. This made compiling a dataset extremely difficult and time consuming for me and I imagine would make finding information on their children’s schools difficult for parents. Instead of obfuscating this data, Florida should publicize it in an easily-digestible way, by school, in place of or in addition to school grades.

These suggestions rely mostly on currently collected data. While increasing the degree to which schools are described to the public, they would do little to change the way that opportunity-to-learn (OTL) factors – such as relative income, food, and housing security – are woven into grades themselves. This is problematic on the one hand because research demonstrates strong correlation between OTL, school resources, and test scores (Archibald,

2006; Berkowitz, Moore, Astor, & Benbenishty, 2017; Cadima, Peixoto, & Leal, 2014; Erbe, 2000; Grodsky, Warren, & Felts, 2008; Welner & Carter, 2013) and on the other because teacher and school effects are not equal across outcomes. In other words, a teacher might be successful in raising student test scores, but have negative impacts on measures like discipline and attendance, or vice versa (Gagnon & Schneider, 2019; Grissom, Loeb, & Doss, 2015; Jackson, 2016; Petek & Pope, 2016).

To account for these differences, in outcomes and opportunities, Florida might – in addition to reporting data – incorporate variables into a revised school quality framework (SQF). Chicago Public Schools and the California Office to Reform Education (CORE) have both done so in recent years. Chicago Public Schools has used survey data to measure “the level of support for teachers, the involvement of families, the safety and orderliness of the school, and the level of academic challenge in classes” into its evaluations (Gagnon & Schneider, 2019, pp. 738-739). The California program – CORE – utilizes a 100-point scale with 60 points dedicated to raw and growth test scores, while 40 points “rely on a broader range of measures, including how many students are missing significant amounts of school, how many are suspended or expelled, and how many English language learners have become fluent” (p. 739). Schneider (2017) adds that in CORE districts, “low-performing schools, which are identified as ‘priority’ schools – signaling a concern with capacity building rather than punishment – are paired with higher-scoring peers, with the goal being to create a ‘community of practice’” (p. 57). This demonstrates how a holistic approach can shape a district or state’s actions pre- and post-evaluation.

Gagnon and Schneider (2019) developed a 32-factor SQF grouped by major categories such as “Relationships,” “Safety,” “Academic Orientation,” and “School Culture.” They rely on focus group and teacher and student surveys, as well as district administrative data to evaluate schools in a major Massachusetts school district. They compare the SQF outcomes to the state’s progress and performance index (PPI) test-score rankings and find that – while the two “roughly align” the PPI system “offers only summative information that cannot be used for school improvement, it fails to capture information about the OTL and social-emotional learning, and it strongly reflects school demographics” (p. 751). When broken down by category, the two metrics diverge. “Teacher perceptions of the usefulness of professional development, student perceptions of the level of teacher interest in students, and principal leadership” all “exhibited moderately negative correlations” with test data. In other words, students in low-performing schools believe teachers are more interested in them. Gagnon and Schneider conclude that “specific strengths and weaknesses [of schools] may be hidden even if an aggregate relationship is positive” (p. 752).

Schneider (2017) explains how such a system might lead to contrary conclusions. Using an early iteration of school quality framework – which he called the Dream School finder – he evaluated schools in Somerville, Massachusetts. He explains:

Somerville High came in at number fifteen [in the state]. To some people, this was a shock. When the Globe released the results of the 2013 MCAS test, Somerville High ranked 271st out of 354 schools in English and 262nd in math...suddenly people were forced to confront the possibility that schools they knew to be “bad” might turn out to be otherwise. Alternatively, many schools assumed to be “good” because of their high test scores might actually have some things to work on (p. 7).

Measuring schools by metrics that include things like relationships and school culture would not only present a more accurate vision of schools, empowering parents to make more educated choices, but also push “good” schools to work on currently under-the-radar deficiencies. This should be a goal of politicians pushing for increased accountability.

While any of my suggestions in this section would be a positive step for Florida, developing a truly holistic system that measures not only growth and performance on standardized testing, but also the qualitative and quantitative factors I discuss in my dissertation, would be both revolutionary and most reflective of the true state of education.

Grading Metrics Ought to Incorporate Teacher Voice and Innovative Goals

One of the most strongly articulated desires of teachers in my study, and in the media, is their desire for policy-input and buy-in to the accountability system. As I stress in Chapter Four, the opposite has historically been true, with teachers and the Florida Education Association strenuously resisting every iteration of A+, school rankings, and teacher merit pay. This dynamic of antagonism between the state and educators is not sustainable, nor beneficial for children. Instead, in addition to treating educators like partners, the state should give them a role in setting reform goals for their schools. I propose that this is done in two ways: first, the state and districts should call together multiple stakeholders, including teachers, to solicit policy feedback and craftwork; second, state policy should build points into the school grading formula that allow teachers and schools to pursue self-developed goals.

“Coherence” has become a buzz-word in education reform circles, with many arguing that the key to educating students is a consistency of goals and methods across multiple disciplines, and that the key to successful education reform is collaboration among multiple stakeholders. Youngs and Bell (2009), for example, argue that between 1985 and 2005, Connecticut instituted successful education reforms by: building political capital with teachers by accompanying higher accountability with higher pay; creating policies that supported one another; and prioritizing capacity and relationship-building in the policy process. Fullan and Quinn (2016) advance a coherence “framework” defined as: “focusing direction, cultivating collaborative cultures, deepening learning, and securing accountability” (p. ix). They argue that, rather than embracing these tenets, since No Child Left Behind in 2001, policymakers have imposed “solutions that were crude and demotivating for the very people who have to help lead the solution – teachers and administrators” (p. 3).

Fullan and Quinn (2016) claim that – like Connecticut – Ontario has transformed a school system with “stagnated student achievement, labor disruptions, and low teacher morale” via a “renewal of public education” that differs from reforms in other places because of “the widely shared ownership of the reform agenda on the part of teachers, schools, and district school board leaders” (p. 37). This renewal is characterized by a 20 percent increase in students passing literacy exams, a 25 percent increase in high school graduation rates, and a significant decrease in the gap between special education and peer performance. They further write:

Educators were viewed as professional partners, and solutions were sought from the field. Promising practices from the early innovative districts and schools were identified and shared. Teachers were engaged to create exemplars, videos, and

resources that were shared widely. This reliance on grown from within, augmented by experts, sent strong messages to teachers and leaders that their input was valued (p. 40).

This philosophy of coherence and shared ownership of reforms contrasts starkly with Florida's emphasis on competition between schools and contests for bonus money. In many instances – such as the recent “Schools of Hope” initiative to incentivize charter school competition – the state has pursued the direct opposite track as that advocated by the Florida Education Association. These contrasting visions can turn teachers into acting resisters of policy instead of reform agents (Paige, 2013).

A first step in moving towards a coherent and shared policy agenda – is talking to one another. Florida attempted this in 2013 when then Governor Rick Scott called an “education summit” composed of lawmakers, parents, teachers, school officials, and business representatives. The summit lasted for three days; Governor Scott did not attend. The main takeaway from the event, according to *the Tampa Bay Times*: change the school grading system. One parent suggests that, instead of grades, “she’d rather see a report card listing how a school does in several areas” while others “suggested tacking pluses and minuses on the grades.” Florida’s teacher of the year at the time fears that the F grade “is a stigma” that “is making teachers leave.” Another teacher calls for “the need to grade programs within schools” such as standard students and International Baccalaureate students, which he calls “two different worlds” (Gardner, 2013).

In short, the 2013 summit produced many of the same findings and recommendations that I propose in 2019. I cannot find any official publication of the outcomes of that summit, nor to my knowledge did it prompt any policy reforms. While the Florida Education

Association and organizations like Impact Florida have held summits in the years after, the state has not. I suggest that the state regularly hold education summits, as should counties. These may succeed in generating innovative ideas and giving teachers a voice and buy-in to policymaking; however, summits must be followed by consideration, research and implementation of ideas.

Second, the state ought to recognize the internal goals set and pursued by schools. In my literature review, I discuss the consensus that – for students – intrinsic motivation is more effective than extrinsic motivation. The same is true for adults (Collinson, Cook, & Conley, 2006; R. Elmore, 2004; R. F. Elmore & Fuhrman, 2001; Firestone, 2014). For years, teachers have been encouraged and sometimes required to form professional learning communities (PLCs) – groups of teachers who share ideas, plan together, and compare data. Fullan and Quinn (2016) argue that the implementation of these communities “has lacked depth.” “If we want our systems to be authentic – energizing environments for students,” they write, “then we must create them for the adults as well” (p. 63). They conclude that, “Deep, collaborative experiences that are tied to daily work, spent designing and assessing learning, and built on teacher choice and input can dramatically energize teachers and increase results” (p. 63). These sorts of meaningful experiences should replace the professional development that several teachers in my study lament as meaningless.

Based on studies of schools in North Carolina and Kentucky, R. F. Elmore and Fuhrman (2001) find that accountability systems based on individual teacher responsibility do not lead to improvements in instruction or learning. R. Elmore (2004) writes that schools will implement “the external accountability system in pro forma ways without ever

internalizing the values of responsibility and efficacy that are the nominal objectives of those systems” (p. 134). Moreover, studies suggest that high-stakes testing makes teachers feel pressured and consequently controlling in ways detrimental to students (G. M. Jones et al., 2007). The alternative is for schools to develop and set internal goals, which have much more potential to impact teacher and student learning (Collinson et al., 2006). Schools – in Florida and across the nation – already do this, of course, but plans are often directed at improving test scores, or are otherwise not incorporated into school grades. Instead, the state might assign a certain number of points in the grading formula to the fulfillment of internally-developed goals.

We have models for what this might look like. Florida statutes require schools that receive a D or F grade to submit a school improvement plan (SIP) to their county school boards for annual approval.⁴⁷ Districts can choose to require a SIP for all schools. These improvement plans are published online and feature identified problem areas and achievement goals for the school, as well as plans for improvement. While these goals are typically related to raising student test performance, the same model could be used to submit any type of policy plans to counties. Schools could develop improvement plans related to discipline infractions, attendance, or teacher separation, among other things. This would be the simplest way to allow schools to self-identify areas for improvement, self-develop reform policies, and track their progress. Plans could be evaluated by counties or by the state and awarded a score based on the extent to which the school reached its goals.

