Teachers’ Motivation and Beliefs in a High-Stakes Testing Context

Dissertation

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By
Heather S. Dawson
Graduate Program in Education

The Ohio State University
2012

Dissertation Committee:
Eric M. Anderman, Advisor
Lynley H. Anderman
Anita Woolfolk Hoy
Ann Allen
Abstract

High-stakes testing has created challenges for teachers, administrators, parents, students, and other related education stakeholders in recent decades (Nichols & Berliner, 2007). While high-stakes tests have a long history (Ravitch, 2009) it was not until No Child Left Behind was signed into law in 2002 that the tests became law for most states (Hursh, 2007). Standardized tests with stakes attached to them (high-stakes tests) follow the prevalent business model for improvement and efficiency (Amrein & Berliner, 2002). There are many concerns associated with testing, including narrowing of the curriculum (Donnelly & Sadler, 2009; Ryan & Brown, 2005), teacher attrition (Darling-Hammond, 2007), invalid test scores (Arenson, 2003; Kleinfield, 2002) student dropout rates (Dobbs, 2003), and increased workload for teachers (Valli & Beuse, 2007) just to name a few.

While the literature on the problems with high-stakes testing is bountiful, few studies have examined teacher and student motivation utilizing a theoretically-driven, empirically-designed method. The effects of high-stakes tests on teachers, students, and classrooms is well-known, however the literature lacks empirical work examining motivation specifically within the context of high-stakes tests.

The purposes of this study are to examine (1) teachers perceptions of test-related disruptions in the classroom, (2) the amount of stress these disruptions cause them, and (3) how the stress teachers perceive as a result of tests is related to their motivation for teaching. For the purposes of this study, motivation is operationalized using two widely-accepted frameworks: Social Cognitive Theory (specifically Teacher Self-Efficacy) and
Self-Determination Theory (specifically teacher autonomy-supportiveness and controllingness). Specifically, the research questions that drive this study are: (1) What is the relation between teachers’ perceptions of stress because of high-stakes tests and their teacher self-efficacy? And (2) What is the relation between teachers’ perceptions of stress because of the test and their autonomy-supportiveness and controllingness?

Using a quantitative, online survey, 550 teachers from all 50 states participated in the study. Teachers in this sample represented all grade levels, experience levels, education levels, and all income levels. The survey that the teachers were asked to complete included 6 parts: a basic demographic portion, the Teacher Self-Efficacy Scale (Tschannen-Moran & Woolfolk Hoy, 2001), administered twice, the Perceived Stress due to High-Stakes Tests Scale (Dawson, 2011), the Teacher Autonomy-Control Scale (Dawson, 2011), and the Problems in Schools Questionnaire (Deci, Schwartz, Sheinman, & Ryan, 1981).

Results indicated that teachers perceive extreme levels of stress as a result of high-stakes tests. Teacher self-efficacy is predicted throughout the school year by their income level, perceptions of administrative support, and experience level. Test-related disruptions also predicted teacher self-efficacy throughout the school year. Teacher self-efficacy in the weeks leading up to a high-stakes testing event is predicted by perceived stress due to the test, the likelihood that the school is going to pass the test, perceptions of administrative support, and teacher income. Teacher autonomy-supportiveness and controllingness is predicted only by the extent to which a teachers’ evaluation depends upon the students’ test scores. Results, implications, and limitations are discussed.
To my grandmother, Barbara Ann Burnett, without whose grace this would not have been possible.
Acknowledgements

This dissertation represents the greatest accomplishment of my life. Given that I am a woman with many faults, it stands to reason that I could not have done this without the patience, grace, and selflessness of several hundred people—I cannot mention all of them here, but please be reassured I have not forgotten anyone. These last six years have been a period of consistent growth, albeit slow at times, in all areas of my life. I suffered, and I grew. I resisted, and then I grew. I fell, and then I grew. I was hurt, I hurt others, and I hurt myself. And then I grew. And the people I mention in these pages were there for me—in some cases, at every step of the way and in other cases, at just the right moment. Regardless of the relative size of the role each of these amazing people played in my life, know that I deeply appreciate and admire each and every one of them for sticking by my side—especially when I abandoned myself.

I cannot imagine the magnitude of the task that Dr. Eric Anderman faced when he agreed to take me on as his student. I was a fledgling second-year Master’s degree student and had very little knowledge about the field, myself, or my future. Yet Eric accepted me and mentored me with great patience and persistence for four years. He went to bat for me, he provided for me, and he guided me. At times my frustration with myself and my direction overwhelmed me, yet he remained ever calm. He never worried about
my future; I worried about my future years in advance. He calmly corrected my mistakes; I repeated them with vigor. He promised he would not make me cry; I cried anyway. We shared tragedies, laughter, and research—and by the end, I found myself molded into the young eager scholar (with a job!) that I had sought him out to help mold me into. I do not know how to show my gratitude for all of this, except to say “thank you”—and you will always be my academic father, no matter where I end up going, and no matter where you end up going.

Behind him, always strong and steadfast, was Dr. Lynley Anderman. Her professional and emotional support over the years has been invaluable. I have learned the importance of a truly encouraging mentor; without Lynley I could not have learned all of the critical implicit lessons that come along with becoming an academic and a woman in today’s age. Because of her I learned how to recognize the importance of being an independent woman and what that means in the academy. Over so many lunches on High Street, we discussed the best ways to navigate this world that at times seems so armed with landmines, especially for women. We’ve both said it before, but it bears repeating: Lynley and I are truly cut from the same cloth. I am a much better person as a result of my relationship with her.

Dr. Anita Hoy is undoubtedly one of the all-time greatest people alive today. She is genuinely the most generous person I’ve known; she gives endlessly of her knowledge, mentorship, and resources. She truly wants to see her students succeed and does whatever is necessary to ensure they do. I am immensely fortunate to have had the opportunity to know her, and even more so to have had the opportunity to work with her. Similarly, Dr. Ann Allen is talented and dedicated, and has given me opportunities at Ohio State that
she could have easily reserved for more deserving students. She patiently devoted herself
to my success, always listening to my ideas and engaging in my dreams, teaching me
about the policy world and helping me figure out how it applied to my work. To both of
these devoted committee members I am grateful.

Without my grandmother, Barbara Burnett, none of this would have been
possible. Her encouragement and constant support for three decades is the foundation for
my achievement. She never wavered in her generosity, grace, and love for me. I gave her
limitless cause for panic from the time I was a teenager, from moving out of my parents
house at 17, to failing out of college during my first, second, third, and fourth semesters
at the University of Toledo at 18, to waking up one morning and moving 650 miles away
at 20 years old, to announcing at age 22 that I would be going back to college and that
yes, I would get a Ph.D….but first, I had to get that elusive Associate’s Degree. I flirted
with danger, with absurdity, and with the impossible. There were bad boys and religious
cults, impulsive decisions and certain uncertainty. She handled what must have been
monumental internal turmoil and dismay with grace and poise, and continued to support
me. Without her, I would be a very different person today.

My grandfather, Samuel Burnett, laid the foundation for the core of my beliefs.
From the time I was born, he affirmed my uniqueness and reminded me that I was
special—a belief that I still hold, however inappropriate it may be. He proved this to me
as a child in copious ways. How many princesses had the power to call off school on a
snow day? Only one—me. How many princesses got to attend special meetings for
leaders in the city and school district? Only one—me. How many princesses were
fortunate enough to enter the zoo from the back instead of the front with the rest of the
visitors? Only one—me. How many princesses got to meet important people like Marcy Kaptur, go into the White House, and take a private tour of the FBI crime unit? Only one—me. My grandfather introduced me to the world, offering me experiences that I still recall in conversation and reflection to this day. With him, I flew on a plane for the first time, saw my first ocean, visited the Eiffel Tower and Buckingham Palace (although I was asleep), and slept in the top room of a four-story townhouse with views of a windmill in the Netherlands. With him, I took communion with the Royal Family. He gave me the hardest math problems in the world, and when I solved them correctly, he told me he had to take my word for it because he wasn’t as smart as I. He pulled money out of my ears, laughed at my first jokes, and took me to the best restaurants in Toledo, such as the Monroe Diner. He angered me on purpose, and then taught me how to control it—and as those closest to me laugh at the thought that I actually can control my anger, I will disclaim that it is a skill that I am yet developing. He gave me the confidence to be part of the world, instead of hiding from it. I am very much like him, and I am heartened by it.

Sometimes support and resources come in many different forms. To that end, my mother may be the single greatest supporter of this journey—and I am incredibly fortunate to have her. I have doubted myself, my abilities, my value, my purpose. My mom always believed in me. She believed in me when there was no reason to. There has not been a moment that I doubted myself that she did not reply, “I don’t know why you’re so down on yourself— you always succeed!” She supported me through painful decisions, and when I made the wrong decision, she supported me still. She let me make my own mistakes—something not all mothers are capable of doing. She watched me enter harmful situations and stood back, patiently waiting for me to emerge on the other
side, where she simply comforted me back to health. She let me get angry when I needed to be angry, and let me be frustrated when I was frustrated, and shared in my happiness when I was happy. She has never once said “I told you so.” I have learned what it means to be a supportive parent because of her.

My father has taught me that in life we can face great adversity and emerge stronger for it. Without that lesson, I could never have finished a doctoral degree. It will serve me well in the future, as I venture out into the world and face life’s challenges once again. The lessons my father has taught me in life reach to the core of my soul and form the basis of my identity to this day. Whatever the future holds, please don’t forget to make your funny face. Without this, none of this would be possible.

To my brothers in academia and in spirit, Michael F. Yough and Deleon F. Gray: Sometimes we fall down for no apparent reason. Sometimes we have nothing more to say, so we say “My monkey made me do it.” Sometimes we compete for the front row seat in R. Lo’s riveting stats class. Sometimes we fight for it. Sometimes the Dalai Lama interrupts our job talk, and sometimes we brave -10°F for a piece of flying pizza. Sometimes we give up and drop the log, and sometimes the dog slides across the floor and crashes into the door. Sometimes our shorts are too short. Sometimes we fall asleep at our desks, and sometimes we lose the world in the mess, or just our keys, whilst still some of us have two desks. Sometimes it takes years to get up the nerve to try the taco truck, and sometimes it’s just too great of a day to pass up Mexican on the patio. Sometimes we throw up in front of our class while we’re teaching. Sometimes our posters are too big. Sometimes we think we heard something else when the lady at Moy’s says, “Thank you, thank you very much.” Sometimes we never zip up our pants.

x
Actually, for some of us, we never, ever zip up our pants. Sometimes it’s a turkey vulture, and sometimes it’s a Great Black-Backed Gull—you can always know because the vulture’s wingtips point to the sky. Sometimes it’s “country thick.” Sometimes in Canada speed metrics change from miles to kilometers. Sometimes we call Phillip Zombardo just to see what’s up. Sometimes tornadoes strike in Toronto when we’re stuck in a traffic jam. Regardless, all roads lead to Deleon Gray. In spite of all of this, my brothers, you will forever be my academic siblings, and we will always have all of this.

It took a collaborative effort of mentoring to prepare me to be who I am today. I have been blessed with a strong skill set that I take with me into the academic world, given to me from my mentors at Ohio State and from gracious and talented individuals around the country and the world. These include Dr. Sharon Nichols from the University of Texas at San Antonio, who is my closest academic peer—that is, our interests are nearly identical. She has both mentored me and provided me with opportunities for collaboration resulting in a prestigious handbook chapter. She also served as part of my dissertation committee for a time, during the most difficult time of my graduate school career. Her guidance through this process was immeasurable. Drs. Helen Patrick (Purdue), Mike Martinez, and AnneMarie Conley (University of California, Irvine) formed a strong and balanced team of mentors during my time in San Diego in 2010. Their feedback on my work and professional guidance shaped my dissertation study in critical ways. Dr. Julie Turner (Notre Dame) was gracious and patient with me during our collaboration on conference presentations and manuscripts, and my experience with her and the other co-authors on that project taught me more than any course in graduate school. Dr. Paul Schutz (UT San Antonio) was a force from the moment I met him,
standing by and always calmly answering all of my frantic Facebook messages, treating me with kindness at every conference, and making me laugh during paper presentations when I should have been paying attention. His lap is a great place to sit when the room is twice as full as it should be.

My peers at Ohio State made my experience fun, memorable, and fruitful. Kate Kovach taught me how to start a fire. Ryan Poirier helped me realize I could study whatever I wanted. Stephanie Levitt showed me how talented I am at recruiting—although she was a bit of a captive audience. Sam Rocha taught me about philosophy, and Brad Rowe about the food and agricultural industries. Bree Frick taught me how to say “no.” Lorie Owens and Virginia Ressa made two years of state-level bureaucracy tolerable, while Deb Zabloudil, Helen Higgins, Chinyere Richardson, and Amber Mances made six years of graduate school tolerable. Kathy Kirkpatrick scraped me off the concrete in Denver and made sure I was returned safely to my bed. Meiling Chang always brightened my day, no matter the day. Nicole Leach and I can tear it up in D.C.—some of the best times of my life.

Brad Bethel helped me see that I could pursue this dream of getting a Ph.D., and allowed me to indulge in the fantasy during the summer of 2005, daring to spend a multi-day trip with me driving around the country visiting the campuses of various prospective graduate schools. Our time at the University of North Carolina-Chapel Hill revealed my passion for higher intellectual pursuits, and crashing the commencement speech was monumental for the both of us—not because of the speaker, but chilling with Mo Rocca. Wandering around the University of Virginia was meaningful to me in a way that I don’t
think Brad ever realized—because of those days, and that time with him, my perspectives on myself began to shift. For that I owe him a debt of gratitude.

Sometimes in life, timing befuddles the soul. Meeting Janelle Alloway a year before leaving Columbus was bittersweet. What more can best friends be than what Janelle and I have created with one another over the last year? Bonding over shared views on various touchy subjects, I know Janelle and she knows me, and I cannot thank her enough for her enthusiasm and support for me in the last year. She will be forever part of my life, regardless of our zip codes.

My oldest friendships, Amanda and Hilary Heisman—our journeys never part, even though time and distance find us frequently. Amanda, you are the person I know I can count on, no matter how varied our perspectives on the world—I know that our reunions will always be full of laughter and love. But mostly laughter. Hilary, we have shared some of the hardest times of our lives and you are always in my heart. For both of you, it does not matter how long it has been since we’ve spoken—we will automatically find that comfort with each other. Never forget that if you are on I-95 in south Florida and you see smoke spewing from under the hood of your car, it’s probably not a big deal. Flames are cause for concern.