⁴⁷ Section 1001.42(18), Florida Statutes

Additionally, the state and federal governments seem to champion the innovation and reforms they see happening at charter and private schools, why not grant public schools the same flexibility and track their success? Massachusetts has implemented such an option. In 2010, the state created a new category of school – innovation schools. Innovation schools are public schools, that remain within districts and are publicly funded, but have flexibility to innovate in areas of curriculum, budget, schedule and calendar, staffing, professional development, and school district policies. Innovation schools can be entirely new schools, converted public schools, or school-within-school academies. They begin life as a detailed innovation plan submitted by parents, teachers, principals, school committees, teacher unions, or other actors. In other words, they draw into the fold the same people I argue here are often excluded from decision-making.

An option like this seems like an ideal fit for Florida, where teachers complain to me that they lack the flexibility of charter and private schools that are immune from many of their laborious requirements. The head of the Innovation Schools program for the Massachusetts Executive Office of Education explains that:

As a former teacher I would often think, ‘If only we could do things this other way,’ Now teachers have a place to go with their ideas. These schools can be an excellent demonstration of teacher professionalism and teacher-initiated reform. There’s no lack of good ideas. Teachers know the school population intimately and can custom-tailor solutions to the kids in front of them (Barrett, 2013).

Teachers bring their ideas to the table to help develop goals; then, each subsequent year, Innovation Schools develop “Measurable Annual Goals” for which they are held accountable by local school committees.

While I believe such a flexible school category would be a good thing for Florida's students, the full embrace of such a program is not needed to improve school ratings. Instead, the state could allow schools the flexibility to create innovative programs or goals. Their plans for doing so could be evaluated by counties, which could then track their progress. Doing so might increase teacher involvement and their sense of agency; alleviate some of the antagonism teachers hold for charter and private schools; spur innovative ideas stemming from teacher expertise – which would make teachers feel like the professionals they are; and make school ratings more holistic and educator-sponsored. Tracking these self-developed goals will not ease all the pressures of standardized testing, but it would provide alternative inputs so that testing is not the be-all, end-all of scores (Welner & Carter, 2013).

Final Thoughts: Chapter Nine

In the introduction to their book *Closing the Opportunity Gap*, Welner and Carter (2013) write:

A narrow focus on the achievement gap predictably leads to policies grounded in high-stakes testing, which in turn leads to narrow thinking about groups of students, their teachers, and their schools. While these assessments attempt to determine where students are, they ignore how they may have gotten there and what alternative pathways might be available for future students. Schools, principals, and teachers are told that they have 'no excuses' and that they will be held accountable for results. Similar pressure is exerted on students. This accountability, however, is rarely extended to those making these demands. Policy makers are not required to provide supports necessary for equitable learning opportunities, nor are they held accountable for the consequences of these tests (p. 3).

I quote Welner and Carter at length because they both capture the essence of my suggestions in this section and provide a warning. While closing the “achievement gap” is an admirable goal, which I discuss throughout this dissertation, it is insufficient. Like school grades, it reduces student achievement to a single metric. I believe that by expanding the metric, by incorporating holistic measures of wellness, Florida might paint a more accurate portrait of schools and improve the lives of teachers and students. Insofar as standardized test scores remain a factor, the state should consider student growth within a year in order to recognize progress and ascertain truer measures of teachers’ value-added. Give those teachers buy-in to the evaluation system by building self-determined goals into the rubric. Recognizing that teachers and administrators can effectively gauge some of the needs of their own schools is a step towards respecting them as professionals, the need for which cannot be overstated.

Key to all these ideas is a recognition that the needs of schools – and perhaps their missions – differ depending on the student populations they serve. Key to different populations is a legacy of segregation in Florida that unfortunately remains a problem today. Florida has too long focused on accountability for schools and teachers, without fulfilling its reciprocal upwards accountability. Throughout the nation, between the 1970s and the 1990s, the accountability movement replaced desegregation and equity as the primary driver of education reform (Borman & Dorn, 2007; Fullan & Quinn, 2016; Goldstein, 2014). While I do not expect that driver to change now, policymakers have no excuses not to recognize the contexts that make teachers, students and their schools “fail.”

APPENDIX A(2): INTERVIEW GUIDE FOR TEACHERS AND PRINCIPALS

1. Tell me about yourself.
2. How long have you been a teacher/principal?
3. What led you to this profession?
4. Where did you start teaching?
5. What led you to your current school?
6. How would you characterize your current school?
7. Tell me about the teachers at your school.
8. How far back do you know your school's grade?
9. What is its current grade?
10. How did you react to finding out the grade for last year?
11. Do you think that grade accurately reflects the school?
12. How do you discuss the grade with family, friends?
13. How about parents? The community?
14. Are there teachers who have left, or might want to leave, because of the school grade?
15. How aware are the students of the school's grade?
16. How do they feel about it?
17. Are there any behavioral problems unique to a failing school?
18. Do you support the school grading system?
19. How do other teachers at your school feel about it?
20. Is there a more accurate or appropriate way to grade schools?
21. Is there a more appropriate measure than an A to F scale?
22. Do you think the school's grade effectively communicates to parents the strengths of the school?
23. What information do you think parents need to know?
24. Is there any evaluation or data that the state could give you as a teacher, or principal, that you don't receive now?
25. In what ways are teachers, or principals, involved in the state's policy-making process?
26. What about county policy?
27. If a policy-maker were asking for your input, what would you tell them?
28. How do students feel about standardized testing?
29. Do stereotypes play a role in their performance?
30. Do you think that school grades have any consequences outside of standardized test scores?
31. Are there any consequences that the state is not paying attention to?
32. One of my goals is to understand how school grades affect teachers and students. Is there anything I have not asked about that you think is important for me to know?

APPENDIX B(2): CONSENT LETTER FOR INTERVIEW PARTICIPANTS

Dear Participant,

I am a doctoral candidate at the University of Massachusetts in Boston. I am conducting research on Florida's schools and impacts of the school grading system. I would like to conduct an interview with you related to school policies, grades, students, and teachers. I'm interested in how students feel about the educations they receive and how students, teachers, and parents interact. I'm also interested in your feelings about your child's school and how you have reacted to school grades, or in your interest in the school system as a non-parent actor.

I would like to ask for your permission to work with you to collect data. This would include taking handwritten notes, audio-taping an interview and then reviewing the data. Interviews will last approximately one hour and no further participation is expected past the interview. I may contact you via e-mail if I have any follow-up questions.

I do not plan on using the real names of participants in any written products. Instead, I will assign interviewees pseudonyms. Names may be attached to a recording and transcript accessible only by me and a transcriptionist. All data such as notes and audio recordings will be kept on a hard-drive protected by me.

Your participation in this study will be completely voluntary which means that you can stop at any time during the study. I anticipate minimal risk to you as a result of participating in this study. You will have the opportunity to keep sensitive information excluded or to discontinue the interview at any time without penalty.

The nature and purpose of this research has been satisfactorily explained to me, and I, _____ am willing to participate in the study. I understand that I am free to discontinue participation at any time if I so choose and that the investigator will gladly answer any questions that arise during the course of the research.

_____ I consent to audio recording of my interview

_____ I consent to being assigned a pseudonym in any written or otherwise-produced products

(Date) (Participant's signature) (Participant's name, printed)

(Date) (Investigator's signature)

You have the right to ask questions about this research before you sign this form and at any time during the study. You can reach me at luke.kupscznk001@umb.edu, or my dissertation chair – Mark Warren – at mark.warren@umb.edu, if you have any questions or concerns. If

you have any questions or concerns about your rights as a research participant, please contact a representative of the Institutional Review Board (IRB), at the University of Massachusetts, Boston, which oversees research involving human participants. The Institutional Review Board may be reached at the following address: IRB, Quinn Administration Building-2-080, University of Massachusetts Boston, 100 Morrissey Boulevard, Boston, MA 02125-3393. You can also contact the Board by telephone or e-mail at (617) 287-5374 or at human.subjects@umb.edu.

APPENDIX C(2): CODE LIST

Code	Description	Count
Community		
Parents	Statements from parents about parenting	8
Demographics and Segregation	Discussions of student demographic attributes (including gender, sexuality, socioeconomic class, race, among others), as well as segregation	50
Press Coverage	Mentions of school grades or events appearing in the press	5
Teacher Association	References to the Florida Education Association (FEA)	4
Internal to Schools		
Administration Pressure	Statements from teachers about pressure they feel from administrators or principals	9
Characterization of School	Teacher and principal descriptions of their own schools, faculty/staffs, and student bodies	23
Principal Transfer	References to principals being transferred from one school to another	6
State/County Intervention	Descriptions of interventions by state or county officials in failing schools (e.g. classroom visits, layoffs, scripted lesson plans)	11
Technology	Discussions of relative technology access in different schools	4
Policy		
A+ Implementation	Discussions of the 1999 implementation of school grading and the A+ education policies	3
County Policy	Discussions of policy set or administered at the county/district level	10
School Cheating	References to schools cheating – or helping students cheat – on standardized tests	3
State Policy	Discussions of state-level policy and policy reforms, including school grading formulas	19
Law	Discussion of federal and state-level law/legislation	5
Policy Suggestions	Recommendations of federal, state, and county-level policy changes – including school grading formulas – by teachers, principals, and parents	36
School Choice and Charters		
State Scholarships/Vouchers	Discussions of state scholarships/vouchers available for students to transfer from failing public schools to private schools, as well as non-profits attempting to assist with applications	9

Charters and School Choice	References to charter schools and school choice, especially comparisons of policies and regulations affecting charters versus public schools	39
Students Leave	Discussions of students leaving a school to attend a charter, private, or different public school	17
School Grades		
School Grades	References to school grades in general, including acknowledgements of a school's history of grades	19
Expression of Shame	Any expression by teachers, principals, or parents of feeling ashamed or embarrassed by association with particular schools, as well as descriptions of defending a school to family or friends	30
Grade Accuracy	Discussion of the relative accuracy or inaccuracy of school grades as measures of school performance	26
Student Reactions to Grades	Teacher, principal, or parent accounts of how students react/reacted to their schools' grades	21
Teacher Reactions to Grades	Teacher, principal, or parent accounts of how teachers react/reacted to their schools' grades	10
Student Performance		
Achievement Gap	Discussions of differences in scores between various student sub-groups, especially between black and white students	7
Attendance	References to student attendance, particularly what factors affect attendance	3
Standardized Testing	Any discussion of standardized testing – either in general or in reference to a specific test (e.g. Florida's FSA or FCAT)	39
Stereotype Threat	Statements with relevance to Claude Steele's theory of "stereotype threat" or general thoughts on how stereotypes might affect student performance	17
Student Out-Of-School Challenges	Discussions of challenges faced by students outside of school, including housing, nutrition, gang and community violence, family life, transportation, among others	30
Discipline	Comments on discipline challenges or practices	26
Student Goals	References to the future goals of a student (e.g. attending college, landing an internship)	9
Student Emotions	Discussions of students' feelings such as stress, alienation, pride, shame, and apathy, among others	15

Student-Teacher Relationship	Descriptions of relationships and communication between students and teachers, either specific or general	18
Teacher Pressure on Students	Discussions of how pressure trickles down from the state's school grades or administration policies to students	5
Teacher Emotions and Job Satisfaction		
Teacher Motivation	Teachers' accounts of their own motivation for becoming a teacher, staying in the profession, or working at particular schools	22
Teachers Discouraged/Leave	Personal accounts – or references to other teachers – of feeling discouraged or leaving a school altogether	34
Teaching Experience	Descriptions of a teacher's past experiences teaching (e.g. how many years, what schools)	21

APPENDIX A(8): DISPLAYS OF SKEW FOR PRE- AND POST-TRANSFORMED VARIABLES

Figure 1

Display of Leftward Skew for Average Daily Attendance Rates 2010-2014

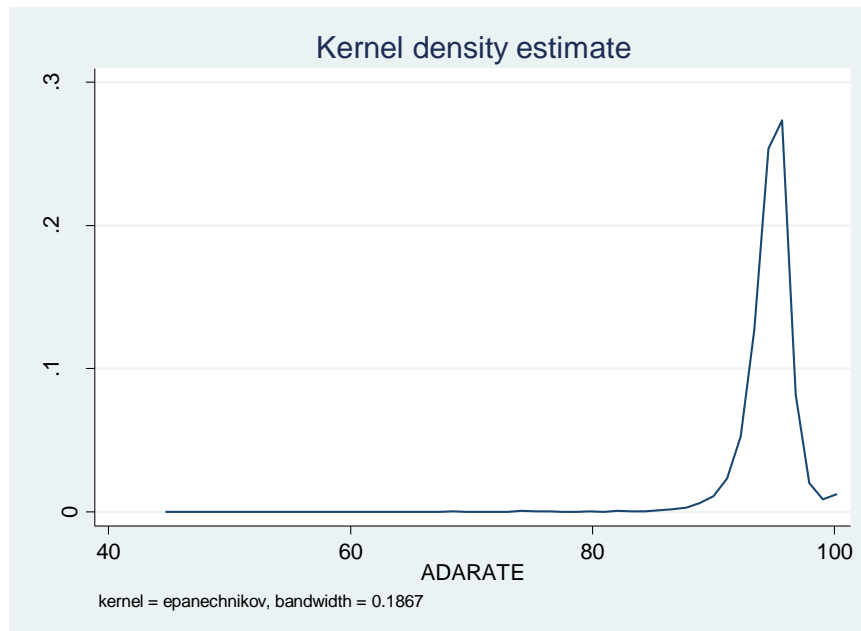


Figure 2

Display of Rightward Skew for Total Drop-Out Rates 2010-2014

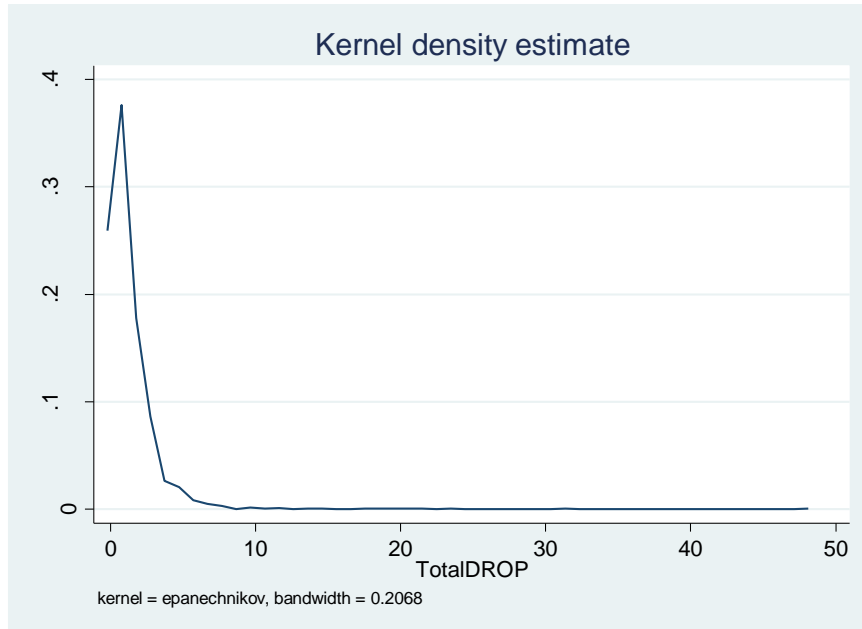


Figure 3

Display of Rightward Skew for Teacher Separation Rates 2010-2014

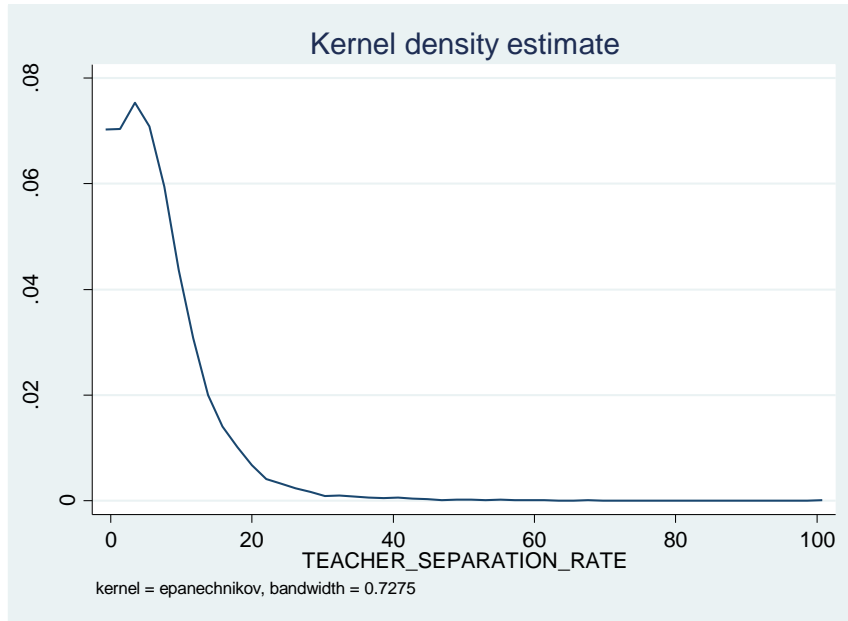


Figure 4

Display of Leftward Skew for Mean Graduation Rates 2010-2014

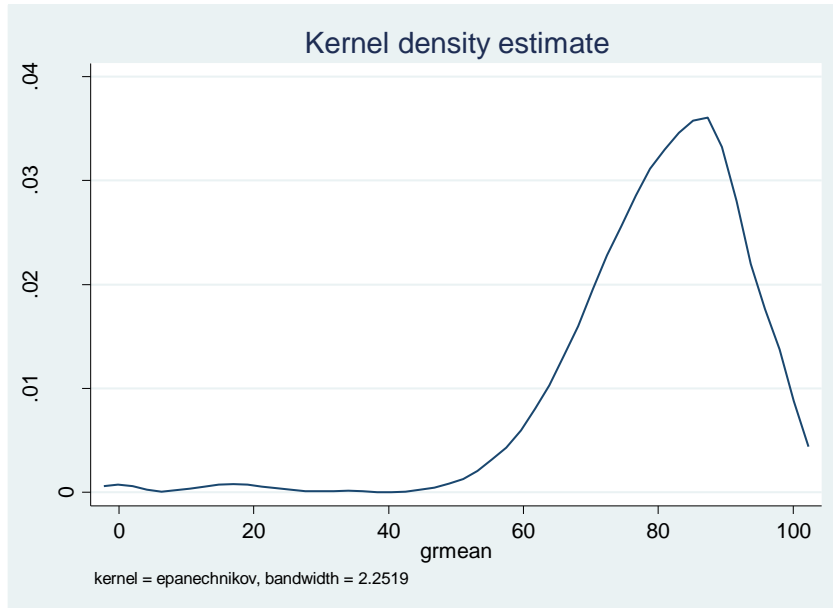


Figure 5