To my favorite little bird: Thank you for being. You move me.
Vita

June 1998…………………………………Whitmer High School

December 2003……………………………A.A., Owens Community College

May 2006………………………………….B.S.Ed. Special Education, Bowling Green State University

August 2008………………………………M.A. Education, The Ohio State University

June 2012…………………………………Graduate Student, School of Education and Human Ecology, The Ohio State University

Publications


Fields of Study

Major Field: Education

Major Field: Psychology
# Table of Contents

Abstract............................................................................................................................p. iii

Recognition......................................................................................................................p. v

Acknowledgements........................................................................................................p. vi

Vita......................................................................................................................................p. xiv

Chapter 1: Introduction....................................................................................................p. 1

Chapter 2: Literature Review.......................................................................................p. 17

Chapter 3: Methods........................................................................................................p. 44

Chapter 4: Results..........................................................................................................p. 72

Chapter 5: Discussion....................................................................................................p. 112

References......................................................................................................................p. 142

Appendix A.....................................................................................................................p. 157

Appendix B.....................................................................................................................p. 159
List of Tables

Table 3.1 Demographic data, frequencies, and applicable means........................................p. 49

Table 3.2 The Perceived Stress due to High-Stakes Test Scale, Known Issues and Stress subscales.............................................................................................................p. 58

Table 3.3 The Teacher Self-Efficacy Scale, Generally, throughout the school year and In the weeks leading up to the test subscales and Cronbach’s reliability alphas for the pilot sample.........................................................p. 60

Table 3.4 The Teacher Autonomy-Support and Control Scale, Final Scale.........................p. 63

Table 3.5 The Problems in Schools Questionnaire, Vignettes and Items............................p. 65

Table 3.6 Dummy Codes for Categorical Variables..............................................................p. 68

Table 4.1 Means, Standard Deviations, and Rotated Component Loadings for the Perceived Stress due to High-Stakes Tests Scale, Known Issues Items.............................................................................................................p. 75

Table 4.2 Means, Standard Deviations, and Component Loadings for the Perceived Stress due to High-Stakes Tests, Stress Items..............................................................p. 76

Table 4.3 Means, Standard Deviations, and Loadings for the Teacher Self-Efficacy Scale, Generally, Throughout the School Year...............................p. 78

Table 4.4 Means, Standard Deviations, and Loadings for the Teacher Self-Efficacy Scale, In the weeks leading up to the test.................................p. 79

Table 4.5 Teacher Autonomy-support and Control Scale, Autonomy-Support Items.................................................................p. 80

Table 4.6 The Problems in Schools Questionnaire and subscales........................................p. 82

Table 4.7 Means, Standard Deviations, and Cronbach’s α for scales included in final analyses .................................................................................................p. 83
Table 4.8 Bivariate correlations between demographic variables.................................p. 86
Table 4.9 Means, Standard Deviations, and Correlations..............................................p. 90
Table 4.10 Means, standard deviations, and ANOVAs, Known Issues scale....................p. 92
Table 4.11 Means, standard deviations, and ANOVAs, Perceived Stress scale..............p. 98
Table 4.12 Means, Standard Deviations, and results on ANOVA and t-test for TSES..............................................................p. 102
Table 4.13 Means and Standard Deviations for the Problems in Schools Questionnaire, according to Teacher Evaluation and Grade Level .................................................p. 103
Table 4.14 Means, Standard Deviations, and Correlations for TSES, Generally throughout the year.............................................................p. 106
Table 4.15 Regression results, including change in R² for the TSESGen Model..............p. 106
Table 4.16 Regression results; coefficients at Step 2 of the TSESGen Model.............p. 107
Table 4.17 Means, Standard Deviations, and Correlations for TSES, in the weeks before the test..........................................................................................p. 109
Table 4.18 Regression results, including change in R² for the TSESTest......................p. 109
Table 4.19 Regression results; coefficients at Step 2 of the TSESTest.......................p. 109
Table 4.20 Regression results for the PIS.....................................................................p. 111
List of Figures

FIGURE 4.1 MEANS ON THE PS-HST, KNOWN ISSUES AND PERCEIVED STRESS SCALES ACCORDING TO SELF-REPORTED FREE/REDUCED LUNCH STATUS.......................................................... P. 93

FIGURE 4.2 MEANS ON THE PS-HST KNOWN ISSUES AND PERCEIVED STRESS SCALES ACCORDING TO GRADE LEVEL TAUGHT.................................................................................................................. P. 94

FIGURE 4.3 MEANS ON THE KNOWN ISSUES AND STRESS SUBSCALES FOR TEACHERS ACCORDING TO THE RELATIVE SUCCESS OF THEIR SCHOOL ON HST.................................................................................. P. 95

FIGURE 4.4 MEANS FOR KNOWN ISSUES AND STRESS SUBSCALES ACCORDING TO TEACHER EVALUATIONS.................................................................................. P. 97

FIGURE 4.5 MEANS FOR THE TSES GENERALLY THROUGHOUT THE SCHOOL YEAR AND IN THE WEEKS LEADING UP TO THE TEST ACCORDING TO HST PERFORMANCE.................................................................................. P. 99

FIGURE 4.6 MEANS ON THE TSES SCALES ACCORDING TO THE ITEM MY ADMINISTRATION IS SUPPORTIVE OF MY WORK AS A TEACHER........................ P. 100

FIGURE 4.7 MEANS FOR TEACHER SELF-EFFICACY ON BOTH SCALES ACCORDING TO TEACHER RACE.................................................................................. P. 101
Chapter One: Introduction

The purpose of the current study is to examine teacher motivation within the context of high-stakes testing policies. High-stakes testing policies have proliferated over the last two decades (Hursh, 2007), causing stress for teachers, administrators, students, and parents (Jones & Egley, 2004). Teacher motivation is operationalized, for the purposes of this study, within Social Cognitive theory and Self-Determination theory. Specifically, teacher motivation is defined as teacher self-efficacy beliefs and teachers’ tendency to adopt an autonomy-supportive or controlling orientation toward teaching will be examined.

Introduction

No Child Left Behind was signed into law on January 8, 2002, with a majority vote in both the House (381-41) and Senate (87-10) (Hursh, 2007). The educational reform was supported by both democrats and republicans whom often support neoliberal and neoconservative policies that aim to perpetuate market forces and competition, privatization, and school choice (Au, 2009; Hursh, 2007; Parkison, 2009). Scholars argue that one cannot understand NCLB without situating it within the larger socioeconomic and sociopolitical sphere (Hursh, 2007). Neoliberalism, dominant in the G.W. Bush administration, emphasizes

“The privatization of the public provision of goods and services—moving their provision from the public sector to the private—along with deregulating how private producers can behave, giving greater
scope to the single-minded pursuit of profit and showing significantly less regard for the need to limit social costs or for redistribution based on nonmarket criteria. The aim of neoliberalism is to put into question all collective structures capable of obstructing the logic of the pure market” (Tabb, 2002, p. 29).

Neoliberal ideology categorizes individuals as “autonomous entrepreneurs” who can always attend to their own needs, who are capable of rationally assessing costs and benefits of their decisions, and who perceive that those individuals who do not “make it” have made poor choices (Apple, 2006). The neoliberal, while arguing for individual rights, suggests that nothing is ever society’s fault—it is always the individual’s fault (Apple, 2006; Harvey, 2005; Hursh, 2007; Foucault, 1979). In the age of globalization, neoliberals argue, competition and free market ideology are a requirement, not an option, and educational reforms that increase efficiency, accountability, fairness, and equality are also therefore a requirement (Hursh, 2007).

To that end, standardized tests are a means to providing conscious consumers with a quality indicator, as well as providing an objective marker of student achievement (Hursh, 2007; U.S. Department of Education, Office of the Secretary, 2003). There is some debate about whether high-stakes testing policies were ever intended to do what they were purported to, or if they were the means to an end—privatization (e.g., Au, 2009; Darling-Hammond, 2007; Hursh, 2007; Ravitch, 2009). Other scholars have argued that these policies are designed to systematically take control of education out of the hands of teachers and teacher educators and place it in the hands of businesses and corporations (Tatto, 2007; Zeichner, 2010). Test scores have become a commodity, subject to market forces, ultimately a signifier of status and privilege that is linked to the
allocation of resources—a mechanism for differentiation (Parkison, 2009; Zeichner, 2010):

“Those performing well on the standardized tests then possess the currency of high test scores and gain power within the political dialogue that produces academic standards. Power comes from producing high test scores. The cultural capital that places individuals in a position to produce these high test scores gets reinforced as academic standards are produced by those who have high test scores” (Parkison, 2009, p. 52).

Several consequences have manifested since 2001 when No Child Left Behind was signed into law as an unfunded mandate. It is debatable whether these consequences were intended or unintended—according to Robertson (2008) high-stakes testing policies were written with very specific goals in mind:

“(1) The redistribution of wealth upward to the ruling elites through new structures of governance, (2), the transformation of educational systems so that the production of workers for the economy is the primary mandate, (3), the breaking down of education as a public sector monopoly opening it up to strategic investment by for-profit firms” (p. 12).

Regardless of intentionality, reforming the education system and requiring sweeping changes (e.g. accountability, adequate yearly progress, highly qualified teachers) without funding to support them inevitably will result in the “haves” continuing to have and the “have nots” continuing to have not.

**What is High-Stakes Testing?**

High-stakes testing is a model for educational improvement borrowed from the business world (Moses & Nanna, 2007), designed to provide a tool to states and school districts that measures the achievement of students, teachers, and school buildings. The purpose of the development of high-stakes tests was originally to reward high-performing
schools and penalize under-performing schools through consequences such as funding cuts, media scrutiny, and staffing reorganization (Amrein & Berliner, 2002). It was thought that schools that were not performing up to the standard would be motivated to improve in order to avoid further sanctions, and that students would be motivated to learn, and teachers forced to “do their job” (Amrein & Berliner, 2002, p. 4). Some have argued that high-stakes testing is ideologically right-of-center, stemming from neoconservative policies that promote the individual as the basis of society, and ignoring the social and economic conditions in which individual is situated (Moses & Nanna, 2007; Ravitch, 2009).

Moses and Nanna (2007) describe high-stakes tests as those “whose scores have a direct impact on a person’s life options and opportunities” (p. 56). This includes using high-stakes test scores to make decisions about academic tracking, grade retention, or graduation eligibility for students. For teachers, this may include using scores to make decisions about salary, promotion, or placement, or to evaluate teacher performance. These authors also identify three primary factors that have contributed to the proliferation of tests: administrative utility, profit motives, and political ideology (Moses & Nanna, 2007). From the perspective of administering education, tests are appealing because they provide a quick, cost-effective way of handling large amounts of data (Moses & Nanna, 2007; Supovitz, 2009). From the perspective of the testing industry, there is much to be gained (both financially and politically) by promoting the use of tests, regardless of the concerns that are had over their benefits and harms. Finally, many politicians, both democratic and republican, have relied on education as a way to influence constituents in elections, and to drive politically-motivated policy decisions (Moses & Nanna, 2007).
The History of High-Stakes Testing

Some argue the impetus for high-stakes testing in schools originated with the Sputnik crisis and the space race of the 1950s, which caused a reactionary American media and political elite to question the effectiveness of American education (Amrein & Berliner, 2002; Nichols & Berliner, 2007). Others argue that standardized tests have been widely used since the beginning of the 20th century, originating with the development of the standard IQ test by Alfred Binet in the years before World War I (Clarke, Madaus, Horn, & Ramos, 2000). Still some claim that testing has held roots in the United States for more than 150 years (Moses & Nanna, 2007).

Radical education reforms of the 1960s and 1970s are often cited as a precursor to the testing movement of recent decades (Ravitch, 2009). From busing to total freedom from authority in the classroom to removing the canon and allowing students to design the curriculum, this period was a time of educational experimentation that seemed to see no limits to the reforms (Ravitch, 2009). Clarke, Madaus, Horn, and Ramos (2000) identify four social “forces” during this era that contributed to the proliferation of testing in the U.S. They are (1) the public dissatisfaction with the quality of education, (2) a wave of federal and state legislation that explicitly mandated or strongly suggested standardized testing programs, (3) a shift from measuring inputs to measuring outputs in education, and (4) bureaucracy in general and in education that made testing technologies more widely available.
A Nation At Risk. Fears of economic failure culminated in a 1983 report released by the Reagan Administration entitled “A Nation At Risk” (U.S. Department of Education, 1983). This report was by all accounts landmark (Darling-Hammond, 2007; Nichols & Berliner, 2007; Ravitch, 2009) and one account went so far as to call the report “sensational” (Ravitch, 2009, p. 24). As a result of the report, policymakers in every state except Iowa developed educational standards, and every state except Nebraska also developed testing policies to be used to assess the policies (Amrein & Berliner, 2002). The necessity of the report to instigate fears among the public is still debated; according to some authors, the report was nothing more than “alarmist rhetoric by the Reagan Administration intended to undermine public education” (Ravitch, 2009, p. 24), a way to perpetuate neoconservative ideology (Moses & Nanna, 2007). This report served as an impetus toward high-stakes testing, privatization, and further neoliberal and neoconservative agendas (Hursh, 2007).

As mentioned above, technology has played an important role in the development and proliferation of testing since the early 1900s (Clarke et al., 2000). The development of various technologies including multiple choice items in the early 1900s, the “scantron” in the 1950s, and various forms of computer-adaptive testing in the 1990s have all had strong influences on testing (Clarke et al., 2000), creating what is now known as the testing “industry.” Additionally, technologies such as computers and computer software programs have proliferated, resulting in need for hardware, software, CD-ROMS that supplement textbook content, programmers, IT specialists, and the like (Burch, 2006). Testing has risen up alongside these technologies, contributing to the newfound industry.
“Industry” is a critical component of testing that cannot be ignored. The testing movement has become a bastion for the business world, with the most recent estimates hovering around $1 billion in testing alone\(^1\) (Check Toch 2006). As noted by Clark et al. (2000), it is difficult to obtain accurate estimates of the size of the industry due to the prolific and largely unregulated nature of the industry, but also due to the secretive nature of the testing industry. Most of this money is directed to relatively few testing corporations such as CBT McGraw-Hill, ETS, NCT Pearson, etc. (Clark et al., 2000; McMillan, 2000; PBS, 2002). In recent years, private vendors have been contracted to develop tests, to align tests, to develop standards, all the way to full-service test development solutions that complete all aspects of curricular design and assessment for school districts (Burch, 2006). NCLB allowed fast and easy access for private sector interests into schools by attaching stakes to test results, the consequences of which led directly to a massive increase in test development revenues (Burch, 2006).

**Problems associated with high-stakes testing**

**Curriculum narrowing.** Curriculum narrowing emerges as a problem for many schools as a result of high-stakes tests (HST) (Donnelly & Sadler, 2009). Teachers report that they feel it is necessary to reduce the material, often eliminating content that will not be tested, as a result of the increased pressure to produce desired outcomes on the tests (Clarke et al., 2003; Donnelly & Sadler, 2009; Ross, 1998; Ryan & Brown, 2005; Urdan & Paris, 1994). Teachers report the greatest levels of curriculum narrowing in low

---

\(^1\) The most recent cited estimate I could find states the industry could be worth $700 million (Frontline, 2008), however while employed for McGraw-Hill I was told numerous times that the testing industry was a $1 billion industry. This figure could reflect global sales for tests.
socioeconomic regions; that is, the areas that are the poorest are suffering the greatest
from these policies (Lomax, West, Harmon, Viator, & Madaus, 2003). Test preparation,
or “teaching to the test,” is becoming a salient practice, resulting in some cases in all non-
tested material being eliminated from the classroom (Donnelly & Sadler, 2009; Ryan &
Brown, 2005; Shaver et al., 2007). Teachers are being “encouraged or required to
reallocate time away from core subjects not tested on state examinations, and to eliminate
or curtail special projects, experiments, library research, extensive writing, or oral
assignments” (Ryan & Brown, 2005, p. 365).

The issue of curricular narrowing has had an effect on teachers in the classroom. For
instance, curricular narrowing as a result of No Child Left Behind had a negative
effect on some beginning teachers’ perceptions about their ability to develop a satisfying
teaching practice (Crocco & Costigan, 2007). In New York City, teachers described the
curriculum as “oppressive and insulting to their developing sense of professionalism”
(Crocco & Costigan, p. 526). Teachers in California were subjected to more narrowing
and constraint in their autonomy in schools that were low-income or minority than their
higher-affluent counterparts (Crocco & Costigan, 2007; Darling-Hammond, 2007). Some
teachers reported leaving the field because they felt their students were capable of much
more than the standards would allow them to teach, and some research suggests that
formal mentoring can improve teachers’ attitudes, efficacy, and autonomy; however,
perpetual funding limitations make implementing mentoring programs like these
challenging (Crocco & Costigan, 2007).

Curriculum narrowing also has a negative effect on students. Teachers teach the
material on which they know their students will be assessed. This is increasingly true
with increasingly high-stakes (Berliner, 2009). As the stakes of tests increase, teachers in many cases will focus more of their instructional time on preparing for the test as a way of bolstering their defenses against negative consequences in which high-stakes testing policies often result, such as threats of funding cuts, public scrutiny in the media, and employment cuts. When these consequences are carried out, schools that have more resources at their disposal can afford to make better preparation than schools that are less affluent. Following this phenomenon to its logical conclusion results in a student body that is further differentiated on socioeconomic lines, as the students from wealthier families will become increasingly more likely to receive the intellectual and cultural capital that non-tested content areas offer, while those from lower socioeconomic families go without (Berliner, 2009; Kozol, 2005).

Through the process of curriculum narrowing, testing has effectively defined what is counted as legitimate knowledge and what is not. In a qualitative study examining high-stakes testing in the classroom, researchers found that high-stakes tests exert control over three specific areas of pedagogic discourse (Au, 2007). They are classroom content, forms of content knowledge, and teacher pedagogies. Some states (e.g., Arizona, Virginia) have moved to a model in which all content areas are tested, thereby minimizing concerns about tested versus non-tested content areas; however, curriculum narrowing occurs in more ways than just subjecting tested content areas against non-tested content areas. For example, noted by Au as the second form of pedagogic control, teachers often react to testing pressure by offering students isolated pieces of knowledge in order to prepare students for the often shallow forms of multiple choice items that will
be presented on the test. Relying on teacher-centered pedagogies such as lecture is the third form of control (Au, 2007).

**Teacher attrition.** Teacher attrition has been linked to testing policies in many studies, and is an especially poignant problem in areas of high poverty, where salaries are often lower and working conditions less desirable (Darling-Hammond, 2007), or among certain types of teachers, particularly those holding negative views of HST (Donnelly & Sadler, 2009). Many teachers attribute the attrition to their dissatisfaction with testing policies (Nichols & Berliner, 2007; Tye & O’Brien, 2002). Stress as a result of many of these issues is likely to contribute to teachers’ decision to leave the field.

There are many variables that teachers identify as stressful; in one study teachers noted that their ability to influence student academic performance was being limited by testing policies (Berryhill et al., 2009). These teachers cited limited time, too many tasks, and the necessity of changing their teaching practices as a result of the tests as leading causes of stress in their work (Berryhill, 2009). Teachers also expressed concerns about the fairness of test scores as stressors in their jobs (Berryhill et al., 2009). Teachers have identified accountability policies as harmful to their well-being (Abrams, Pedulla, Madaus, 2003; Berryhill, Linney, Fromewick, 2009; Jones & Egley, 2004; 2007). The harm teachers identified is a result of perceived stress that is emerging as a result of tests.

**Arbitrary test scores.** Standardized test scores in many states are meaningless, manipulated by politicians with an agenda to fulfill. For example, in New York cutoff scores are raised or lowered depending on whether the state wants to increase the graduation rate or appear more rigorous (Hursh, 2007). In other cases tests were so poorly constructed that entire cohorts of scores were discarded (Arenson, 2003). In yet
another instance, one English exam was revised post-construction to remove any reference to race, religion, ethnicity, sex, nudity, alcohol, profanity, and all references to Judaism, yet students were still questioned on deleted portions of the test (Kleinfield, 2002). Teachers are expected to prepare students to pass these tests on a regular basis.

Other issues related to test validity have been researched as well. As Darling-Hammond notes, it is impossible to reach 100% proficiency on a norm-referenced test—the kind of test that many states adopted to meet the requirements of NCLB (2007). The implication here is that the policy was written with impossible outcomes in mind. Adequate yearly progress dictates that a school will be declared “failing” if it does not reach the minimum standards. High-stakes are attached to this, based on the notion that low-performing schools, teachers, or students will be motivated to improve by shame; indeed, there often is an implicit assumption that being shamed publicly, finding your school de-funded, or having your students drop out will be the impetus for change.

Students are disappearing (dropping out, being held back, or asked to leave) in many states, resulting in inflated test scores (Darling-Hammond, 2007; Dobbs, 2003). Many of those are students of color, second-language learners, and students with disabilities. The state of Massachusetts experienced a 300% increase in dropout between the years 1997-2000, just a few years after the state implemented a new high-stakes accountability system (Darling-Hammond, 2007). In New York City between 1999 and 2001 over 55,000 students were “discharged” from school without graduating (New York Times, 2001). In Texas, a study of one 9th-grade cohort revealed that test scores reported to the media and the state reflected that the rate of dropout was kept below the mandated 3.5%, and that most of the students had passed the Texas high-stakes exam; real figures
examined by the researchers of the study found that 30% of the students had been retained in 9\textsuperscript{th} grade and 49% had dropped out, withdrew, or disappeared (Heilig & Darling-Hammond, 2008).

Rhoades and Madaus identified two types of human error in high-stakes testing: active errors and latent errors (2003). Their discussions of these two types of human errors did not include errors of measurement; random measurement error has been well documented elsewhere (Rhoades & Madaus, 2003). Active human error refers to the kinds of errors that individuals commit and are easy to identify. Latent errors refer to those errors committed by humans that are rooted in managerial decision-making and offer the greatest threat to complex systems (Rhoades & Madaus, 2003). These kinds of errors “stem from poorly conceived legislation or policy mandates” (p. 6) and include the use of a single test score to determine critical life decisions, requiring test scores be reported faster than can be dependably accomplished by the vendor, or mandates that school test scores increase by a certain figure per year (Rhoades & Madaus, 2003). When active errors are discovered, it is logical to question if latent errors are at play.

These errors are costly. For example, in 1997 the Educational Testing Service (ETS) discovered an error on the SAT that lowered the test score of 15,000 students by up to 100 points (Rhoades & Madaus, 2003; Sandham, 1998). In 1999, dropped scores were reported on the TerraNova (McGraw-Hill) in Indiana, Tennessee, and New York City (Sandham, 1998). Officials in New York City, specifically New York State Education Commissioner Richard Mills and New York City Mayor Rudy Giuliani, used the scores to order students scoring below the 15\textsuperscript{th} percentile into summer school (Rhoades & Madaus, 2003). Thirty-five thousand NYC students attended summer school
that year. Afterward, McGraw-Hill admitted the drop in test results had been the due to an error in scoring, and determined that 8,668 students had been mistakenly ordered into summer school. The TerraNova scores for this group of students had in fact risen during the time in question (Hartocollis, 1999; Rhoades & Madaus, 2003). Many more of these kinds of errors are being reported around the country as well.²

**Increased workload.** Teachers’ workload has become more stressful as a result of testing policies. In one study, researchers found that teachers were working more and with more intensity: “the number of tasks [teachers] were asked to assume increased in number, expanded in scope, and intensified as the policy climate in their school district became more high-stakes. Both the magnitude and the rate of change were greater than we anticipated” (Valli & Buese, 2007, p. 550). Teachers’ roles also shifted dramatically, becoming more hierarchically controlled, increasingly regulated through assessment and data analysis expectations, losing most of the flexibility in their teaching practice (Valli & Buese, 2007). In lieu of the de-skilling of teaching, the act of teaching became increasingly demanding as well as difficult (2007). The problem of stress due to high-stakes testing is not unfamiliar to researchers of education, and stands at the center of this dissertation study.

**Student achievement.** In many studies, high-stakes tests are not found to have any relation to student achievement. In a study conducted by Marchant, Paulson, and Shunk (2006), high-stakes testing policies were related to higher scores on NAEP tests,

---

² I saw this firsthand a handful of times during my time of employment at the Ohio Department of Education.
but these differences disappeared once demographic data was controlled. In other studies, no connection between testing and student achievement was noted (Nichols, 2007). In most studies, caution is warranted while interpreting results due to the large probability of teaching to the test (Nichols, 2007).

Other studies have mixed results. In one study by Carnoy and Loeb, student achievement was examined in relation to the “strength” of the accountability in the state (2002). Achievement was determined by the National Assessment of Educational Progress (NAEP) math test. For this study, achievement was positively related to accountability, suggesting that states with stricter high-stakes testing policies also saw increases in student achievement (Carnoy & Loeb, 2002). However, in one other study, no achievement gains in lieu of high-stakes testing policies were evident (Lee & Wong, 2004). The evidence is inconclusive for the relations between accountability and achievement, and is most likely based greatly upon methods used to examine the question.

**Other test-related problems.** Other issues have arisen that appear to be related in some aspect to testing. In one study, bullying was shown to have an interaction with testing. As teachers focused intensely on preparation for tests, their attention was distracted from student behavior in the classroom, allowing bullying to occur more frequently (Hazel, 2010). Additionally, school-wide attention on testing and test preparation has led teachers to “compartmentalize” student academic needs (e.g., from student emotional needs to exclusively student academic [test-related] needs) (Hazel, 2010). The consequence of this, Hazel argues, was an increase in bullying activity. This article presents evidence that testing is harmful indeed; however, testing has become so
prolific that it is increasingly difficult to determine what problems in education are test-related, as it seems increasingly so that all problems in education are test-related.

Cheating has been identified as one problem that has increased since testing policies have proliferated. Fifty percent of teachers reported cheating at some point in order to help their students pass the test, while fifty percent of teachers reported knowing at least one teacher who had done the same (Amrein-Beardsley, Berliner, & Rideau, 2010). For the teachers in this study, the cheating behaviors were conducted in response to increased pressure to ensure passing scores on the test, and were often a protective measure against the related threats and consequences (Amrein-Beardsley et al., 2010).

Nichols and Berliner (2007) suggest this form of cheating actually acts as a form of resistance to the test, a consequence of “being placed by our politicians in untenable positions, environments where pressures encourage questionable behavior” (p. 34). Teachers may resent their situation and feel they have little control over it. As a result, teachers may feel that cheating is justified (Nichols & Berliner, 2007).

Cheating occurs on a larger scale as well. In 2003, it was reported that the state of Massachusetts cheated on the state test results, reporting to the public and to the media that 90% of the students whom had taken the Massachusetts Comprehensive Assessment System (MCAS; National Board on Educational Testing and Public Policy, 2003). When appropriate enrollment calculations were taken into account, the percentage of students who had actually passed the test was closer to 70% (NBETPP, 2003). The state had failed to account for attrition of students, calculating the passing rate using only then currently enrolled students in the 12th grade (NBETPP, 2003).
High-stakes testing and teacher motivation

While many authors, scholars, critics, intellectuals, researchers, teachers, administrators, and parents have criticized high-stakes testing policies they still draw broad support politically (Supovitz, 2009), and from various other scholars in related fields (e.g., Supovitz, 2009). Many of the criticisms, while valid, fail to frame empirical work in theories of motivation that can help to provide evidence that high-stakes testing is harming teacher motivation for teaching. For example, in one article by Richard Ryan and Netta Weinstein (2009), these authors, lay out a concise, detailed, and very accurate argument against high-stakes testing framed in self-determination theory. However, the authors did not collect any data to validate their arguments—it was a purely theoretical piece. This article will be described in the literature review below. My hope is to begin a career examining teacher and student motivation that demonstrates empirically what all of the above stakeholders have been arguing intellectually for years.

Summary

High-stakes testing has elicited many unintended consequences for teachers, students, and administrators. These include curricular narrowing, teacher attrition, arbitrary or useless test scores, and increased workload for teachers, among others. Additionally, students are experiencing pressure to produce passing test scores and those who are (for a variety of reasons) not able to are being pushed out of schools, reinforcing dropout behavior, and resulting in falsely inflated test scores. These issues are responsible for increasing stress among teachers (e.g., Reich & Bally, 2010), the effects of which will be described next.
Chapter Two: Literature Review
In pastures across the heartland  
Dairy farmers could tell with a glance  
If their cows were unusually skinny  
No prognosis was left up to chance

In time there were remedies  
Recipes tested  
Farmers were trained  
New diets suggested

Despite some progress  
The farmers were concerned  
And did their best to uncover  
What the cows had learned

About eating proper doses  
And grazing in the field  
So that they’d be healthy  
Their stoutness revealed

Then the government issued a warning  
“Our cows are underweight!  
The ones in Europe and Asia  
Get fat at a faster rate!”

Experts jumped into action  
Dairy farmers were urged to comply  
With new strict diet guidelines  
And discouraged from questioning why

A law was passed to move things along  
“No Cow Left Behind!” was its name  
Norms were issue by age and weight  
To ensure all cows’ growth was the same

Food was measured in exact proportions  
Cows must be fed in troughs and rows  
Thus ignoring the farmers’ wisdom  
For how a scrawny cow grows

The law had another component  
To measure the cows with precision  
The scales must be digitally calibrated  
To accomplish our fattened up vision

The farmers were told to get cows on those scales  
Whether they liked it or not  
And to measure the girth of their bellies  
To gather the data they sought

The cows, reluctant and unhappy,  
Attempted to stand on those scales
They ate the bland diet we fed them
But hated that stale hay in bales
They remembered the fresh grassy pastures
Where they used to munch and play
And longed for the freedom of choosing
How to spend a warm summer’s day

Then one night at the dinner table
Two farm children wondered aloud,
“Why don’t we feed our cows better food?
And not pack them into a crowd?

We should let them run around
Enjoy the fresh air outside
They we’ll see how those bellies
Begin to get really wide!”

So they tested their theory
Prepared tasty concoctions
With inventive ingredients
And culinary options

One day it was oatmeal
With wheatgrass and soy
Dinner was crunchier—
Roasted peanuts and bok choy

The cows on that farm
Just feasted for weeks
On exotic combinations
Like chocolate covered leeks

The children discovered
A foregone conclusion
The cows’ weight loss crisis
Was just an illusion

The cows were depressed
Not too skinny or thin
A variety diet
Was the way to begin

To lift up their spirits
Give them lofty ambitions
The children painted cow portraits
In full color renditions

They sent them to Washington
With detailed description
But the question remains
Will the politicians listen?

~Miletta & Morris, 2006
Introduction

The purpose of this review is to illustrate the relations of perceived stress because of high-stakes tests and teacher as defined by Social Cognitive Theory (specifically teacher self-efficacy). Self-Determination Theory (autonomy-supportiveness and controllingness) will also be examined. I will begin with a detailed review of stress from the psychological literature, emphasizing the relevant the work that has been done that specifically examines teacher stress. Then I will synthesize each of the theoretical frameworks that this study utilizes, beginning with Social Cognitive Theory (teacher self-efficacy), and finishing with Self-Determination Theory (autonomy-supportiveness and controllingness).