Display of Skew for Transformed (Log) Absence Rate

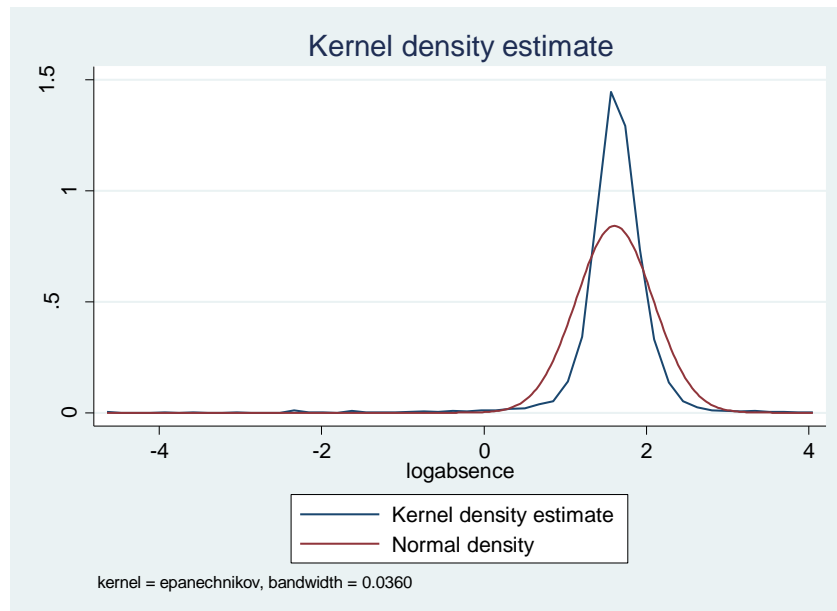


Figure 6

Display of Skew for Transformed (Log) Drop-Out Rate

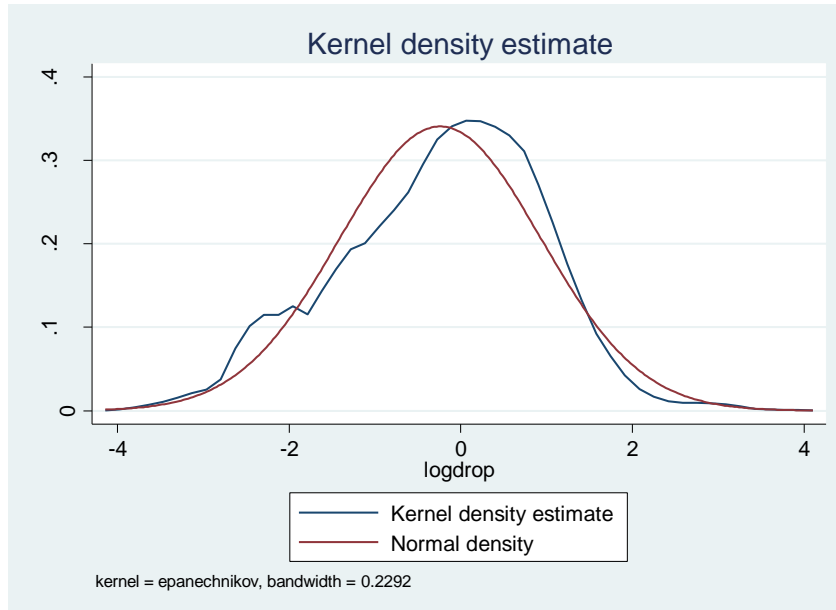


Figure 7

Display of Skew for Transformed (Log) Teacher Separation Rate

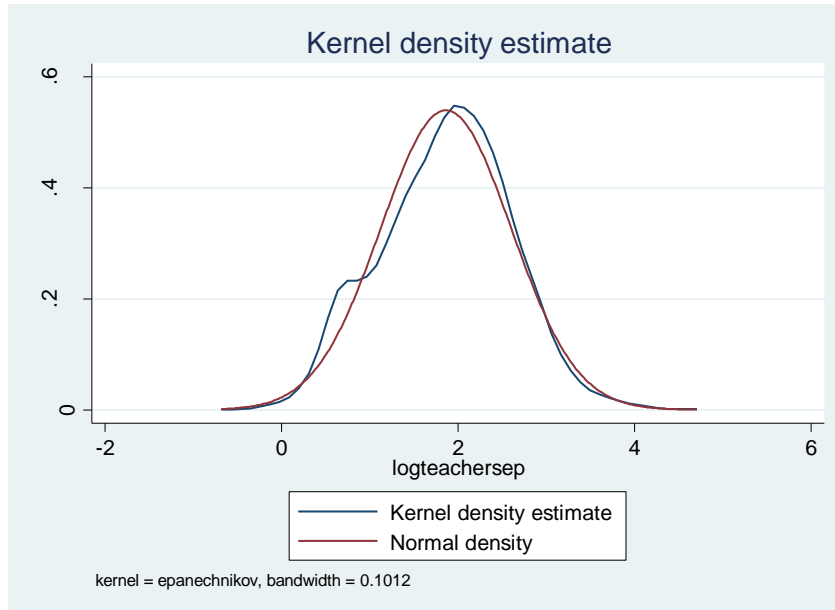


Figure 8

Display of Skew for Transformed (Log) White Failure Rate

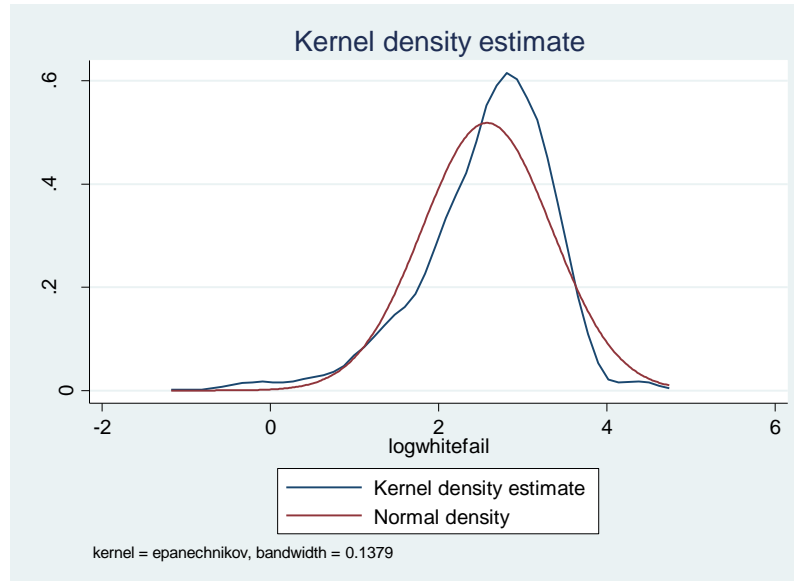


Figure 9

Display of Skew for Transformed (Log) Black Failure Rate

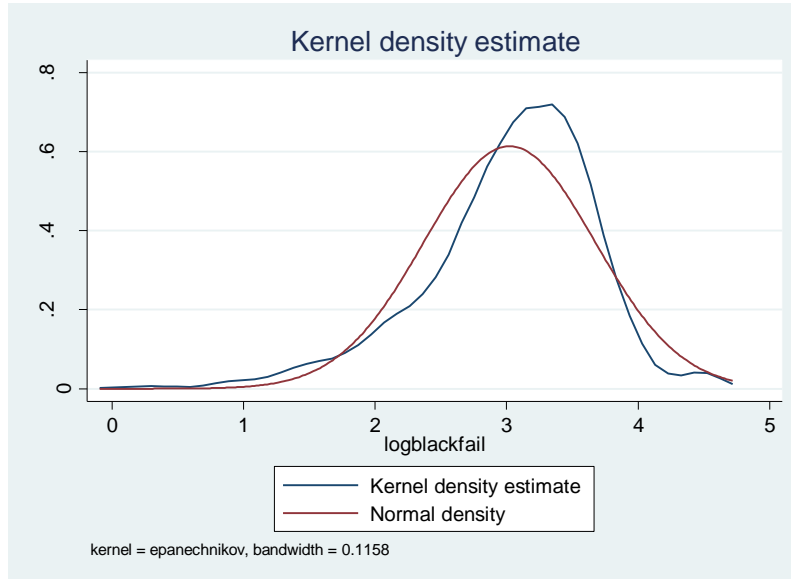


Figure 10

Display of Skew for Transformed (Log) Hispanic Failure Rate

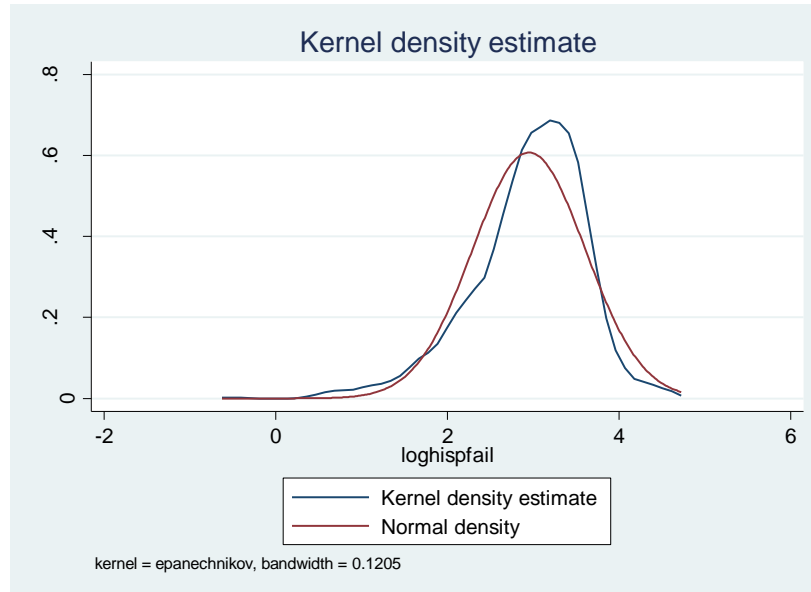
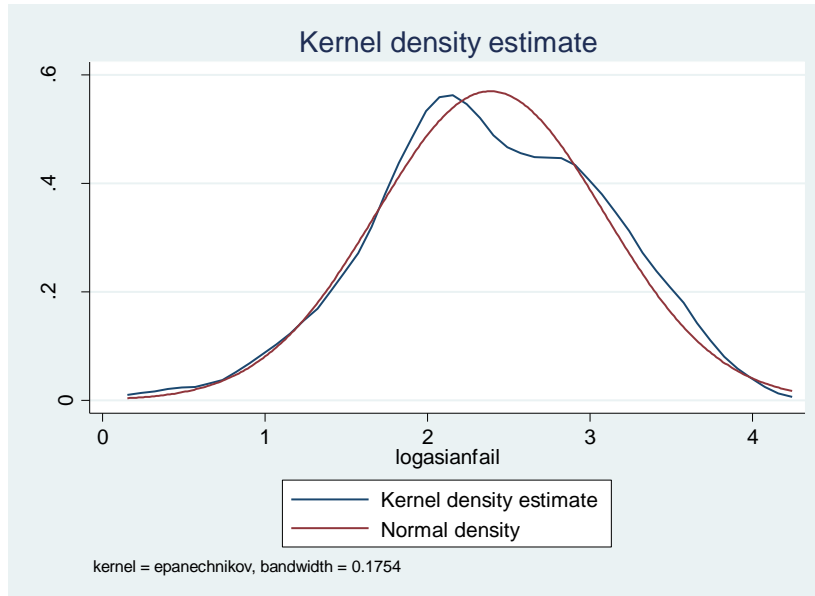


Figure 11

Display of Skew for Transformed (Log) Asian Failure Rate



APPENDIX B(8): RESULTS OF SCHEFFE MULTIPLE COMPARISONS TEST,
TEACHER SEPARATION RATE BY SCHOOL GRADE

Source	Analysis of Variance			F	Prob > F
	SS	df	MS		
Between groups	17065.9258	5	3413.18517	67.54	0.0000
Within groups	766950.484	15177	50.5337342		
Total	784016.41	15182	51.641181		

Chi2(5) = 1.8e+03					
prob>chi2 = 0.0					
School Grade	F	D	C	B	A
D	-2.399 0.0				
C	-3.314 0.0	-0.915 0.0			
B	-3.992 0.0	-1.593 0.0	-0.679 0.0		
A	-4.925 0.0	-2.526 0.0	-1.611 0.0	-0.933 0.0	
Ungraded	-2.485 0.0	-0.086 1.0	0.829 0.3	1.507 0.0	2.440 0.0