Stress

Stress is defined as a particular kind of relationship between person and environment where the demands of the relationship exceed the resources available to the person (Lazarus, 1990). Stress does not exist exclusively in the person or exclusively as an environmental input, but it is a dynamic, multivariate conjunction between the two and cannot be reduced to a single, linear event:

“Stress is a multivariate process involving inputs, outputs, and the mediating activities of appraisal and coping; there is constant feedback from ongoing events, based on changes in the person-environment relationship, how it is coped with and, therefore, appraised” (Lazarus, 1990, p. 4).

Appraisals and coping are the mediators between the event and emotional outcomes (Folkman & Lazarus, 1986). Appraisal is the process through which the significance of the event and the options for coping are evaluated (Folkman & Lazarus, 1986); or as Lazarus pointed out in his 1993 piece, appraisal is a
negotiation between the demands, constraints, and resources of the environment and the goal hierarchy and beliefs of the individual (1993). Coping refers to the thoughts and actions that individuals engage with to manage the internal or external demands of a stressful event (Folkman & Lazarus, 1986). According to the cognitive-phenomenological theory of stress and coping, there are two commonly accepted methods of coping in which an individual can use to deal with stress: emotion-focused coping and problem-focused coping (Folkman & Lazarus, 1986; Lazarus, 1990). The two will be described in brief below.

**Two Types of Coping**

Problem-focused coping refers to any strategy used to deal with psychological stress that changes the person’s stress-related relationship with the environment, resulting in a change in conditions for the better (Lazarus, 1993). This coping strategy reduces psychological stress by eliminating or changing the stressor (e.g., psychological stress due to procrastination is eliminated by tackling the task) (Folkman & Lazarus, 1988). The terms of the relationship are changed with the application of problem-focused coping, ultimately affecting the individual’s appraisal of the situation (Lazarus, 1990).

Emotion-focused coping efforts are directed at regulating emotional distress that is associated with a stressful event (Folkman & Lazarus, 1986; Lazarus, 1990). This coping strategy changes the way a person interprets the threat, and does not change the threat itself (Lazarus, 1993). This strategy involves reappraisal of the threat into non-threatening terms (e.g., he said what he said because he’s having a bad day). This form of coping is a powerful method to control psychological stress even if it does not change the external circumstances for the individual, because it changes their construal of the

These two conceptually distinct forms of coping have been established empirically and validated in many studies (Folkman & Lazarus, 1986; 1988; Lazarus 1990; 1993; 2000; Lazarus & Folkman, 1984), although Lazarus notes that it is a flaw to keep them separated in analyses, as they are both equally part of the overall dynamic stress system, informing each other as well as the emotional outcomes of the stress response (2000). Support for this exists in results showing that emotion-focused and problem-focused coping co-occur in a study that examined daily coping with rheumatoid arthritis (Tennen et al., 2000). Emotion-focused coping was predicted by the previous day’s problem-focused coping (Tennen et al. 2000), and the interaction between emotion-focused and problem-focused coping predicted the emotional outcome to the stressor. Individuals that were not able to control their pain with problem-focused coping on a given day then adjusted their emotion-focused coping the next day (Tennen et al., 2000).

Measurement of Stress

Operationalizing stress as a dynamic, fluid system and not a single variable presents several measurement challenges. Lazarus notes the challenges throughout his writings (1990; 1993; 2000), boldly stating that accepting that stress is a complex system of factors, and not a single or linear construct, dooms measurement to failure (1990). He suggests instead the development of a series of measurements designed to capture various facets of the stress process: key inputs, antecedents including person-centered motivational patterns and belief systems, mediating appraisals and coping, and emotional outputs (Lazarus, 1990). Engaging in process research is difficult, costly, and time
consuming, as Lazarus notes, but is superior to most other measurement options. This manner of measuring all of the variables affiliated with the stress relationship is difficult, Lazarus notes, and he acknowledges that methodology is not yet sophisticated enough to completely capture the stress relationship in a given moment; however, current methodology is still useful for informing the scientific community, even in spite of limitations (1990).

A decade later, Lazarus returned to offer commentary on studies that attempt to measure the stress process longitudinally, in a day-to-day fashion, with an interpersonal as well as intrapersonal approach (Lazarus, 2000). Lazarus argues that longitudinal research on the stress process offers the possibility of examining both state and trait characteristics of stress (2000). However, measuring both interpersonal and intrapersonal variables every day with a sample size large enough to warrant adequate power is an unrealistic expectation for a doctoral dissertation, a notion suggested by Lazarus. As I learned through this process, the issue of measurement of stress in studies without funding is a limitation that we as researchers must accept.

Lazarus notes elsewhere that the positivist approach to examining psychological phenomenon has posited itself toward developing generalizable laws and developing invariance, and with ambivalence toward the more nuanced individual differences which must be understood if true prediction is to be possible (1993). This perspective is likely to attribute individual differences in stress processing, or variance, as errors of measurement (1993). SAY SOMETHING ABOUT THE ETHICS OF TRYING TO PREDICT HUMAN BEHAVIOR.

**Stress and Outcomes**
Individual differences in beliefs, traits, and environment will all contribute to an individual’s likelihood of processing stress in an adaptive manner or maladaptive manner. Examples of the individual differences that are known to empirically mediate the relation between stress and emotional outcomes include constructive thinking (Epstein & Meier 1989), hope (Snyder et al., 1991), and self-efficacy (Bandura 1982), among others. Constructive thinking, as operationalized by Epstein & Meier in the 1980s, was related to adaptive coping processes, resulting in improved stress management. Snyder et al. (1991) examined hope in relation to stressors and the interactions between the two and found that individuals scoring higher on the hope measures were more likely to set more challenging goals for themselves and persist in the face of difficulty. Bandura noted that individuals with low self-efficacy beliefs likely perceive events as more stressful than they are, and also perceive their capacity for handing the event as less than it really is (1982).

Stress and coping, and stress and motivation, are synergistic in their relations to one another (Folkman & Lazarus, 1986). The relation between an individual’s motivational beliefs and his or her perceived stress and consequential emotions is very complex and nuanced. Research has shed some light on these issues, however. In work on stress and self-determination in the workplace, Weinstein found that individuals given greater choice at work, or higher autonomy, experienced less job stress (Weinstein & Ryan, 2010). Similarly, individuals under controlling conditions perceived constant pressure at work (Weinstein & Ryan, 2010). These findings echo other studies on motivation in the workplace indicating that when individuals are autonomous at work, enjoying and valuing the work rather than focusing on external rewards, they experience
a greater sense of challenge with respect to work-related stressors and less pressure; as a result, they are more productive (Amabile, Hill, Hennessey, & Tighe, 1994). This translates to teaching as well. When teachers feel autonomous in the classroom, they are likely to enjoy the work and likely to value teaching for the sake of teaching rather than for the sake of test scores. Studies comparing autonomous and controlled motivational styles showed that controlled individuals are less persistent and perform less satisfactorily when receiving negative feedback (Weinstein & Ryan 2010). Additionally, Weinstein and Hodgins (2009) demonstrated that autonomously functioning individuals, either as an individual difference or as a result of motivational priming that elicits autonomy versus control orientations in the moment, are more capable of effectively processing stressful events.

**Teacher Stress**

Testing has been shown to be stressful for teachers in several studies (e.g. Berryhill, Linney, Fromewick, 2009; Darling-Hammond, 2007), providing evidence to suggest that stress is going to become a more salient issue for teachers as accountability policies are bolstered in the United states. For example, teacher stress was inversely related to teachers’ job satisfaction for American teachers (Klassen, Usher, & Bong, 2010), contributing to a lower perceived quality of life for some teachers.

Teacher stress is defined as “the experience by a teacher of unpleasant negative emotions, such as anger, anxiety, tension, frustration, or depression resulting from some aspect of their work as a teacher” (Kyriacou, 2001, p. 28). Teacher stress more generally is considered a negative emotional experience that is triggered by teachers’ perception that their work situation has constituted a threat to their self-esteem or well-being
(Kyriacou, 2001). Additional definitions exist that describe teacher stress in terms of the level of pressure placed on teachers, the dissonance between the demands placed on an individual and the individual’s ability to cope, and in terms of teacher burnout (Kyriacou, 2001). Teacher stress can contribute to teacher burnout, depression, poor performance, absenteeism, low levels of job satisfaction, and teacher attrition (Klassen et al., 2010).

Job satisfaction was inversely related to teacher stress for some American teachers (Kyriacou, 2001) whereas teacher-efficacy for student engagement was positively related to teacher stress and social support (Sass, Seal, & Martin, 2010). McCormick and Barnett found that teachers who attributed greater stress to their personal inadequacies as teachers experienced a lower sense of personal accomplishment (2010). Stress was significantly correlated to student behaviors (Chaplain, 2008) including indolence, disrespect for the teacher, and noisiness (Geving, 2007) and mediated by perceived social support. That is, social support or perceived social support acts as a buffer to stress for many people; having a supportive social network reduces the effects of stress (Bakker, Hakanen, Demerouti, & Xanthopoulou, 2007; Berryhill, Linney, & Fromewick, 2009; Margolis & Nagel, 2006; Sass, Seal, & Martin, 2010).

As previously mentioned, stress is known to affect individuals’ perceptions of the situation as well as beliefs and motivation (e.g., Amabile, Hill, Hennessey, & Tighe, 1994; Bandura 1982; Kyriacou, 2001, Weinstein, 2010). Teacher beliefs are known to affect their behavior (Tschannen-Moran & Woolfolk Hoy, 2001), so there is reason to hypothesize that perceived stress because of high-stakes tests will affect teachers’ self-efficacy, their goal orientations, and their autonomy-supportiveness and controllingness toward teaching, among many other possible factors.
Teacher Self-Efficacy

In the mid-1900s, a theory of human behavior began to surface that would challenge behaviorism as the dominant intellectual paradigm of the time. Social Cognitive Theory triggered an evolution in the field of psychology by incorporating the roles of self-beliefs, self-regulation, and cognitive processes into the framework of human functioning (Bandura, 1997). Reciprocal determinism, or the notion that there is a dynamic system of influences that are engaged with one another perpetuating human functioning, was the first attempt at explaining the relations between individuals’ behaviors, their personal characteristics such as cognitive and affective processes, and the environment around them (Bandura, 1977; 1986; 1997). At the center of this idea was the construct of self-efficacy.

Self-efficacy is defined as a person’s beliefs about her or his ability to succeed at a given task (Bandura, 1986; 1997). According to Bandura, the definition of self-efficacy is “the belief in one’s capabilities to organize and execute the courses of action required to manage prospective situations” (1997, p. 2). These beliefs are important because they have been found to have a profound effect on outcomes; that is, one’s self-efficacy for a task can, essentially, change the result of that task—it is one of the greatest predictors of human behavior to be identified by psychologists (Bandura, 1997; Zimmerman, 2000). Bandura noted the importance of this in the theory, stating, “unless people believe that their actions can produce the outcomes they desire, they have little incentive to act or persevere in the face of difficulties” (1986, p. 391). Of particular importance is the notion that self-efficacy beliefs will affect the choices people make and the courses of action they pursue, how much effort they will spend on a task, the amount of time they will
persevere on the task in the face of obstacles, and how much resilience they will possess (Bandura, 1986). Additionally, self-efficacy can shape the perceptions individuals will hold about difficult tasks; that is, individuals with high self-efficacy will view a challenging task as something to be mastered whereas the individual with low self-efficacy will view the task as a threat to be avoided (Bandura, 1986).

Additionally, self-efficacy beliefs are multidimensional, are sensitive to variations in context, and mastery criterion-referenced. That is, these beliefs vary from task to task, may shift according to the environmental conditions in which the task is conducted, and are not affected by comparative judgments (Zimmerman, 2000). Self-efficacy beliefs are always oriented toward future functioning, and do not reflect past performance (Bandura, 1986; 1997; Zimmerman, 2000). To be clear, self-efficacy beliefs derive their sources from previous experience, such as mastery experiences; however, the actual self-efficacy belief is always oriented toward a future task.

There are four sources of self-efficacy: mastery experiences, vicarious experiences, verbal persuasion, and physiological/affective states (Bandura, 1997). Mastery experiences, the most influential of the four sources, predicted 24% of the variance in self-efficacy in science among middle school students in one recent study (Britner & Pajares, 2006). Similarly, Tschannen-Moran and her colleagues found mastery experiences to be significant predictors of teacher-efficacy in professional development settings, but the other three sources were not (Tschannen-Moran & McCaster, 2009). Mastery experiences, in combination with other contextual variables, accounted for 49% of teacher self-efficacy for novice teachers (Tschannen-Moran & Woolfolk Hoy, 2007).
**Teacher-efficacy**

Teacher-efficacy is defined as a teacher’s belief that s/he has the capacity to influence how well students learn, or their beliefs about their ability to teach (Guskey & Passaro, 1994; Tschannen-Moran, Woolfolk Hoy, & Hoy, 1998). This construct emerged from work done at the RAND research institute—originally conceived of as the extent to which teachers believed that they could “control the reinforcement of their actions,” or whether “control of reinforcement lay within themselves or within the environment” (Tschannen-Moran et al. 1998, p. 202). Through the perspective of social learning theory, researchers assumed student performance and motivation acted as significant reinforcers for teaching behavior (Tschannen-Moran et al., 1998). Approximately a decade after this, Bandura began working through the social cognitive framework, identifying teacher-efficacy as similarly to self-efficacy: a process in which people construct beliefs about their capacity to perform at a given level of attainment (Bandura, 1977; Tschannen-Moran, 1998). According to Tschannen-Moran and colleagues, these two perspectives are largely the dominant perspectives on teacher-efficacy (1998).

From the very earliest studies, there was evidence to suggest that teacher-efficacy was strongly correlated with student achievement (Tschannen-Moran et al. 1998). Ashton and Webb pioneered a qualitative study in the area of teacher self-efficacy using vignettes and asking teachers to rate the possible responses (1984). Additional findings from this study revealed that stress and teacher efficacy were moderately correlated (-.05 to -.82, with an average of -.39), however it was determined that stress could not be used as a proxy for teacher efficacy and vice versa (Ashton & Webb, 1984; Tschannen-Moran & Woolfolk Hoy, 2001).
In several studies, student achievement scores were significantly related to teacher-efficacy (Gibson & Dembo, 1984; Moore & Esselman, 1992; Tschannen-Moran et al., 1998). Specifically, personal teaching efficacy (PTE) was strongly related to student achievement, including reading and math scores (Ross, 1992; Tschannen-Moran et al. 1998). Additionally, and importantly, teacher’s sense of efficacy for teaching was related to students’ affective development and attitudes toward school, content area, and teachers (Tschannen-Moran et al., 1998; Woolfolk, Rosoff, and Hoy, 1990).

Teacher-efficacy affects issues in classroom management (Poulou, 2007), goal setting, and motivation (Tschannen-Moran, et al. 1998). Teachers with higher levels of self-efficacy for teaching are also more likely to experiment with new teaching methods, as if providing a sense of security or bravery in the classroom (Berman, McLaughlin, Bass, Pauly, & Zellman, 1977; Guskey, 1988; Stein & Wang, 1988). There is evidence to support the notion that teachers with higher self-efficacy are also more likely to engage in adaptive instructional practices, such as adopting mastery-goal orientations in their classrooms (Wolters & Daugherty, 2007). Teachers with higher levels of teaching efficacy also have a tendency to be less critical of their students and more supportive, both instructionally and emotionally to their students, are more enthusiastic, and have more commitment to the field than other teachers (Ashton & Webb, 1986; Gibson & Dembo, 1984; Tschannen-Moran, et al., 1998).

Teacher self-efficacy varies according to the number of years of experience a teacher has in the classroom in a non-linear fashion, increasing through early and mid-career, and then declining into late career (Klassen & Chiu, 2010), as well as varying according to which grade a teacher teaches; teachers in higher grades (middle and high
school) report lower teacher self-efficacy for engaging students meaningfully in the learning process (Midgley, Anderman, & Hicks, 1995; Wolters & Daugherty, 2007). This finding, coupled with teacher reports of using more performance-oriented teaching practices in the same grade levels, suggests there may be a relation between teacher self-efficacy and instructional practices that promote certain goal orientations (Wolters & Daugherty, 2007). Teacher-efficacy is a contributing factor in a teacher’s decision to stay in the field or leave due to burnout (Burley, Hall, Villeme, & Brockmeier, 1991; Fives, Hamman, & Olivarez, 2007; Glickman & Tamashiro, 1982, Skaalvik & Skaalvik, 2007).

Teacher self-efficacy can help determine if a teacher will persevere in the face of difficulty or give up. Teacher self-efficacy also helps to determine if a teacher will try a new strategy with a student who did not grasp the concept the first time around, or rely on old strategies that s/he may find more comfortable (Gibson & Dembo, 1984). These are not small distinctions—they have a lasting impact on the education system and on individual students, and their academic progress.

Teacher-efficacy, like its parent self-efficacy, relies heavily on an individual’s prior experiences, and these experiences are critical to teacher education programs. Particularly challenging for teacher preparation programs is how to handle the demand for experience that preservice teachers need before entering the field. Preservice teachers rely on the information they have acquired from their coursework, and to a lesser extent, their field experiences, whereas inservice teachers rely on their experiences in the classroom seemingly at the expense of the theory and research they learned in their classes (Burke & Shutherland, 2004). Almog and Schechtman (2007) found that teachers were more likely to rely on restrictive responses (those that are punitive, restrictive,
threatening, preaching, withholding, or punishing) and less likely to rely on helpful responses, contributing to increased distance between students and teachers and minimizing the amount of challenge the student experienced. There was a positive correlation between helpful response styles and teacher self-efficacy. Teachers with democratic beliefs were also more likely to rely on a helpful response style (Almog & Schechtman, 2007).

**General and personal teacher efficacy**

Specifically, the RAND studies differentiated between general teaching efficacy (GTE, “*When it comes right down to it, a teacher really can’t do much because most of a student’s motivation and performance depends on his or her home environment*”) and personal teaching efficacy (PTE, “*If I try really hard, I can get through to even the most difficult or unmotivated students.*”). Together, these two types of teacher-efficacy accounted for up to 46% of the variance in student achievement in the earliest studies (Ashton & Webb, 1986; Tschannen-Moran et al. 1998).

Support for the differentiation between personal teacher-efficacy (teacher self-efficacy) and general teaching efficacy (teaching efficacy) exists elsewhere in the literature. A distinction was made between the constructs that reflected individuals’ beliefs about their own abilities (Personal efficacy or Outcome efficacy) versus their beliefs about the effectiveness of teaching (Teaching Efficacy) in work by Soodak and Podell in the 1990s. These researchers argued that an individual can hold conceptually different beliefs about these aspects of efficacy simultaneously, and that only personal teacher self-efficacy was related to outcomes (Soodak & Podell, 1996).
Gibson and Dembo found that teachers with higher teaching efficacy on both personal and general teaching efficacy measures were less likely to criticize students, more likely to persist with a failing student, and more likely to divide students into small groups for instruction (Gibson & Dembo, 1984). Teacher-efficacy has also been correlated with positive attitudes about teaching, willingness to implement innovations, and teacher stress, addressed in the next section (Greenwood et al., 1990; Guskey 1984). Other work in this area found teacher-efficacy was linked to instructional experimentation, instructional innovation, organization, enthusiasm, and teacher fairness (Allinder, 1994; Tschannen-Moran et al., 1998).

**Teacher Self-Efficacy and Stress**

Klassen (2011) found that “occupational commitment” was directly influenced by classroom stress and teacher self-efficacy for instructional strategies for practicing teachers, and “intention to leave the profession” was directly influenced by teaching grade level (teachers of younger students were more likely to intend to quit) overall stress (greater stress was related to greater intention to quit), and occupational commitment (less occupational commitment was related to greater intention to quit) (Klassen, 2011, p. 125). In another study, teacher self-efficacy was strongly negatively correlated with teacher burnout (Skaalvik, 2007). Having low teacher self-efficacy will likely contribute to a person’s maladaptive pattern of behavior, dwelling on the deficiencies in the situation, magnifying the severity of the threat, or even resorting to escapist modes of coping that will increase stress and strain (Bandura, 1997). However, it cannot be determined if teacher self-efficacy leads to teacher burnout or if burnout predicts teacher self-efficacy. Most likely, there is a reciprocal relation between the two (Skaalvik, 2007).
The relation between teacher self-efficacy and stress is complex. Teacher self-efficacy is often used in research studies to predict stress or stressful outcomes (e.g. Klassen, 2010). However, Social Cognitive Theory (specifically reciprocal determinism) posits that the relation between stress and teacher self-efficacy is dynamic and synergistic (Bandura, 1997). Stress, defined as a physiological arousal state, can be perceived as a threat or a stimulant depending on the level of efficacy of the individual (Karadag & Baloglu, 2009). Alternatively, stress could affect an individual’s self-efficacy, according to the manner in which the stress event serves as a source of efficacy for the individual.

Because stress is a relationship between the environment and individual and does not reside entirely within the individual (Lazarus, 1988), it will be examined as a predictor of self-efficacy for teachers. This is due to the likelihood that the stress perceived by teachers would influence their self-efficacy beliefs, depending upon whether the individual perceived the event as within or out of their control.

**Self-Determination Theory**

Self-Determination Theory (SDT) posits that individuals vary in their intent to act from intentional to non-intentional. The distinguishing characteristic that SDT possesses is the extent to which an individual is acting of her own volition or whether her intent is being controlled to some extent (Deci, Vallerand, Pelletier, & Ryan, 1991; Deci & Ryan, 2000). The individual acting “wholly volitionally” is self-determined, implying the person perceives an internal locus of causality. The individual acting without volition perceives the locus of causality to be external to her. These differences are significant and represent different components of regulation; that is, the self-determined individual
possesses a regulatory process of choice, whereas the controlled individual possesses a regulatory process that is compliant (Deci et al., 1991; Ryan & Deci, 2000).

SDT incorporates needs-based aspects of human life in its framework. Needs are central to human behavior and the human psyche, and the opportunity to satisfy these needs contribute to an individual’s likelihood to be motivated (Deci et al., 1991; Deci & Ryan, 2000). According to the theory, there are three innate psychological needs each human possesses: they are the need for competence, the need for relatedness, and the need for autonomy or self-determination (Deci et al. 1991; Deci & Ryan, 2000; Ryan & Deci, 2000). Competence is described as an individual’s need for attainment of various outcomes and efficacy for doing so; relatedness is described as an individual’s need for secure and satisfying connections with others, and autonomy is described as the ability to regulate one’s own actions and behaviors (Deci et al., 1991).

Perhaps the most pervasive and wide-reaching constructs in psychology and education, arguably, are the notions of intrinsic versus extrinsic motivation. Deci and colleagues describe intrinsically motivated behaviors as those that individuals engage in for their own sake, “for the pleasure and satisfaction derived from their performance” (Deci et al., 1991, p. 328). This differs from extrinsically motivated behaviors, which individuals engage in because of some external reward or consequence; these behaviors are described as “instrumental” in nature and are related to some other, separable, outcome (Deci et al., 1991; Deci & Ryan, 2000).

Internalization is term used to describe the process of transforming regulation from the external to the internal. According to SDT, an individual is motivated to a greater degree with increased internalization of processes. In other words, the more
“intrinsic” an individual’s motivation is, the more self-determined he is; being self-determined is a positive outcome. When this process of regulation is fully integrated into the self, the individual is said to be optimally self-determined (Deci et al., 1991; Deci & Ryan, 2000; Ryan & Deci, 2000; Ryan & Deci, 2006).

**Control and Autonomy**

Autonomy is an inner endorsement of one’s own actions, or the sense that one’s actions emanate from oneself and are one’s own (Deci & Ryan, 1987). Autonomous individuals perceive themselves as the initiators of their own behavior, selecting desired outcomes and deciding how to achieve them (Deci & Ryan, 1987; Ryan & Deci, 2006). Individuals who perceive that they are controlled, on the contrary, do not feel they are the initiators of their own behaviors. These individuals feel some level of coercion to engage in the behavior, regardless of the extent to which they agree with the intended outcomes (Deci & Ryan, 1987).

A supportive environment is crucial to the individual. Deci and Ryan (1987) note that of primary concern to the issue of autonomy and control is the context that an individual perceives around him, and whether that context is supportive of autonomy. There are a series of events that research has shown could potentially have an effect on the extent to which an individual will feel autonomy-supported or controlled. They include rewards (Deci & Ryan, 1980; 1987; Harackiewicz, 1979), threats, deadlines, evaluation, surveillance (Lepper & Greene, 1975), choice (Zuckerman, Porac, Lathin, & Deci, 1978), and positive feedback (Ryan, 1982; Vallerand & Reid, 1984). Additionally, a controlling event can influence an individual’s self-determination: at risk include the individual’s interest or enjoyment (Ryan et al. 1983), creativity (Amabile, 1979),

Rewards can be perceived as autonomy-supporting or controlling depending upon the orientation of the individual administering them (Joussemet, Koestner, Lekes, & Houfourt, 2004). This supports the argument in the current study indicating that teachers’ autonomy-supportiveness and controllingness may have an impact on students’ perceptions of rewards and highlighting the importance of teachers’ orientations toward these constructs. Additionally, this study speaks to teachers’ perceptions of the sanctions that are administered under controlling conditions such as high-stakes tests; detrimental effects of controlling rewards were detected almost immediately among study participants (Joussemet et al., 2004).

The presence of threats, deadlines, surveillance, and evaluation resulted in significantly decreased interest in the task for students (Lepper & Greene, 1975). As these authors noted, “The knowledge that one’s performance at a task is being observed and evaluated by someone else, even with there is no explicit expectation of any tangible reward for engaging in the activity, appears sufficient to decrease later interest in the task” (p. 484). The implications of this for the current study are serious due to the heavily influential nature of the high-stakes test: most agree that testing is rife with threats, deadlines, surveillance, and evaluation—all highly controlling activities.

**Self-Determination and High-Stakes Testing**

The effects of policies such as high-stakes testing on motivation have been documented theoretically (Ryan & Weinstein, 2009). Events such as assessments,
feedback, rewards, punishments, and curriculum decisions can be perceived as informational, controlling, or amotivating, which will shape the motivation of individuals (Ryan & Weinstein, 2009). An amotivated person has neither autonomous nor controlled motivations (Ratelle, Guay, Vallerand, Larose, & Senecal, 2007). Tests can be perceived as informational, a positive outcome for motivation, if they enhance feelings of competence and autonomy (Ryan & Weinstein, 2009). Reports from teachers to date generally do not indicate that they feel high-stakes tests are informational (Ryan & Weinstein, 2009). Tests and other events can be perceived as controlling if they include some element of pressure, or as an attempt to control behavior, such as tests that include rewards or sanctions (Ryan & Weinstein, 2009). Controlling events are likely to undermine motivation as individuals will be more likely to exert as little effort as possible to get by, resulting in a de-valuing of the event for the individual, enhancement of superficial forms of learning, and ego-focused motivation (Ryan & Weinstein, 2009). As much of the literature, both academic and otherwise, has indicated, testing exerts great pressure not only on teachers but students and administrators as well. Finally, tests or other events can be perceived as amotivating when they are too challenging, discouraging, or are overwrought with negative feedback, which elicits helplessness or incompetence in individuals (Ryan & Weinstein, 2009).

Increased pressure to raise test scores is one of the more salient issues reported by teachers (Abrams et al., 2003; Reese, Gordon, & Price, 2004; Urdan & Paris, 1994). Referencing the Massachusetts Comprehensive Assessment System (MCAS), one teacher reported, “The MCAS is one of the reasons I’m leaving, because of the pressure it puts on the kids and everyone else. And, we don’t even have the right materials to teach for it”
(Nichols & Berliner, 2007, p.145). Many times, teachers are threatened proximally with the possibility of termination, transfer, or pay adjustments, and they are threatened more distally with the possibility that the school district will be sanctioned in some way. Teachers are more likely to engage in controlling instructional practices when faced with increased pressure, including lecturing, criticizing, and directing (Reeve, 2009; Ryan & Weinstein, 2009). Three conditions characterize the controlling teaching style: the tendency of teachers to adopt only their own perspective; the tendency to intrude into students’ thoughts, feelings, and actions; and increasing pressure for students to think, feel, or behave in a particular way (Reeve, 2009). Researchers have identified three kinds of pressure teachers face: pressure from above (such as policy and administration), pressure from below (from students; poor classroom management would be an example), and pressure from within, or the pressure that arises from the teachers’ own values, beliefs, and personality (Pelletier, Seguin-Levesque, & Legault, 2002; Reeve, 2009).

Nichols and Berliner report two main problems with using pressure to motivate people (2007). First, using pressure often results in the unintended consequence that individuals (teachers) will short-circuit the process in favor of the outcome, and this may not contribute to lasting results (Nichols & Berliner, 2007). This phenomenon is not reflective of teachers’ professional shortcomings--it is reflective of the shortcomings in the system. As reported in Nichols and Berliner, Margaret Spellings, former Secretary of Education, is on the record stating, “Bribery is the cornerstone of good parenting. And good management.” (p. 148). The problem with this from a motivational perspective is that bribery (extrinsic rewards) undermines intrinsic motivation, and this style of

The second problem with using pressure to motivate, according to Nichols and Berliner (2007), is that teachers’ work conditions are being altered in ways that are beyond their control, undermining autonomy and efficacy. In many work environments an individual’s performance is judged according to her own work or productivity. In education, with the implementation of high-stakes testing, “teachers are put in the precarious position of putting their livelihoods squarely on the shoulders of their students” (Nichols & Berliner, 2007, p. 150). While teachers may understand that their students may do their very best on the test, they are aware this is irrelevant to the testing policy. This drastically undermines teachers’ sense of autonomy, professionalism, job satisfaction, and empowerment (Nichols & Berliner, 2007).