APPENDIX C(8): RESULTS OF SCHEFFE MULTIPLE COMPARISONS TEST,
SEPARATED TEACHER EXPERIENCE BY SCHOOL GRADE

Source	Analysis of Variance			F	Prob > F
	SS	df	MS		
Between groups	7776.57177	5	1555.31435	111.69	0.0000
Within groups	208395.895	14965	13.9255526		
Total	216172.467	14970	14.4403785		

Chi2(5) = 148.721					
prob>chi2 = 0.0					
School Grade	F	D	C	B	A
D	.953 0.0				
C	1.460 0.0	0.507 0.0			
B	2.185 0.0	1.232 0.0	0.725 0.0		
A	2.512 0.0	1.559 0.0	1.052 0.0	0.326 0.0	
Ungraded	-0.273 0.96	-1.226 0.0	-1.733 0.0	-2.458 0.0	-2.785 0.0

APPENDIX D(8): RESULTS OF SCHEFFE MULTIPLE COMPARISONS TEST, DROP-OUT RATES BY SCHOOL GRADE

Source	Analysis of Variance			F	Prob > F
	SS	df	MS		
Between groups	2336.36395	5	467.272789	118.98	0.0000
Within groups	9088.11382	2314	3.92744763		
Total	11424.4778	2319	4.92646734		

Chi2(5) = 1.8e+03					
prob>chi2 = 0.0					
School Grade	F	D	C	B	A
D	-1.280 0.13				
C	-1.820 0.0	-0.541 0.22			
B	-2.333 0.0	-1.053 0.0	-0.512 0.0		
A	-2.750 0.0	-1.47 0.0	-0.929 0.0	-0.417 0.0	
Ungraded	3.097 0.0	4.376 0.0	4.917 0.0	5.429 0.0	5.846 0.0

APPENDIX E(8): RESULTS OF SCHEFFE MULTIPLE COMPARISONS TEST,
GRADUATE RATES OF FOUR RACIAL SUBGROUPS, BY SCHOOL GRADE

Figure 1

White Graduation Rates

Source	Analysis of Variance			F	Prob > F
	SS	df	MS		
Between groups	47507.3106	5	9501.46212	84.60	0.0000
Within groups	214505.868	1910	112.306737		
Total	262013.179	1915	136.821503		

Chi2(5) = 199.2828					
prob>chi2 = 0.0					
School Grade	F	D	C	B	A
D	-0.562 1.0				
C	1.612 0.99	2.174 0.74			
B	7.956 0.352	8.517 0.0	6.344 0.0		
A	12.998 0.01	13.560 0.0	11.386 0.0	5.043 0.0	
Ungraded	-7.054 0.63	-6.492 0.10	-8.666 0.0	-15.009 0.0	-20.052 0.0

Figure 2

Black Graduation Rates

Source	Analysis of Variance			F	Prob > F
	SS	df	MS		
Between groups	88942.6663	5	17788.5333	103.42	0.0000
Within groups	299971.404	1744	172.001952		
Total	388914.07	1749	222.363676		

Chi2(5) = 187.2062					
prob>chi2 = 0.0					
School Grade	F	D	C	B	A
D	15.148 0.0				
C	15.908 0.0	0.761 0.99			
B	20.126 0.0	4.978 0.05	4.218 0.0		
A	26.846 0.0	11.698 0.0	10.937 0.0	6.719 0.0	
Ungraded	-10.650 0.16	-25.798 0.0	-26.558 0.0	-30.776 0.0	-37.495 0.0

Figure 3

Hispanic Graduation Rates

Source	Analysis of Variance			F	Prob > F
	SS	df	MS		
Between groups	70608.4092	5	14121.6818	87.00	0.0000
Within groups	274163.366	1689	162.322893		
Total	344771.776	1694	203.525251		

Chi2(5) = 115.5328					
prob>chi2 = 0.0					
School Grade	F	D	C	B	A
D	6.986 0.65				
C	11.252 0.08	4.265 0.224			
B	16.358 0.0	9.371 0.0	5.106 0.0		
A	24.219 0.0	17.232 0.0	12.967 0.0	7.861 0.0	
Ungraded	-2.766 0.99	-9.753 0.01	-14.018 0.0	-19.124 0.0	-26.985 0.0

Figure 4

Asian Graduation Rates

Source	Analysis of Variance			F	Prob > F
	SS	df	MS		
Between groups	5242.18146	5	1048.43629	10.44	0.0000
Within groups	78838.576	785	100.431307		
Total	84080.7574	790	106.431339		

Chi2(5) = 24.6400					
prob>chi2 = 0.0					
School Grade	F	D	C	B	A
D	-7.785 0.90				
C	-5.098 0.98	2.687 0.95			
B	-2.658 0.99	5.127 0.43	2.440 0.49		
A	1.265 1.0	9.050 0.01	3.363 0.0	3.923 0.0	
Ungraded	-0.358 1.0	7.427 0.77	4.740 0.94	2.300 0.99	-1.623 1.0

APPENDIX F(8): TABLE OF SAMPLE MEANS, STANDARD DEVIATIONS,
MINIMUM AND MAXIMUM MEASURES

Variable	Observations	Mean	Std. Deviation	Minimum	Maximum
ADA Rate	15,196	94.581	2.638	44.88	100
White Drop-Out Rate	2,252	1.262	2.042	0	21.2
Black Drop-Out Rate	2,221	1.318	2.479	0	43.6
Hispanic Drop-Out Rate	2,204	1.382	2.595	0	66.7
Asian Drop-Out Rate	1,660	0.478	1.456	0	16.67
Total Drop-Out Rate	2,320	1.300	2.220	0	47.9
White Graduation Rate	1,916	83.766	11.697	0	100
Black Graduation Rate	1,750	76.235	14.912	0	100
Hispanic Graduation Rate	1,695	78.030	14.267	0	100
Asian Graduation Rate	791	89.805	10.317	41.667	100
Total Teachers	15,183	55.542	29.423	1	306
Teacher Separation Rate	15,183	6.832	7.186	0	100
Separated Teacher Experience (in years)	14,971	10.712	3.800	0	27
Log White Failure Rate	1,854	2.576	0.769	-1.040	4.605
Log Black Failure Rate	1,690	3.022	0.650	0.020	4.605
Log Hispanic Failure Rate	1,616	2.952	0.657	-0.504	4.605
Log Asian Failure Rate	587	2.389	0.700	0.329	4.066
Log Teacher Separation	12,293	1.864	0.739	-0.580	4.605

*Rates are listed in terms of percent, out of 100.

REFERENCES

- Adler, F. (1999, July 2). Our schools need to stay focused on education [Letter to the editor]. *St. Petersburg Times*. Retrieved from <https://www.tampabay.com>
- Agee, J. (2004). Negotiating a Teaching Identity: An African American Teacher's Struggle to Teach in Test-Driven Contexts. *Teachers College Record*, 106(4), 747-774. doi: 10.1111/j.1467-9620.2004.00357.x
- Akkaraju, S., Atamturktur, S., Broughton, L., & Frazer, T. (2019). Ensuring Student Success: Using Formative Assessment as the Key to Communication and Compassion Among Faculty, Students, and Staff. *New Directions for Community Colleges*, 2019(186), 71-79. doi: 10.1002/cc.20358
- Alispahić, S. (2013). Motivational Function of Plans and Goals. *Psychological Thought*, 6(2), 196-203. doi: 10.5964/psyc.v6i2.80
- Allison, P. D. (2005). Fixed Effects Regression Methods for Longitudinal Data Using SAS. SAS Institute. March 2005.
- Alvarez, L. (2014, November 9). States listen as parents give rampant testing an F." *New York Times*. Retrieved from <https://www.nyt.com>
- Amrein, A. L., & Berliner, D. C. (2002). An Analysis of Some Unintended and Negative Consequences of High-Stakes Testing. *Education Policy Research Unit, Arizona State University*.
- Anyon, Y., Jenson, J. M., Altschul, I., Farrar, J., McQueen, J., Greer, E., Simmons, J. (2014). The persistent effect of race and the promise of alternatives to suspension in school discipline outcomes. *Children & Youth Services Review*, 44, 379-386. doi: 10.1016/j.childyouth.2014.06.025
- Archibald, S. (2006). Narrowing in on educational resources that do affect student achievement. *Peabody Journal of Education*, 81(4), 23-42.
- Au, W. (2010). The Idiocy of Policy: The anti-democratic curriculum of high-stakes testing. *Critical Education*, 1(1).
- Au, W. W. (2009). High-stakes testing and discursive control: The triple bind for non-standard student identities. *Multicultural Perspectives*, 11(2), 65-71. doi: 10.1080/15210960903028727
- Ballou, D., Sanders, W., & Wright, P. (2004). Controlling for student background in value-added assessment of teachers. *Journal of Educational and Behavioral Statistics*, 29(1), 37-65.
- Ballou, D., & Springer, M. G. (2015). Using student test scores to measure teacher performance: Some problems in the design and implementation of evaluation systems. *Educational Researcher*, 44(2), 77-86.
- Barksdale-Ladd, M. A., & Thomas, K. F. (2000). What's at stake in high-stakes testing; Teachers and parents speak out. *Journal of Teacher Education*, 51, 384-397.
- Barrett, L. (2013, February 27). More districts try innovation schools. National Education Association. Retrieved from <https://www.nea.org>
- Barshay, J. (2015). The gap between rich and poor schools grew 44 percent over a decade. *The Hechinger Report, Education by the Numbers: April 6, 2015.*, Retrieved From:

- <http://hechingerreport.org/the-gap-between-rich-and-poor-schools-grew-44-percent-over-a-decade/>.
- Bates, L. A., & Glick, J. E. (2013). Does it matter if teachers and schools match the student? Racial and ethnic disparities in problem behaviors. *Social Science Research*, 42(5), September, 2013.
- Berends, M., & Koretz, D. M. (1995). Reporting minority students' test scores: How well can the national assessment of educational progress account for differences in social context? *Educational Assessment*, 3(3), 249.
- Berkowitz, R., Moore, H., Astor, R. A., & Benbenishty, R. (2017). A research synthesis of the associations between socioeconomic background, inequality, school climate, and academic achievement. *Review of Educational Research*, 87(2), 425-469.
- Betebenner, D. (2008). Book Review: Feeling the Florida Heat? How Low-Performing Schools Respond to Voucher and Accountability Pressure *Urban Institute's National Center for Analysis of Longitudinal Data in Education Research*, Review.
- Blaich, C. F., & Wise, K. S. (2010). Moving from assessment to institutional improvement. *New Directions for Institutional Research*, 2010, 67-78. doi: 10.1002/ir.373
- Booher-Jennings, J. (2005). "Below the bubble: 'Educational triage' and the Texas Accountability System." *American Educational Research Journal* Summer 2005, Vol. 42, No. 2, pp. 231-268.
- Borges, N. (2013, July 31). Why school grades in Florida are full of controversy. NPR. Retrieved from <https://www.wlrn.org>
- Borman, K.M., & Dorn, S. (Eds.). (2007). *Education reform in Florida: Diversity and equity in public policy*. New York, NY: State University of New York Press.
- Borman, K. M., & Dorn, S. (2007). Introduction: Issues in Florida education reform. In Kathryn M. Borman, Sherman Dorn (Eds.), *Education reform in Florida: Diversity and equity in public policy* (pp. 1-18). New York, NY: State University of New York Press.
- Bowen, D. H., & Trivitt, J. R. (2014). Stigmas without sanctions: The (lack of) impact of private school vouchers on student achievement. *Education Policy Analysis Archives*, 22(87).
- Boykin, A. W., & Noguera, P. A. (2011). *Creating the opportunity to learn: Moving from research to practice to close the achievement gap*. Alexandria, VA: ASCD.
- Brenneman, R. (2016, March 9). Florida lawmakers renew teacher bonus-pay program over bipartisan objections. *Education Week*. Retrieved from <https://www.edweek.org>
- Brevetti, M. (2014). Reevaluating narrow accountability in american schools: The need for collaborative effort in improving teaching performances. *Delta Kappa Gamma Bulletin*, 81(1), 32-35.
- Burns, E. C., Martin, A. J., & Collie, R. J. (2019). Understanding the role of personal best (PB) goal setting in students' declining engagement: A latent growth model. *Journal of Educational Psychology*, 111(4), 557-572. doi: 10.1037/edu0000291
- Burrell, T. (2010). *Brainwashed: Challenging the myth of black inferiority*. New York, NY: Smiley Books.

- Bush, J. (2018, April 18). Florida's rising NAEP scores show education reforms are working for all students in the state. *The 74 Million*. Retrieved from <https://www.the74million.org>
- Cadima, J., Peixoto, C., & Leal, T. (2014). Observed classroom quality in first grade: Associations with teacher, classroom, and school characteristics. *European Journal of Psychology of Education*, 29, 139-158.
- Camera, L. (2017, June 15). DeVos racks up more than a dozen visits to Florida schools." *US News*. Retrieved from <https://www.usnews.com>
- Cannon, C.M. (2019, October 2). RCP Poll: K-12 Education Falls Short, and Hope for gains Lags. Real Clear Politics. Retrieved from <http://www.realclearpolitics.com>
- Carr, P. B., & Steele, C. M. (2009). Stereotype threat and inflexible perseverance in problem solving. *Journal of Experimental Social Psychology*, 45(4), 853-859. doi: 10.1016/j.jesp.2009.03.003
- Carver-Thomas, D. and Darling-Hammond, L. (2017, August 16). Teacher turnover: Why it matters and what we can do about it. *Learning Policy Institute*. Retrieved from <https://www.learningpolicyinstitute.org>
- Chakrabarti, R. (2013). Vouchers, public school response, and the role of incentives: Evidence from Florida. *Economic Inquiry*, 51(1), 500-526. doi: 10.1111/j.1465-7295.2012.00455.x
- Cobb-Roberts, D., & Shircliffe, B. (2007). *The legacy of desegregation in Florida*. In Kathryn M. Borman, Sherman Dorn (Eds.), *Education reform in Florida: Diversity and equity in public policy* (pp. 19-52). New York, NY: State University of New York Press.
- Collinson, V., Cook, T. F., & Conley, S. (2006). Organizational learning in schools and school systems: Improving learning, teaching, and leading. *Theory Into Practice*, 45(2), 107-116. doi: 10.1207/s15430421tip4502_2
- Conley, D. (2001). A Room with a view or a room of one's own? Housing and social stratification. *Sociological Forum*, 16(2).
- Creed, D., Hudson, B., Okhuysen, G., & Smith-Crowe, K. (2014). Swimming in a sea of shame: Incorporating emotion into explanations of institutional reproduction and change. *Academy of Management Review*. doi: 10.5465/amr.2012.0074
- Creswell, J. W., & Miller, D. (2002). Determining validity in qualitative inquiry. *Theory into Practice*, 39(3).
- Davis, J. W. (2015). Gov. Scott orders less testing in Florida schools. *MyNews13 February 18, 2015*. Retrieved From: http://www.mynews13.com/content/news/cfnews13/news/article.html/content/news/articles/cfn/2015/2/18/fl_fewer_tests.html.
- De Witte, K., & Csillag, M. (2014). Does anybody notice? On the impact of improved truancy reporting on school dropout. *Education Economics*, 22(6), 549-568. doi: 10.1080/09645292.2012.672555
- DiCarlo, M. (2017, July 13). For Florida's school grading system, a smart change with unexpected effects. *Albert Shanker Institute*. Retrieved from <http://www.Shankerinstitute.org>

- DiCarlo, M. (2014, November 3). The bewildering arguments underlying Florida's fight over ELL test scores. *Albert Shanker Institute*. Retrieved from <http://www.Shankerinstitute.org>
- Dittmer, J. (2004). Assessing school assessment. *Southeastern Geographer*, 44(2), 170-185. doi: 10.1353/sgo.2004.0021
- Downer, J. T., Goble, P., Myers, S. S., & Pianta, R. C. (2016). Teacher-child racial/ethnic match within pre-kindergarten classroom and children's early school adjustment. *Early Childhood Research Quarterly*, 37, 4th Quarter 2016.
- Dreilinger, D. (2016, February 22). Louisiana voucher students did worse at new schools, study says. *New Orleans Times Picayune (NOLA)*. Retrieved from <http://www.nola.com>
- Dynarski, M. (2018, May 3). Is the high school graduation rate really going up? Brookings. Retrieved from www.brookings.edu
- Education Next (2014, November 12). Jeb Bush speaks out. *Education Next*. Retrieved from <https://educationnext.org>
- Egalite, A. J., Kisida, B., & Winters, M. A. (2015). Representation in the classroom: the effect of own-race teachers on student achievement. *Economics of Education Review*, 45, April 2015.
- Elmore, R. (2004). *School reform from the inside out: Policy, practice, and performance*. Cambridge, MA: Harvard University Press.
- Elmore, R. F., & Fuhrman, S. H. (2001). Holding schools accountable: Is it working? *Phi Delta Kappan*, 83(1), 67. doi: 10.1177/003172170108300114
- Erbe, B. M. (2000). Correlates of school achievement in Chicago elementary schools. *Paper presented at the Annual Meeting of the American Educational Research Associations*, New Orleans, Louisiana.
- Failure factories: The story of how one school board turned five average schools into some of the worst in Florida (2015). *The Tampa Bay Times*. Retrieved from <https://www.tampabay.com>
- Feinberg, C. (2018, August 27). Why is Florida the only state that's not in compliance with federal education standards? *The Miami Herald*. Retrieved from <https://www.miamiherald.com>
- Ferguson, R. F. (2008). *Toward excellence with equity: An emerging vision for closing the achievement gap*. Boston, MA: Harvard Education Press.
- Figlio, D. N., & Lucas, M. E. (2004). What's in a grade? School report cards and the housing market. *The American Economic Review*, 94(3), 591-604.
- Figlio, D. N., & Rouse, C. E. (2005). Do accountability and voucher threats improve low-performing schools? *Journal of Public Economics*, 90((2006) 239-255).
- Firestone, W. A. (2014). Teacher evaluation policy and conflicting theories of motivation. *Educational Researcher*, 43(2), 100-107.
- Florida Department of Education (2018, February). Florida's High School Cohort 2016-17 Dropout Rate. Retrieved from <https://www.fldoe.org>

- Florida Department of Education (2017). 2017-2018 Guide to calculating school grades, district grades, and federal percent of points index. Retrieved from <https://www.fldoe.org>
- Foundation for Excellence in Education. (2012). Florida's Education Revolution: Accountability, Choice and High Expectations Create Student Success. www.excelined.org.
- Frontline (2019, June 18). The Reality of School Grades. June 18, 2019. Frontline News. Retrieved from <https://www.feaweb.org>
- Fullan, M., & Quinn, J. (2016). *Coherence: The right drivers in action for schools, districts, and systems*. Thousand Oaks, CA: Corwin: A Sage Company.
- Gagnon, D. J., & Schneider, J. (2019). Holistic school quality measurement and the future of accountability: Pilot-test results. *Educational Policy*, 33(5), 734-760. doi: 10.1177/0895904817736631
- Gartner, L. (2013, August 28). Summit's message to Gov. Scott: Change the school grading system." *Tampa Bay Times*. Retrieved from <https://www.tampabay.com>
- Gershenson, S., Holt, S. B., & Papageorge, N. W. (2016). Who believes in me? The effect of student-teacher demographic match on teacher expectations. *Economics of Education Review*, 52, June 2016.
- Goldhaber, D. (2015). Exploring the Potential of Value-Added Performance Measures to Affect the Quality of the Teacher Workforce. *Educational Researcher*, 44(2), 87-95.
- Goldhaber, D., & Hannaway, J. (2004). Accountability with a kicker: Observations on the Florida A+ accountability plan. *Phi Delta Kappan*, 85(8), 598-605.
- Goldschmidt, P., Roschewski, P., Choi, K., Auty, W., Hebbler, S., Blank, R. K., & Williams, A. (2005). Policymakers' Guide to Growth Models for School Accountability: How do Accountability Models Differ? *CCSSO Accountability Systems and Reporting, State Collaborative on Assessment and Student Standards*.
- Goldstein, D. (2014). *The teacher wars: A history of america's most embattled profession*. New York, New York: Doubleday.
- Gonzalez, S. (2012, April 16). Grading Florida schools: Opportunities lag for rural students." *State Impact, NPR*, Retrieved from <https://www.stateimpact.npr.org>
- Gosa, T. L., & Alexander, K. L. (2007). Family (dis)advantage and the educational prospects of better off African American youth: How race still matters. *Teachers College Record*, 109(2), 285-321.
- Greene, J. P. (2001). An Evaluation of the Florida A-Plus Accountability and School Choice Program. *Center for Civic Innovation at the Manhattan Institute, Program on Education Policy and Governance, Harvard University*.
- Greifner, L. (2013). "Suspensions Kept Fla. Low Scorers Out of Tests, Study Says." *Education Week*, 25(41), pp. 30-30.
- Grissom, J. A., Loeb, S., & Doss, C. (2015). The multiple dimensions of teacher quality: Does value-added capture teachers' nonachievement contributions to their schools? In: J.A. Grissom and P. Young (Eds.), *Improving teacher evaluation systems: Making the most of multiple measures* (pp. 37-50). New York, NY: Teachers College Press.