Another problem with pressure exists: the pressure that teachers and administrators are under to improve their schools’ ratings serves as a perverse incentive for schools to weed out underperformers. Often, this is done by classifying low achievers, minority students, and non-English speaking students as special education students, urging these students to move to other schools, failing them, or encouraging them to drop out (Darling-Hammond, 2007; Ryan & Weinstein, 2009). The added sanctions placed on schools have not curtailed these behaviors, but rather exacerbated them, adding to teacher and administrator tendencies to alter test scores and also drop out and attrition rates (Ryan & Weinstein, 2009).
Summary of literature

High-stakes testing presents a controlling context for both teachers and students. Controlling environments cause individuals to experience increased stress, reductions in self-efficacy, greater likelihood of adopting a controlling-orientation, and that these issues may also influence each other. The literature reviewed in this chapter highlight aspects of the research that has been conducted on stress, teacher self-efficacy, and autonomy and control, and ties together the various aspects of these constructs that are especially relevant to teachers.

Stress is defined as a dynamic relationship between an individual and their personal characteristics and their environment (Lazarus, 1990) that presents particular challenges for measurement (Lazarus, 2000) due to its complex nature. Teacher stress is a negative emotional experience that is triggered by teachers’ perception that their work situation is threatening in some way (Kyriacou, 2001). Stress is known to affect individuals’ beliefs and motivation (e.g., Amabile, Hill, Hennessey, & Tighe, 1994; Bandura 1982; Kyriacou, 2001, Weinstein, 2010).

Teacher beliefs are known to affect their behavior (Tschannen-Moran & Woolfolk Hoy, 2001). Teacher self-efficacy in particular is one of the most predictive and widely-studied forms of beliefs that currently exist in the literature (Bandura, 2001; Tschannen-Moran & Woolfolk Hoy, 2001). Teacher-efficacy is defined as a teacher’s belief that s/he has the capacity to influence how well students learn, (Guskey & Passaro, 1994; Tschannen-Moran, Woolfolk Hoy, & Hoy, 1998), and is related to student achievement, classroom management, and teacher persistence ((Gibson & Dembo, 1984; Moore & Esselman, 1992; Tschannen-Moran et al., 1998). The sources of self-efficacy are mastery
experiences, vicarious experiences, verbal persuasion, and physical/affective responses (Bandura, 2001). These beliefs are likely to vary according to the varying conditions that they are situated within, and therefore are central constructs to the current study.

Equally central to the purpose of the current study are teachers’ autonomy-supportiveness and controllingness. Autonomous individuals perceive regulate their own behavior, choosing their own outcomes and deciding how to achieve them (Deci & Ryan, 1987; Ryan & Deci, 2006). Controlled individuals feel they have to do the behavior, regardless of the extent to which they agree with the intended outcomes (Deci & Ryan, 1987). Tests and other events can be perceived as controlling if they include some element of pressure, or as an attempt to control behavior, such as tests that include rewards or sanctions (Ryan & Weinstein, 2009). Teachers are more likely to engage in controlling instructional practices when faced with increased pressure, including lecturing, criticizing, and directing (Reeve, 2009; Ryan & Weinstein, 2009). Therefore, it was critical to include the examination of teachers’ autonomy-supportiveness and controllingness in this study.

The present study

The present study is an examination of teachers’ beliefs as a result of high-stakes testing policies. Specifically, teachers’ self-efficacy beliefs are examined in the form of self-reports that ask currently practicing teachers to report their own teacher efficacy beliefs generally, and then to report their efficacy beliefs in the weeks leading up to the test. Along with this, teachers are asked to rate several vignettes intended to measure their tendency to adopt an autonomy-supportive orientation or a controlling orientation in the classroom. Specifically, the following research questions are identified:
R1: What is the relation between teachers’ perceptions of stress due to high-stakes tests and teachers’ self-efficacy for teaching?

R2: What is the relation between teachers’ perceptions of stress as a result of high-stakes tests and their autonomy-supportiveness and controllingness?
CHAPTER THREE

METHODS

The purpose of the present study is to examine teacher motivational beliefs in relation to perceived stress due to high-stakes tests. Whereas much is known about the effects of high-stakes tests on student achievement (Amrein & Berliner, 2002), curricular design (Crocco & Costigan, 2007), and teacher satisfaction in their jobs (Darling-Hammond, 2007), little is known about how the testing policies have impacted teacher motivation from an empirical perspective. Specifically, this study focuses on the following three hypotheses:

• $H_1$: Teachers that report higher levels of stress as a result of high-stakes tests are more likely to report lower levels of teacher self-efficacy.

• $H_2$: Teachers that report higher levels of stress as a result of high-stakes tests are less likely to report adopting an autonomy-supportive orientation toward teaching.

• $H_3$: Teachers that report higher levels of stress as a result of high-stakes tests are more likely to report adopting a controlling-orientation toward teaching.
Participants and Procedures

In order to investigate the hypotheses presented above, it was necessary to gather information from practicing teachers of all varieties; thus, the sample for this study was teachers of all grades, of all content areas, from all geographic regions of the country, with all levels of experience in the classroom, and with a wide variety of education and certification. Additionally, I sought teachers who represented a variety of schools and neighborhoods, political persuasions, races and genders. The sample that was recruited possessed these various characteristics and more. This enabled me to work with data that represented the teaching population as closely as possible. The sample and its demographic characteristics are described in detail below.

Teachers from all grades levels and content areas were invited to participate in this study. Recruitment for this study occurred between July and October of 2011. IRB approval was obtained to collect data from practicing teachers throughout the United States through two primary vehicles: email recruitment to teachers directly, and snowball recruitment through participating teachers. Four undergraduate students assisted me with the recruitment of teachers for the study, identifying prospective participants around the country from school district and education department websites, collecting email addresses and delivering them to me in order to facilitate recruitment. These students were trained to identify participants from rural, urban, and suburban areas as well as from varying content areas, grade levels, and experience levels. For recruitment, each of the four students was assigned 11 states, and I targeted the remaining 6 states. In the recruitment email, participants were asked to volunteer for the study and then forward the
information to their peers and colleagues. Teachers were provided a link to the online survey and data was subsequently entered into an online survey tool, Surveygizmo. The incentives for participation included entry into a drawing for one of four $50 Barnes and Noble gift cards or one Apple iPad. The online survey took approximately 20 minutes to complete.

The four undergraduate assistants and I spent approximately three hours per week from July 2011 through October 2011 identifying prospective participants for the study. I took sole responsibility for sending all emails to prospective participants. Although detailed figures were not recorded regarding how many recruitment emails were sent, estimates range between 20,000 and 24,000 across 50 states. This figure does not include the estimation of teachers who responded as a result of the snowball portion of the recruitment. Of those, approximately 750 chose to respond, and of those, 550 were completed. Listwise deletion was utilized to account for missing data, resulting in a final \( N \) of 529.

The recruitment email can be found in Figure 3.1, located in the Appendix.

**The Sample**

All demographic information on the study participants can be found in Table 3.1. Due to IRB protocol, participants were not required to answer every item on the survey. Of the total \( N \), 127 self-reported that they most closely identified as male (24%), 399 (75.4%) identified as female, and three (0.6%) did not respond.

In order to determine the racial breakdown of the sample, the categories utilized by the US Census Bureau were selected to ask participants to identify their race and
ethnicity. Of the total N of 529, four-hundred and ninety-four (494) of the respondents identified as White (93.4%), nine (9) identified as Black (1.7%), three (3) as Asian (0.6%), nine (9) as Hispanic (1.7%), and eight (8) did not report (1.5%). One hundred sixteen teachers (116) reported teaching in an inner city or urban setting (21.9%), 180 reported teaching in a suburban setting (34%), and 228 reported teaching in a rural setting (43.1%). Other descriptive data can be found in Table 3.1.

Just over a third (35.3%; N = 187) earned a Bachelor’s degree, 320 (60.5%) possessed a Master’s degree, and 19 (3.6%) had earned a Ph.D. or Specialist degree. Three respondents did not provide an answer (0.6%).

Of the total 529 teachers, fifteen (15) of them had one year of teaching experience or less (2.8%), twenty-one (21) held 2 years of teaching experience (4%), eighteen teachers (18; 3.4%) held three years and the same amount had four years experience, and thirty-three (33; 6.2%) had five years of teaching experience. One hundred forty (140) teachers held between 6-15 years of teaching experience (31.1%). The remainder held over 15 years of teaching experience. The breakdown of teaching experience can be seen in detail in Table 3.1.

Nearly one-third of the teachers in this sample reported teaching in a suburban area (180; 34%). Only one hundred-sixteen (116) reported teaching in an urban setting (21.9%) whereas two hundred twenty eight (228; 43.1%) reported teaching in a rural setting.

Teachers in this sample reported that their teacher evaluations did not depend on student performance on the test 43.2% of the time ($n = 227$). For 206 of these teachers
(39.2%), teacher evaluations were dependent upon student performance on the test some of the time, whereas 75 (14.2%) teachers reported that their teacher evaluations were dependent on student performance on the test nearly half of the time. For eighteen of these teachers \(n = 18; \ 3.4\%\) over half of their teacher evaluation depended upon student performance on the test. Means, standard deviations, and total \(N\) for all demographic items can be found in Table 3.1.
Table 3.1.
Demographic data, frequencies, and applicable means

<table>
<thead>
<tr>
<th>With which gender do you most closely identify?</th>
<th>N</th>
<th>%</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>127</td>
<td>24</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>399</td>
<td>75.4</td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>3</td>
<td>.6</td>
<td>529</td>
</tr>
<tr>
<td>With which race or ethnicity do you most closely identify?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>494</td>
<td>93.4</td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>9</td>
<td>1.7</td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>3</td>
<td>.6</td>
<td></td>
</tr>
<tr>
<td>Native American/Pacific Islander</td>
<td>6</td>
<td>.11</td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>9</td>
<td>1.7</td>
<td></td>
</tr>
<tr>
<td>Arab</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>8</td>
<td>1.5</td>
<td>529</td>
</tr>
<tr>
<td>How many years of teaching experience do you have?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>15</td>
<td>2.8</td>
<td>16.17</td>
</tr>
<tr>
<td>2</td>
<td>21</td>
<td>4.0</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>18</td>
<td>3.4</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>18</td>
<td>3.4</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>33</td>
<td>6.2</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>26</td>
<td>4.9</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>17</td>
<td>3.2</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>17</td>
<td>3.2</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>10</td>
<td>1.9</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>23</td>
<td>4.3</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>13</td>
<td>2.5</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>13</td>
<td>2.5</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>13</td>
<td>2.5</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>8</td>
<td>1.5</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>24</td>
<td>4.5</td>
<td></td>
</tr>
<tr>
<td>16-20</td>
<td>78</td>
<td>14.7</td>
<td></td>
</tr>
<tr>
<td>21-29</td>
<td>118</td>
<td>22.3</td>
<td></td>
</tr>
<tr>
<td>Over 30</td>
<td>61</td>
<td>11.5</td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>18</td>
<td>3.4</td>
<td>529</td>
</tr>
<tr>
<td>Bachelor’s</td>
<td>187</td>
<td>35.3</td>
<td></td>
</tr>
<tr>
<td>Master’s</td>
<td>320</td>
<td>60.5</td>
<td></td>
</tr>
<tr>
<td>Ph.D.</td>
<td>19</td>
<td>3.6</td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>3</td>
<td>.6</td>
<td>529</td>
</tr>
</tbody>
</table>

1 For all tables, percentages are given for figures that add up to 100%. In some cases, teachers may have selected more than one option, yielding a total percentage of more than 100%.
Table 3.1.
Demographic data, frequencies, and applicable means continued.

<table>
<thead>
<tr>
<th>How does your school usually score on high-stakes tests?</th>
<th>N</th>
<th>%</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>-Below the appropriate level (e.g., we are a “failing” school or in “continuous improvement”).</td>
<td>156</td>
<td>29.5</td>
<td></td>
</tr>
<tr>
<td>-Right about at the appropriate level.</td>
<td>194</td>
<td>36.7</td>
<td></td>
</tr>
<tr>
<td>-My school generally performs well on the standardized tests.</td>
<td>174</td>
<td>32.9</td>
<td></td>
</tr>
<tr>
<td>-Neither.</td>
<td>5</td>
<td>.9</td>
<td>529</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Which content areas do you primarily teach?</th>
<th>N</th>
<th>%</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>-Tested content areas (math, reading).</td>
<td>402</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>-Non-tested content areas.</td>
<td>251</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>-Both.</td>
<td>115</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>-Neither.</td>
<td>23</td>
<td>-</td>
<td>529</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>What grades do you teach?</th>
<th>N</th>
<th>%</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>K</td>
<td>37</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>48</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>45</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>61</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>55</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>60</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>56</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>85</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>92</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>170</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>196</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>211</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>200</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Special</td>
<td>69</td>
<td>-</td>
<td>529</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>In what kind of neighborhood is the school you teach located?</th>
<th>N</th>
<th>%</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>-Urban</td>
<td>116</td>
<td>21.9</td>
<td></td>
</tr>
<tr>
<td>-Suburban</td>
<td>180</td>
<td>34</td>
<td></td>
</tr>
<tr>
<td>-Rural</td>
<td>228</td>
<td>43.1</td>
<td></td>
</tr>
<tr>
<td>-Missing</td>
<td>5</td>
<td>.9</td>
<td>529</td>
</tr>
</tbody>
</table>

2 For all tables, percentages are given for figures that add up to 100%. In some cases, teachers may have selected more than one option, yielding a total percentage of more than 100%.
The administration in my school is supportive of my work as a teacher.

<table>
<thead>
<tr>
<th>Agreement Level</th>
<th>Count</th>
<th>Percentage</th>
<th>Total</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strongly Agree</td>
<td>170</td>
<td>32.1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Agree</td>
<td>239</td>
<td>45.2</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Neutral</td>
<td>68</td>
<td>12.9</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Disagree</td>
<td>39</td>
<td>7.4</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Strongly Disagree</td>
<td>11</td>
<td>2.1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Missing</td>
<td>2</td>
<td>.4</td>
<td>529</td>
<td>100%</td>
</tr>
</tbody>
</table>

What is your average class size?