- Grodsky, E., Warren, J. R., & Felts, E. (2008). Testing and social stratification in American education. *Annual Review of Sociology*, 34(1), 385-404. doi: doi:10.1146/annurev.soc.34.040507.134711
- Hagerty, S. (1999, May 17). School grading draws criticism." *Tampa Bay Times*. Retrieved from <http://www.tampabay.com>
- Hamilton, L. (2003). Assessment as a policy tool. *Review of Research in Education*, 27, 25-68. doi: 10.2307/3568127
- Harlen, W., & Crick, R. D. (2003). Testing and motivation for learning. *Assessment in Education: Principles, Policy & Practice*, 10(2), 169. doi: 10.1080/0969594032000121270
- Harris, D. N., & Herrington, C. D. (2015). The use of teacher value-added measures in schools: New evidence, unanswered questions, and future prospects. *Educational Researcher*, 44(2), 71-76.
- Harris, M. (2017). *Kids these days: Human capital and the making of millennials*. New York, NY: Little, Brown and Company.
- Harrison, C., & Cohen-Vogel, L. (2012). The politics of teacher reform in Florida: Analyzing causal narratives surrounding state adoption of performance-based evaluations, performance pay, and tenure elimination. *Peabody Journal of Education* (0161956X), 87(5), 517-534. doi: 10.1080/0161956X.2012.723489
- Haughbrook, R., Hart, S. A., Schatschneider, C., & Taylor, J. (2017). Genetic and environmental influences on early literacy skills across school grade contexts. *Developmental Science*, 20(5), n/a-N.PAG. doi: 10.1111/desc.12434
- Hegarty, S. (1999, May 17). School grading draws criticism. *St. Petersburg Times*. Retrieved from <https://www.tampabay.com>
- Herrington, C. D. (2005). Educational accountability and assessment in Florida: Where have we been? Where are we going? *Florida Institute of Education at the University of North Florida, FIE Policy Brief 1*(June 2005).
- Hopkins, M., Thompson, K. D., Linquanti, R., Hakuta, K., & August, D. (2013). Fully accounting for English learner performance: A key issue in ESEA reauthorization. *Educational Researcher*, 42(2), 101-108.
- Jackson, C. K. (2016). What do test scores miss? The importance of teacher effects on non-test score outcomes (Working Paper No. 22226). *National Bureau of Economic Research, Cambridge, MA*.
- Jacob, B. A. (2004). Public housing, housing vouchers, and student achievement: Evidence from public housing demolitions in Chicago. *The American Economic Review*, 94(1), 233-258. doi: 10.2307/3592777
- Jacobsen, R., Snyder, J. W., & Saultz, A. (2014). Informing or shaping public opinion? The influence of school accountability data format on public perceptions of school quality. *American Journal of Education*, 121(1 (November 2014)), 1-27.
- Jiang, J. Y., Sporte, S. E., & Luppescu, S. (2015). Teacher perspectives on evaluation reform: Chicago's REACH students. *Educational Researcher*, 44(2), 105-116.
- Jimenez-Castellanos, O. (2010). Relationship between educational resources and school achievement: A mixed method intra-district analysis. *Urban Review*, 42, 351-371.

- Johnson, S. M. (2015). Will VAMS reinforce the walls of the egg-crate school? *Educational Researcher*, 44(2), 117-126.
- Johnson, S. M., Kraft, M. A., & Papay, J. P. (2012). How context matters in high-need schools: The effects of teachers' working conditions on their professional satisfaction and their students' achievement. *Teachers College Record*, 114(10), 1-39.
- Jones, G. M., Jones, B. D., & Hargrove, T. Y. (2007). The unintended outcomes of high-stakes testing. *Journal of Applied School Psychology*, 23(2), 65-86. doi: 10.1300/J370v23n02_05
- Jones, N. D., Buzick, H. M., & Turkan, S. (2013). Including students with disabilities and english learners in measures of educator effectiveness. *Educational Researcher*, 42(4), 234-241.
- Kantor, H., & Lowe, R. (2013). Educationalizing the welfare state and privatizing education: The evolution of social policy since the new deal. In Carter, P.L. and Welner, K.G. (Ed.) *Closing the Opportunity Gap* (pp. 1-10). New York, New York: Oxford University Press.
- Klein, A. (2016, March 31). The Every Student Succeeds Act: An ESSA overview. *Education Week*. Retrieved from <https://www.edweek.org>
- King, G., Keohane, R. O., & Verba, S. (1994). *Designing social inquiry: Scientific inference in qualitative research*. Princeton, NJ: Princeton University Press.
- Kurtz, M. D. (2018). Value-added and student growth percentile models: What drives differences in estimated classroom effects? *Statistics & Public Policy*, 5(1), 1-8. doi: 10.1080/2330443X.2018.1438938
- Ladd, H. F., & Lauen, D. L. (2010). Status versus growth: The distributional effects of school accountability policies. *Journal of Policy Analysis and Management*, 29(3 (Summer 2010)), 426-450.
- Lawrence, J. S., Marks, B. T., & Jackson, J. S. (2010). Domain identification predicts black students' underperformance on moderately-difficult tests. *Motivation & Emotion*, 34(2), 105-109. doi: 10.1007/s11031-010-9159-8
- Lee, R. S., Borman, K. M., & Tyson, W. (2007). *Florida's A+ Plan: Education reform policies and student outcomes*. Borman, K. and Dorn, S. (Eds.). New York, NY: State University of New York Press.
- Linn, R. L. (2008). Methodological issues in achieving school accountability. *Journal of Curriculum Studies*, 40(6), 699-711. doi: 10.1080/00220270802105729
- Loubert, L. (2005). Housing markets and school financing. *Journal of Education Finance*, 30(4), 412-429. doi: 10.2307/40704244
- Martin, A. and Postal, L. (2018, April 25). Florida probes 3 private schools that hired felon teachers. *The Orlando Sentinel*. Retrieved from <https://www.orlandosentinel.com>
- Mack, S. (2014, May 12). Remembering Early Desegregation in Florida. *State Impact, NPR*. Retrieved from <http://stateimpact.npr.org>
- Mack, S. (2013, August 14). What a school grade means to parents. *State Impact, NPR*. Retrieved from <http://stateimpact.npr.org>
- Maxwell, J. A. (1996). Qualitative Research Design. *Applied Social Research Methods Series*, 41(Sage Publications).

- Maxwell, L. A. (2012). School Grading Changes Spark Outcry in Florida. *Education Week*, 31(32), 19-20.
- Maxwell, S (2019, August 13). 3,500 teacher vacancies in Florida. This is what happens when you abuse public education. *The Orlando Sentinel*. Retrieved from <https://www.orlandosentinel.com>
- Maxwell, S. (2017, March 26). Florida needs to listen to its teachers. *The Orlando Sentinel*. Retrieved from <https://www.orlandosentinel.com>
- Maxwell, S. (2017, March 24). With teachers fleeing, Florida must start listening.” *The Orlando Sentinel*. Retrieved from <https://www.orlandosentinel.com>
- McKay, A. (2018, May 29). The problem with ‘Great Schools.’ *Medium Education*. Retrieved from <https://www.medium.com>
- McMahon, R. (1999, August 16). State officials didn’t play fair in grading schools [Letter to the editor]. *St. Petersburg Times*. Retrieved from <https://www.tampabay.com>
- Mehmetoglu, M., & Jakobsen, T. G. (2017). *Applied statistics using STATA: A guide for the social sciences*. London, England: Sage Publications Ltd.
- Michael, D. L., & Dorn, S. (2007). *Accountability as a means of improvement*. Borman, K. and Dorn, S. (Eds.). New York, NY: State University of New York Press.
- Miggins, E. M. (2014). “No Crystal Stair”: The Cleveland Public Schools and the Struggle for Equality, 1900–1930 (Vol. 40, pp. 671-698).
- Miller, J. J. (2011). The Education Ex-Governor. (Cover story). *National Review*, 63(3), 32-36.
- Minnich, C. (2018, December 13). What Detroit and Chicago can teach us about student growth and school rating systems.” Learning Policy Institute. Retrieved from <http://www.learningpolicyinstitute.org>
- Mitchell, K. (1999, July 2). Our schools need to stay focused on education [Letter to the editor]. *St. Petersburg Times*. Retrieved from <https://www.tampabay.com>
- Moreno, I. (2017, December 3). US charter schools put growing numbers in racial isolation. Retrieved from <https://www.ap.org>
- Morris, M. W. (2016). *Pushout: The Criminilization of Black Girls in Schools*. New York, NY: The New Press.
- Murnane, R. J. (2007). Improving the education of children living in poverty. *The Future of Children*, 17(2), 161-182. doi: 10.2307/4495065
- Murphy, Paul (2018, July 3). Teachers are tired of Robert Marzano. *Teacher Habits*. Retrieved from <https://www.teacherhabits.com>
- Noguera, P. A., & Wing, J. Y. (2006). *Unfinished business: Closing the racial achievement gap in our schools*. San Francisco, CA: Jossey Bass.
- Oberle, M. (1999, July 2). Our schools need to stay focused on education [Letter to the editor]. *St. Petersburg Times*. Retrieved from <https://www.tampabay.com>
- Orfield, G. (2013). Housing segregation produces unequal schools: Causes and solutions. In Carter, P.L. and Welner, K.G. (Ed.) *Closing the Opportunity Gap* (pp. 40-60). New York, New York: Oxford University Press.

- Orfield, G., & Ee, J. (2017). Tough Choices Facing Florida's Governments: Patterns of Resegregation in Florida's Schools. *A Report for the LeRoy Collins Institute, Florida State University*. September 27, 2017.
- Osborne, J. W. (2001). Testing stereotype threat: Does anxiety explain race and sex differences in achievement? *Contemporary Educational Psychology*, 26(3), 291-310. doi: 10.1006/ceps.2000.1052
- Paige, M. (2013). Applying the “paradox” theory: A law and policy analysis of collective bargaining rights and teacher evaluation reform from selected states. *Brigham Young University Education & Law Journal*(1), 21-43.
- Parke, C. S., & Kanyongo, G. Y. (2012). Student attendance, mobility, and mathematics achievement in an urban school district. *Journal of Educational Research*, 105(3), 161-175. doi: 10.1080/00220671.2010.547231
- Pear, R. (2004, February 24). Education chief calls union “terrorist,” then recants. *The New York Times*. Retrieved from <https://www.nyt.com>
- Perry, B. L., & Morris, E. W. (2014). Suspending progress: Collateral consequences of exclusionary punishment in public schools. *American Sociological Review*, 79(6), 1067-1087. doi: 10.1177/0003122414556308
- Petek, N., & Pope, N. G. (2016). The multidimensional impact of teachers on students (Working Paper). Retrieved From: http://home.uchicago.edu/~npope/Nolan_Pope_JMP.pdf.
- Porfeli, E., Chuang, W., Audette, R., McColl, A., & Algozzine, B. (2009). Influence of social and community capital on student achievement in a large urban school district. *Education & Urban Society*, 42(1), 72-95.
- Postal, L. (2012, January 30). Educators criticize latest Florida school rankings. *The Orlando Sentinel*. Retrieved from <https://www.orlandosentinel.com>
- Postal, L., Kassab, B. and Martin, A. (2018, June 1). Schools without rules: Private schools’ curriculum downplays slavery, says humans and dinosaurs lived together. *The Orlando Sentinel*. Retrieved from <https://www.orlandosentinel.com>
- Prothero, A. (2017). In Florida, laissez-faire approach to monitoring private school vouchers. *Education Week*, 37(13), 1-13.
- Ravitch, D. (2010). Stop the madness. *Education Review*, 23(1), 27-34.
- Ready, D. D. (2013). Associations between student achievement and student learning: Implications for value-added school accountability models. *Educational Policy*, 27(1), 92-120. doi: 10.1177/0895904811429289
- Reingold, D. A., Ryzin, G. G. V., & Ronda, M. (2001). Does urban public housing diminish the social capital and labor force activity of its tenants? *Journal of Policy Analysis and Management*, 20(3), 485-504. doi: 10.2307/3326133
- Robinson, S.J. (2002, June 18). FCAT ignores key factors in school success [Letter to the editor]. *Tampa Bay Times*. Retrieved from <http://www.tampabay.com>
- Rossmann, G. B., & Rallis, S. F. (2003). *Learning in the field: An introduction to qualitative research, (Second Edition)*. Thousand Oaks, CA: Sage Publications.
- Rouse, C. E., Hannaway, J., Goldhaber, D., & Figlio, D. (2007). Feeling the Florida Heat? How Low-Performing Schools Respond to Voucher and Accountability Pressure.

- American Economic Journal: Economic Policy*, 5(2), 251-281. doi: doi: 10.1257/pol.5.2.251
- Rubin, H. J., & Rubin, I. S. (2012). *Qualitative interviewing: The art of hearing data (Third Edition)*. Thousand Oaks, CA: Sage.
- Sampson, H., Silva, C. & Marshall, T. (2010, April 15). Gov. Crist Vetoes Teacher Tenure Bill. *Tampa Bay Times*. Retrieved from <https://www.tampabay.com>
- Schneider, J. (2017). *Beyond test scores: A better way to measure school quality*. Cambridge, MA: Harvard University Press.
- Schunk, D. H., Meece, J. R., & Pintrich, P. R. (2013). *Motivation in education: Theory, research, and applications*. Boston, MA: Pearson.
- Schutt, R. K. (2012). *Investigating the social world: The process and practice of research (7th Edition)*. Thousand Oaks, CA: Sage.
- Schwartz, A. E., McCabe, B. J., Ellen, I. G., & Chellman, C. C. (2010). Public Schools, Public Housing: The Education of Children Living in Public Housing. *Urban Affairs Review*, 46(1), 68-89. doi: 10.1177/1078087410367780
- Shannon, E. W., & Saatcioglu, A. (2018). School characteristics and teacher policy influence: Evidence from New York City. *Educational Policy*, 32(7), 1041-1069.
- Sherman, A. (2017, August 10). Scott delivered on promise to study standardized tests. Politifact. Retrieved from <https://www.politifact.com>
- Shih, M., & Pittinsky, T. L. (1999). Stereotype susceptibility: Identity salience and shifts in quantitative performance. *Psychological Science (Wiley-Blackwell)*, 10(1), 80.
- Shubart, N. (2002, June 18). FCAT ignores key factors in school success [Letter to the editor]. *Tampa Bay Times*. Retrieved from <http://www.tampabay.com>
- Skiba, R. J., Arredondo, M. I., & Williams, N. T. (2014). More than a metaphor: The contribution of exclusionary discipline to a school-to-prison pipeline. *Equity & Excellence in Education*, 47(4), 546-564. doi: 10.1080/10665684.2014.958965
- Smith, B.S. (2002, June 18). FCAT ignores key factors in school success [Letter to the editor]. *Tampa Bay Times*. Retrieved from <http://www.tampabay.com>
- State Impact. (2014). How Does the State Grade Schools? *State Impact, NPR*. Retrieved From: <https://stateimpact.npr.org>
- Steele, C. M. (2010). *Whistling Vivaldi: How stereotypes affect us and what we can do*. New York, NY: Norton.
- Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality & Social Psychology*, 69(5), 797-811.
- Stephenson, H. (2017, October 14). Arizona schools could get a do-over on their A-through-F grades. *Arizona Daily Star*. Retrieved from <https://tucson.com>
- Strauss, V. (2017, June 15). Florida governor is signing K-12 public education bill that critics call a “scam” at a Catholic school. *The Washington Post*. Retrieved from <https://www.washingtonpost.com>
- Strauss, V. (2013, Feb 24). Does school reform perpetuate inequity?. *The Washington Post*. Retrieved from <https://www.washingtonpost.com>
- Stretesky, P. B., & Lynch, M. J. (2002). Environmental hazards and school segregation in Hillsborough County, Florida 1987-1999. *Sociological Quarterly*, 43(4), 553. doi: 10.1111/j.1533-8525.2002.tb00066.x

- Sullivan, A. L., Van Norman, E. R., & Klingbeil, D. A. (2014). Exclusionary discipline of students with disabilities: Student and school characteristics predicting suspension. *Remedial & Special Education, 35*(4), 199-210. doi: 10.1177/0741932513519825
- Tampa Bay Times Editorial Board (2019, January 9). Change formula for grading schools: Florida's school grades ignore too many groups of students. *Tampa Bay Times*. Retrieved from <https://www.tampabay.com>
- Taras, M. (2005). Assessment: Summative and formative: Some theoretical reflections. *British Journal of Educational Studies, 53*(4), 466-478.
- The reality of school grades. (2019, June 18), Retrieved from <https://feaweb.org>
- Tienken, C. H., & Zhao, Y. (2013). How Common Standards and Standardized Testing Widen the Opportunity Gap. In Carter, P.L. and Welner, K.G. (Ed.) *Closing the Opportunity Gap* (pp. 111-122). New York, New York: Oxford University Press.
- Toppo, G. (2012, October 11). Florida's schools' race-based plan draws criticism. *USA Today*. Retrieved from <https://www.usatoday.com>
- Tyson, K. (2013). Tracking, Segregation, and the Opportunity Gap. In Carter, P.L. and Welner, K.G. (Ed.) *Closing the Opportunity Gap* (pp. 1-10). New York, New York: Oxford University Press.
- Ujifusa, A. (2012). Waiver plans push school grading systems. *Education Week, 31*(30), 21-23.
- Valant, J. (2017, April 18). As states submit ESSA plans, policymakers must design responsible school rating systems. Brookings Institute, <https://www.brookings.edu>
- Van de Grift, W. (2009). Reliability and validity in measuring the value added of schools. *School Effectiveness & School Improvement, 20*(2), 269-285. doi: 10.1080/09243450902883946
- Van Loo, K. J., Boucher, K. L., Rydell, R. J., & Rydell, M. T. (2013). Competition in stereotyped domains: Competition, social comparison, and stereotype threat. *European Journal of Social Psychology, 43*(7), 648-660. doi: 10.1002/ejsp.1977
- Veiga, C. (2015, March 1). Controversial standardized tests debut in Florida schools. *The Miami Herald*. Retrieved from <https://www.miamiherald.com>
- Vogell, H. and Fresques, H. (2017, February 21). "Alternative" education: Using charter schools to hide dropouts and game the system. Propublica, USA Today. Retrieved from <https://www.usatoday.com>
- Waite, M. (2004, April 18). Politics of F-Cat shadow election. *St. Petersburg Times*. Retrieved from: <http://www.tampabay.com>
- Weiss, R. S. (1994). *Learning from strangers: The art and method of qualitative interview studies*. New York, NY: Free Press.
- Welner, K. G., & Carter, P. L. (2013). Achievement Gaps Arise from Opportunity Gaps. In Carter, P.L. and Welner, K.G. (Ed.) *Closing the Opportunity Gap* (pp. 1-10). New York, NY: Oxford University Press.
- Wout, D., Danso, H., Jackson, J., & Spencer, S. (2008). The many faces of stereotype threat: Group- and self-threat. *Journal of Experimental Social Psychology, 44*(3), 792-799. doi: 10.1016/j.jesp.2007.07.005

- Wright, D. B. (2018). Estimating school effectiveness with student growth percentile and gain score models. *Journal of Applied Statistics*, 45(14), 2536-2547. doi: 10.1080/02664763.2018.1426742
- Yoder, J., & Lopez, A. (2013). Parent's perceptions of involvement in children's education: Findings from a qualitative study of public housing residents. *Child & Adolescent Social Work Journal*, 30(5), 415-433. doi: 10.1007/s10560-013-0298-0
- Youngs, P., & Bell, C. (2009). When policy instruments combine to promote coherence: an analysis of Connecticut's policies related to teacher quality. *Journal of Education Policy*, 24(4), 435-460. doi: 10.1080/02680930802669920