<table>
<thead>
<tr>
<th>Class Size</th>
<th>Count</th>
<th>Percentage</th>
<th>Total</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Under 20</td>
<td>110</td>
<td>20.8</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>20-25</td>
<td>227</td>
<td>42.9</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>25-35</td>
<td>185</td>
<td>35</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Over 35</td>
<td>7</td>
<td>1.3</td>
<td>529</td>
<td>100%</td>
</tr>
</tbody>
</table>
Table 3.1 continued
*States represented by sample*

<table>
<thead>
<tr>
<th>State</th>
<th>N</th>
<th>%</th>
<th>State</th>
<th>N</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alabama</td>
<td>5</td>
<td>.9</td>
<td>Montana</td>
<td>7</td>
<td>1.3</td>
</tr>
<tr>
<td>Alaska</td>
<td>8</td>
<td>1.5</td>
<td>Nebraska</td>
<td>11</td>
<td>2.1</td>
</tr>
<tr>
<td>Arizona</td>
<td>13</td>
<td>2.5</td>
<td>Nevada</td>
<td>9</td>
<td>1.7</td>
</tr>
<tr>
<td>Arkansas</td>
<td>3</td>
<td>.6</td>
<td>New Hampshire</td>
<td>2</td>
<td>.4</td>
</tr>
<tr>
<td>California</td>
<td>11</td>
<td>2.1</td>
<td>New Jersey</td>
<td>8</td>
<td>1.5</td>
</tr>
<tr>
<td>Colorado</td>
<td>9</td>
<td>1.7</td>
<td>New Mexico</td>
<td>5</td>
<td>.9</td>
</tr>
<tr>
<td>Connecticut</td>
<td>7</td>
<td>1.3</td>
<td>New York</td>
<td>7</td>
<td>1.3</td>
</tr>
<tr>
<td>Delaware</td>
<td>10</td>
<td>1.9</td>
<td>North Carolina</td>
<td>11</td>
<td>2.1</td>
</tr>
<tr>
<td>Florida</td>
<td>9</td>
<td>1.7</td>
<td>North Dakota</td>
<td>3</td>
<td>.6</td>
</tr>
<tr>
<td>Georgia</td>
<td>25</td>
<td>4.7</td>
<td>Ohio</td>
<td>26</td>
<td>4.9</td>
</tr>
<tr>
<td>Hawai’i</td>
<td>8</td>
<td>1.5</td>
<td>Oklahoma</td>
<td>17</td>
<td>3.2</td>
</tr>
<tr>
<td>Idaho</td>
<td>12</td>
<td>2.3</td>
<td>Oregon</td>
<td>16</td>
<td>3.0</td>
</tr>
<tr>
<td>Illinois</td>
<td>7</td>
<td>1.3</td>
<td>Pennsylvania</td>
<td>1</td>
<td>.2</td>
</tr>
<tr>
<td>Indiana</td>
<td>22</td>
<td>4.2</td>
<td>Rhode Island</td>
<td>4</td>
<td>.8</td>
</tr>
<tr>
<td>Iowa</td>
<td>8</td>
<td>1.5</td>
<td>South Carolina</td>
<td>16</td>
<td>3.0</td>
</tr>
<tr>
<td>Kansas</td>
<td>23</td>
<td>4.3</td>
<td>South Dakota</td>
<td>8</td>
<td>1.5</td>
</tr>
<tr>
<td>Kentucky</td>
<td>12</td>
<td>2.3</td>
<td>Tennessee</td>
<td>9</td>
<td>1.7</td>
</tr>
<tr>
<td>Louisiana</td>
<td>6</td>
<td>1.1</td>
<td>Texas</td>
<td>11</td>
<td>2.1</td>
</tr>
<tr>
<td>Maine</td>
<td>28</td>
<td>5.3</td>
<td>Utah</td>
<td>8</td>
<td>1.5</td>
</tr>
<tr>
<td>Maryland</td>
<td>13</td>
<td>2.5</td>
<td>Vermont</td>
<td>9</td>
<td>1.7</td>
</tr>
<tr>
<td>Massachusetts</td>
<td>14</td>
<td>2.6</td>
<td>Virginia</td>
<td>10</td>
<td>1.9</td>
</tr>
<tr>
<td>Michigan</td>
<td>14</td>
<td>2.6</td>
<td>Washington</td>
<td>14</td>
<td>2.6</td>
</tr>
<tr>
<td>Minnesota</td>
<td>2</td>
<td>.4</td>
<td>West Virginia</td>
<td>9</td>
<td>1.7</td>
</tr>
<tr>
<td>Mississippi</td>
<td>5</td>
<td>.9</td>
<td>Wisconsin</td>
<td>8</td>
<td>1.5</td>
</tr>
<tr>
<td>Missouri</td>
<td>9</td>
<td>1.7</td>
<td>Wyoming</td>
<td>13</td>
<td>2.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Missing</td>
<td>4</td>
<td>.8</td>
</tr>
</tbody>
</table>
Instruments

Teachers completed an online survey consisting of 98 total items. The survey consisted of five sections: demographics, the *Perceived Stress Due to High-Stakes Tests* scale (PS-HST) (Dawson, 2011), the *Teacher Self-Efficacy Scale* (TSES) (Tschannen-Moran & Woolfolk Hoy, 2001), the *Teacher Autonomy and Control Scale* (TACS) (Dawson, 2011), and the *Problems in Schools Questionnaire* (PIS) (Deci, Schwartz, Sheinman, and Ryan 1981). The *Teacher Self-Efficacy Scale* was administered twice—the first administration was qualified with the following specification: “*Generally, throughout the school year...*” and the second administration was qualified with the following: “*In the weeks leading up to the test...*”

**Demographic Items.** The items listed in this section have been described above; see Table 3.1.

**The Perceived Stress due to High-Stakes Tests Scale.** In order to measure teachers’ perceptions of the pressure and stress due to high-stakes testing, I created sixteen items to capture the various dimensions of teaching framed from the perspective of stress and testing. I gathered several doctoral students in Educational Psychology in the spring of 2011 for a focus group to determine the kinds of teaching tasks that are most likely to be affected by testing. Many of these doctoral students were former practicing teachers and therefore held firsthand experiences with the kinds of issues teachers experienced as a result of high-stakes tests. The focus group began with a discussion about pressure due to high-stakes tests. Specifically, I solicited discussion surrounding tests and lesson planning, administration, parents, behavior management, professional development,
media, and personal stress. Thirty-two (32) items were written as a result of the focus group targeting the areas of instruction and lesson planning (8 items), administration and principals, testing coordinators and curriculum coordinators (4 items), emotional well-being (3 items), school climate (3 items), behavior management (3 items), media/society (3 items), teacher/professional relationships (5 items), and self (3 items). These items were originally designed then to be responded to three times each: the first to a prompt “Generally, throughout the school year;” the second prompt “In the weeks leading up to the test;” and the third “How much stress does this cause you?”

Piloting of the measures revealed the survey, originally written as described above, required approximately 57 minutes to complete. This was not acceptable, so I discarded these items and the three-part prompts for a shorter, less unwieldy format. The eight items that were finalized for the study originated from those constructed in the focus group, and resulted from discussions with other scholars, practitioners, and mentors. These discussions resulted in agreement that the items that most globally represent each of the targeted areas would be included, and a two-part format was preferable to the three-part format. This decision to reduce the items was a pragmatic decision based on the practicality of teachers’ ability to complete the survey, and not a decision based on content.

All of these items were scored on a 5-point Likert-type scale from 1 (Strongly Disagree) to 5 (Strongly Agree). Each final item contained two parts: the first was designed to target some aspect of teaching that is most likely to be interrupted during test-preparation time or otherwise susceptible to pressures related to high-stakes tests. For
example, “In my school, I feel pressure to improve standardized test scores.” The second part of the item, “This causes me stress.” This two-part item design was intended to minimize potential confounding of meaning in the items; for example, asking teachers “Does pressure to improve standardized test scores cause you stress?” or “How much stress do you experience as a result of pressure to utilize practice standardized tests?” are examples of double-barreled items, and using items such as these could result in measurement error. In order to be as precise as possible in the measurement of stress due to high-stakes tests, I determined that it would be necessary to measure the extent to which teachers actually perceive tests as disruptive before measuring their self-reported stress. Separating the constructs, or latent variables, into two discrete items was the most practical way to handle this. This decision resulted in two intended potential subscales: the Known Issues subscale, designed to measure actual observable changes teachers perceive as a result of the test, and the Perceived Stress subscale, designed to measure the extent to which teachers perceive these issues as stressful.

When the final eight items on the Known Issues subscale were selected, they were sent out for “expert review” among a group of peers consisting of six individuals ranging in experience from early doctoral study in Educational Psychology to Assistant Professor in Educational Psychology, in widely varied areas of the U.S. These expert reviewers were asked to provide feedback on the items, including rephrasing the items to convey what they perceived the item to mean in their own words (Karabenick et al., 2007), reviewing for potential problems in the construction of the items, and grammatical issues. Items were subjected to minor modification upon receipt of expert reviewer feedback for
words that were not entirely clear, or small grammatical suggestions. Upon completion
the items were included in the survey for final administration.

Piloting of the instrument began in July 2011. I collected the first 50 responses online
and then analyzed them to determine how the items were performing and if further
refinement was necessary. Preliminary results from reliability estimates during piloting
revealed the measure was acceptable for use in the final administration (Part 1: \( n = 48 \),
Cronbach’s \( \alpha = .85 \); Part 2: \( n = 48 \); Cronbach’s \( \alpha = .88 \)). A Pearson-product moment
correlation was calculated to determine the strength of the relation between the two
potential sub-scales (part 1 and part 2). These were strongly, positively correlated \( (r = .88, n = 48, p < .000) \).

The final version of the stress inventories included 16 items across two independent
measures (Known Issues; Perceived Stress). A five-point, Likert-type scale was used with
anchors at 1 (Strongly Disagree) to 5 (Strongly Agree). Items with a negative valence
were reverse-coded for analysis. The additional eight items (“This causes me stress”)
reflect post-proposal discussions between advisor and student and were added to capture
stress specifically, rather than more global issues that were happening in schools
exclusively. For example, the item “I feel pressure to improve standardized test scores”
captures testing pressures, issues in schools and classrooms, etc. However, it does not
imply the teacher is experiencing stress, which was the original intent of the study. This,
combined with the previously discussed issue of double-barreled items, resulted in the
final decision to implement a two-part measure to capture with precision the stress that is
experienced by teachers specifically due to high-stakes tests. Items can be found in Table

56
3.2.
Table 3.2.
The Perceived Stress due to High-Stakes Test Scale, Known Issues and Stress subscales.

<table>
<thead>
<tr>
<th>Known Issues</th>
<th>Perceived Stress</th>
</tr>
</thead>
<tbody>
<tr>
<td>In my school, I feel pressure to improve standardized test scores.</td>
<td>This causes me stress.</td>
</tr>
<tr>
<td>Tested content areas are often treated as more important than non-tested content areas.</td>
<td>This causes me stress.</td>
</tr>
<tr>
<td>I feel pressured to make sure that my students pass the test.</td>
<td>This causes me stress.</td>
</tr>
<tr>
<td>There is increased pressure in the weeks leading up to the test.</td>
<td>This causes me stress.</td>
</tr>
<tr>
<td>Using practice standardized tests is common.</td>
<td>This causes me stress.</td>
</tr>
<tr>
<td>There is pressure to maintain or improve our image to the public.</td>
<td>This causes me stress.</td>
</tr>
<tr>
<td>There is a sense of relief when the tests are over.</td>
<td>This reduces my stress.</td>
</tr>
<tr>
<td>Tensions rise as testing time approaches.</td>
<td>This causes me stress.</td>
</tr>
</tbody>
</table>

**Teacher Self-Efficacy Survey**

The TSES (Tschannen-Moran & Hoy, 2001) was originally developed to accurately measure teacher self-efficacy, consisting of a long form (24 items) and a short form (12 items). An interval, 9-point response scale was used for each item, with anchors at 1—nothing, 3—very little, 5—some influence, 7—quite a bit, and 9—a great deal. Example items include “To what extent can you use a variety of assessment strategies?” “How much can you do to control disruptive behavior in the classroom?” “How much can you do to motivate students who show low interest in schoolwork?” The scale consists of three factors, or subscales, designed to measure teacher self-efficacy for instruction, classroom management, and engagement. Reliability estimates on the measure for both the long form and the short form indicate these are reliable for the purposes of measuring teacher self-efficacy (Cronbach’s α = .94 and .90, respectively; Tschannen-Moran & Hoy, 2001).

For the purposes of this study, items were anchored from 1 (I can do nothing) to 5 (I can do a great deal). The short form of the scale was used for this study and each item
was administered two times: the first set of items included the following qualifier—

“Generally, throughout the school year…” The second set of items included this
qualifier—“In the weeks leading up to the test….” I selected this method to parse out the
variability in beliefs teachers hold about their ability to teach in a general classroom
setting and in a classroom setting approaching testing time. The decision to add a second
set of TSES items and differentiate between general, throughout the year beliefs and
leading up to the test beliefs was made post-proposal defense, in collaboration with my
advisor.

These measures were piloted to ensure internal consistency and reliability before full
administration using the same strategy employed with the Perceived Stress due to High-
Stakes Tests items. The piloting on this already-existing and well-validated measure was
conducted due to the unorthodox method in which the scale was being utilized—that is,
asking participants to answer each item two times, representing different times
throughout the school year. The first fifty responses were analyzed and examined, and
reliability estimates were strong, indicating the measures were acceptable for use in the
final study (n = 46, Cronbach’s α = .83 for the “Generally, throughout the school year”
items; n = 46, Cronbach’s α = .89 for the “In the weeks leading up to the test” items).

The full scale and the alphas for each sub-scale can be found in Table 3.3.
Table 3.3.
The Teacher Self-Efficacy Scale, Generally, throughout the school year and In the weeks leading up to the test subscales and Cronbach’s reliability alphas for the pilot sample

<table>
<thead>
<tr>
<th>Generally, throughout the school year… (α = .83)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Instruction</strong></td>
<td></td>
</tr>
<tr>
<td>To what extent can you use a variety of assessment strategies, generally throughout the school year?</td>
<td></td>
</tr>
<tr>
<td>To what extent can you provide an alternative explanation or example when students are confused, generally throughout the school year?</td>
<td></td>
</tr>
<tr>
<td>To what extent can you craft good questions for your students, generally throughout the school year?</td>
<td></td>
</tr>
<tr>
<td>How well can you implement alternative strategies in your classroom, generally throughout the school year?</td>
<td></td>
</tr>
<tr>
<td><strong>Classroom Management</strong></td>
<td></td>
</tr>
<tr>
<td>How much can you do to control disruptive behavior in the classroom, generally throughout the school year?</td>
<td></td>
</tr>
<tr>
<td>How much can you get students to follow classroom rules, generally throughout the school year?</td>
<td></td>
</tr>
<tr>
<td>How much can you do to calm a student who is disruptive or noisy, generally throughout the school year?</td>
<td></td>
</tr>
<tr>
<td>How well can you establish a classroom management system with each group of students, generally throughout the school year?</td>
<td></td>
</tr>
<tr>
<td><strong>Student Engagement</strong></td>
<td></td>
</tr>
<tr>
<td>How much can you do to get students to believe they can do well in schoolwork, generally throughout the school year?</td>
<td></td>
</tr>
<tr>
<td>How much can you do to help your students value learning, generally throughout the school year?</td>
<td></td>
</tr>
<tr>
<td>How much can you do to motivate students who show low interest in schoolwork, generally throughout the school year?</td>
<td></td>
</tr>
<tr>
<td>How much can you do to assist families in helping their children do well in school, generally throughout the school year?</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>In the weeks leading up to the test… (α = .89)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Instruction</strong></td>
<td></td>
</tr>
<tr>
<td>To what extent can you use a variety of assessment strategies, in the weeks leading up to the test?</td>
<td></td>
</tr>
<tr>
<td>To what extent can you provide an alternative explanation or example when students are confused, in the weeks leading up to the test?</td>
<td></td>
</tr>
<tr>
<td>To what extent can you craft good questions for your students, in the weeks leading up to the test?</td>
<td></td>
</tr>
<tr>
<td>How well can you implement alternative strategies in your classroom, in the weeks leading up to the test?</td>
<td></td>
</tr>
<tr>
<td><strong>Classroom Management</strong></td>
<td></td>
</tr>
<tr>
<td>How much can you do to control disruptive behavior in the classroom, in the weeks leading up to the test?</td>
<td></td>
</tr>
<tr>
<td>How much can you get students to follow classroom rules, in the weeks leading up to the test?</td>
<td></td>
</tr>
<tr>
<td>How much can you do to calm a student who is disruptive or noisy, in the weeks leading up to the test?</td>
<td></td>
</tr>
<tr>
<td>How well can you establish a classroom management system with each group of students, in the weeks leading up to the test?</td>
<td></td>
</tr>
<tr>
<td><strong>Student Engagement</strong></td>
<td></td>
</tr>
<tr>
<td>How much can you do to get students to believe they can do well in schoolwork, in the weeks leading up to the test?</td>
<td></td>
</tr>
<tr>
<td>How much can you do to help your students value learning, in the weeks leading up to the test?</td>
<td></td>
</tr>
<tr>
<td>How much can you do to motivate students who show low interest in schoolwork, in the weeks leading up to the test?</td>
<td></td>
</tr>
<tr>
<td>How much can you do to assist families in helping their children do well in school, in the weeks leading up to the test?</td>
<td></td>
</tr>
</tbody>
</table>
Autonomy-Support versus Control

The Teacher Autonomy-support and Control Scale (TACS). In order to measure the tendency of teachers to adopt a controlling orientation toward teaching or an autonomy-supportive orientation toward teaching, I created items from the self-determination perspective. In the spring of 2011, I wrote items to assess the autonomy-supporting or controlling orientation of teachers, based on the literature’s operationalization of these constructs (Assor, Kaplan, & Roth, 2002; Roth, Assor, Kanat-Maymon, & Kaplan, 2007; Reeve, 1998; 2009). Ten items were written and then sent to approximately 10 experienced graduate students or new professors in Educational Psychology around the country for review. Utilizing the feedback I received, I then revised the items for grammar and conceptual clarity. The revisions included minor grammatical corrections and linguistic modifications for clarity, as well as reviewing items that were confounded according to feedback received. For example, the item “I make my presence visible to my students as they work” was interpreted as controlling to some reviewers and as autonomy-supportive to other reviewers. This item was not immediately dropped, because the internal consistency of the scale was stronger with the item included, thus it was kept for the interim. These items were scored on a 5 point Likert-type scale, where (1) = Strongly Disagree, and (5) = Strongly Agree.

Piloting of these items occurred in tandem with the other scales in this study. Initial estimates using Principal Components Analysis with varimax rotation (N = 50) suggested that the items loaded on two components; upon further analysis with the pilot data, seven items clearly represent autonomy-support (n = 50, Cronbach’s α = .75) and five items
clearly represent control \((n = 50, \text{Cronbach’s } \alpha = .62)\). The two subscales are not
strongly correlated \((n = 50, r = .16, p = .28)\) suggesting they are measuring different
latent variables or constructs.

The results from the reliability analyses on the scale designed to measure controlling
orientations in teachers presented an interesting problem. These items were not forming a
strong scale according to traditional estimates of reliability (i.e., Cronbach’s alpha). With
this information in hand, the items were once again modified for conceptual clarity. This
process consisted of rewriting items to improve clarity, and asking Educational
Psychologists, former teachers, and writers to review the items for meaning and clarity.

Items, as well as the full scale and alphas for each factor can be found in Table 3.4.
Table 3.4. 
The Teacher Autonomy-Support and Control Scale, Final Scale

<table>
<thead>
<tr>
<th>Autonomy-Support (α = .75)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>I generally give my students feedback on their work as quickly as possible.</td>
<td></td>
</tr>
<tr>
<td>I always try to understand the students’ perspectives when I am working through difficult problems with them.</td>
<td></td>
</tr>
<tr>
<td>I include my students in the process of designing lessons.</td>
<td></td>
</tr>
<tr>
<td>I am flexible with deadlines for projects.</td>
<td></td>
</tr>
<tr>
<td>I usually give my students a choice when assigning projects.</td>
<td></td>
</tr>
<tr>
<td>I encourage my students to set their own goals and deadlines.</td>
<td></td>
</tr>
<tr>
<td>I include my students in the process of designing rules for my classroom.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Control (α = .62)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>I offer my students rewards for good behavior.</td>
<td></td>
</tr>
<tr>
<td>I prefer to have my students learn to solve problems the way I have taught them.</td>
<td></td>
</tr>
<tr>
<td>I expect my students follow the directions that I explain to them.</td>
<td></td>
</tr>
<tr>
<td>I make it clear to my students that there are negative consequences for failing in school.</td>
<td></td>
</tr>
<tr>
<td>I make my presence visible to my students as they work.</td>
<td></td>
</tr>
</tbody>
</table>

Problems in Schools Questionnaire. Additionally, to provide validity to the items constructed to measure autonomy-support and controlling orientations toward teaching, I chose to administer the Problems in Schools Questionnaire (PIS) (Deci, Schwartz, Sheinman, and Ryan, 1981). This scale is a compilation of eight vignettes describing a potential classroom situation. The scale consists of four subscales that target teachers’ orientations toward teaching—Highly Autonomous, Moderately Autonomous, Moderately Controlling, and Highly Controlling. After each individual vignette, there are four responses, each representing one of four different orientations/subscales toward control and autonomy-support (high control; moderate control; high autonomy; moderate autonomy). The participants rates each of the four possible responses according to how appropriate they think it is, on a scale from 1 to 5, where (1) = Strongly Disagree, and (5) = Strongly Agree (see Table 3.5). For the purposes of this study, all relevant vignettes were included (6) and two were not included. The two vignettes that were not included in
the final survey were eliminated as a result of cognitive interviewing with current and former teachers. Items can be found in Table 3.5.
Table 3.5.
The Problems in Schools Questionnaire, Vignettes and Items

Jim is an average student who has been working at grade level. During the past two weeks he has appeared listless and has not been participating during reading group. The work he does is accurate but he has not been completing assignments. A phone conversation with his mother revealed no useful information. The most appropriate thing for Jim's teacher to do is:

1. She should impress upon him the importance of finishing his assignments since he needs to learn this material for his own good.
2. Let him know that he doesn't have to finish all of his work now and see if she can help him work out the cause of the listlessness.
3. Make him stay after school until that day's assignments are done.
4. Let him see how he compares with the other children in terms of his assignments and encourage him to catch up with the others.

At a parent conference last night, Mr. and Mrs. Greene were told that their daughter Sarah has made more progress than expected since the time of the last conference. All agree that they hope she continues to improve so that she does not have to repeat the grade (which the Greene's have been kind of expecting since the last report card). As a result of the conference, the Greenes decide to:

5. Increase her allowance and promise her a ten-speed if she continues to improve.
6. Tell her that she's now doing as well as many of the other children in her class.
7. Tell her about the report, letting her know that they're aware of her increased independence in school and at home.
8. Continue to emphasize that she has to work hard to get better grades.

Donny loses his temper a lot and has a way of agitating other children. He doesn't respond well to what you tell him to do and you're concerned that he won't learn the social skills he needs. The best thing for you to do with him is:

9. Emphasize how important it is for him to control himself in order to succeed in school and in other situations.
10. Put him in a special class which has the structure and reward contingencies which he needs.
11. Help him see how other children behave in these various situations and praise him for doing the same.
12. Realize that Donny is probably not getting the attention he needs and start being more responsive to him.

The Rangers spelling group has been having trouble all year. How could Miss Wilson best help the Rangers?

13. Have regular spelling bees so that Rangers will be motivated to do as well as the other groups.
14. Make them drill more and give them special privileges for improvements.
15. Have each child keep a spelling chart and emphasize how important it is to have a good chart.
16. Help the group devise ways of learning the words together (skits, games, and so on).

In your class is a girl named Margy who has been the butt of jokes for years. She is quiet and usually alone. In spite of the efforts of previous teachers, Margy has not been accepted by the other children. Your wisdom would guide you to:

17. Prod her into interactions and provide her with much praise for any social initiative.
18. Talk to her and emphasize that she should make friends so she'll be happier.
19. Invite her to talk about her relations with the other kids, and encourage her to take small steps when she's ready.
20. Encourage her to observe how other children relate and to join in with them.

For the past few weeks things have been disappearing from the teacher's desk and lunch money has been
Analyses

**Frequency analyses.** Data in this study were first and foremost subjected to descriptive and frequency analyses.

**Reliability of scales.** When descriptive analyses were complete, the scales that were used and described in this chapter previously were then analyzed for reliability and validity. All of the scales utilized in this study were subjected to Principal Components Analysis with Varimax rotation (PCA) to determine how the items were loading onto the various components, providing additional evidence to their internal consistency; that is, to bolster the reliability of the measures utilized in this study. I then subjected each scale to reliability testing with Cronbach’s alpha used as the primary indicator. Each scale was examined for acceptable levels of variance in the items using Bartlett’s Test of Sphericity and the Kaiser-Meyer-Olkin Test (Meyers & Gamst, 2005). This test is an initial step in principal components analysis that determines if sufficient variability exists in a set of items and if principal components analysis is warranted. The emerging subscales were all tested for reliability with Cronbach’s alpha and the subscales were reported. The exception to this is the *Problems in Schools Questionnaire*, which was not subjected to
PCA due to its well-established validity. The scale was not altered for use in this study; save for the elimination of eight items (two vignettes), so I determined that the scale would not need to be subjected to this analysis. Cronbach’s alphas were reported for this scale.

**Correlations.** Bivariate correlations were then examined in order to determine the strongest relations among the variables, and to provide preliminary evidence to support the hypotheses of the study.

**ANOVAS.** Analysis of variance was conducted on dichotomous and categorical predictor variables to determine if differences existed between groups of participants. These included gender, race, neighborhood, and tested versus non-tested content areas. Statistical significance was reported as well as effect size for statistically significant findings.

**Multiple regression analyses.** Regression analyses were then performed that examined the dependent (outcome) variables when controlling for demographic variability and perceived stress. Four separate multiple regressions were conducted to analyze the four dependent (outcome) variables. The first task to be conducted in this analysis was to dummy code all of the categorical data so they could then be entered into the regressions. The variables that were subjected to dummy coding included gender, race, neighborhood, and content areas taught. The dummy codes for these variables are detailed in Table 3.6.
Table 3.6  
* Dummy Codes for Categorical Variables *

<table>
<thead>
<tr>
<th></th>
<th>Male</th>
<th>Female</th>
<th>White*</th>
<th>Black</th>
<th>Asian</th>
<th>Native Am</th>
<th>Hisp</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>0</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Content Area</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tested*</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-Tested</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Both</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neither</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neighborhood</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urban</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Suburban*</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rural</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Comparison group denoted in bold.

Following the dummy coding of categorical variables, the regression equations were modeled. The first step in the multiple regression model was to enter the demographic predictor variables, including years of teaching experience, school high-stakes test scores, non/tested content area, grades taught, highest education attained, teacher salary, school neighborhood, free/reduced lunch, class size, and administrative support. These are the control variables for each of the four equations. State and political persuasion were not included in the analyses. The first step of the model is as follows:

Step 1: \[ \text{DV} = b_0 + b_1(\text{years experience}) + b_2(\text{HST score}) + b_3(\text{non-tested content area}) + b_4(\text{grades taught}) + b_5(\text{highest education}) + b_6(\text{income}) + b_7(\text{neighborhood}) + b_8(\text{free and reduced}) + b_9(\text{class size}) + b_{10}(\text{admin}) + b_{11}(\text{gender}) + b_{12}(\text{race}). \]

The two stress subscale variables were entered into the regression model at Step 2. First, the Known Issues variable was entered, followed by the Perceived Stress variable.
The equation then became:

\[
\text{Step 2: } \text{DV} = \beta_0 + b_1(\text{years experience}) + b_2(\text{HST score}) + b_3(\text{non/tested content area}) + b_4(\text{grades taught}) + b_5(\text{highest education}) + b_6(\text{income}) + b_7(\text{neighborhood}) + b_8(\text{free and reduced}) + b_9(\text{class size}) + b_{10}(\text{admin}) + b_{11}(\text{gender}) + b_{12}(\text{race}) + b_{13}(\text{PS-HST, Known Issues}) + b_{14}(\text{PS-HST, Stress}).
\]

And following the PS-HST Known Issues subscale, the PS-HST Stress subscale was entered. These steps were included to provide evidence to indicate if teacher perceptions of known issues due to high-stakes tests and the resultant stress explained additional variance in the outcome variables, after having controlled for demographics.

\[
\text{DV} = \beta_0 + b_1(\text{years experience}) + b_2(\text{HST score}) + b_3(\text{non/tested content area}) + b_4(\text{grades taught}) + b_5(\text{highest education}) + b_6(\text{income}) + b_7(\text{neighborhood}) + b_8(\text{free and reduced}) + b_9(\text{class size}) + b_{10}(\text{admin}) + b_{11}(\text{gender}) + b_{12}(\text{race}) + b_{13}(\text{PS-HST, Known Issues}) + b_{14}(\text{PS-HST, Stress}).
\]

Interactions that were entered and examined in the regression equations included “How much of your evaluation as a teacher depends on your students’ high-stakes test scores?” multiplied by the Perceived Stress scale. The rationale for this interaction is that teachers that are evaluated to a greater extent based upon their students’ test scores may be experiencing greater amounts of stress than teachers whose evaluation does not depend on test scores. A new variable was created to represent this interaction by multiplying the standardized item measuring teacher evaluations and the standardized version of the Perceived Stress scale. The variable was then entered into the equation as a predictor in Step 3.

The second interaction to be examined included the variables for perceived administrative support and the tendency of the school to pass the high-stakes test.
Administrative support mediates teacher stress when administrators are perceived as supportive (Margolis & Nagel, 2006), and teacher stress is increased when teachers are pressured to improve scores with no outlet for their frustrations (Margolis & Nagel, 2006). The interaction between perceived administrative support and perceived stress as a result of high-stakes tests was examined in order to determine if there was a moderation effect in addition to the mediation effect already documented by Margolis et al. (2006). A new variable was created to represent this interaction by multiplying the standardized item measuring perceived administrative support and the standardized Perceived Stress scale. This variable was then entered into the equation as a predictor in Step 4.

Finally, the third interaction included teachers’ perceived stress as a result of high-stakes tests multiplied by extent to which teachers are evaluated based upon student test scores, creating an interaction variable that was entered into Step 4 of the equation. The purpose for examining teachers’ self-report regarding how much of their teacher evaluation depended upon their students’ high-stakes test scores with self-reports about the extent to which their school passes the test was based upon the same arguments as the previous interactions: it is thought that teachers whose evaluations depend on student scores may be experiencing higher stress, and for those teachers who also teach in schools that are less likely to pass the test, the stress may be even higher. This interaction was examined accordingly.

The final regression models were as follows:

\[ DV = \beta_0 + b_1(\text{years experience}) + b_2(\text{HST score}) + b_3(\text{non/tested content area}) + b_4(\text{grades taught}) + b_5(\text{highest education}) + b_6(\text{income}) + b_7(\text{neighborhood}) + b_8(\text{free and reduced}) + b_9(\text{class size}) + b_{10}(\text{admin}) + b_{11}(\text{gender}) + \]

70
The separate models are as follows:

**Teacher Self-Efficacy: Generally, throughout the school year**
\[ \beta_0 + b_1(\text{years experience}) + b_2(\text{HST score}) + b_3(\text{non/tested content area}) + b_4(\text{grades taught}) + b_5(\text{highest education}) + b_6(\text{income}) + b_7(\text{neighborhood}) + b_8(\text{free and reduced}) + b_9(\text{class size}) + b_{10}(\text{admin}) + b_{11}(\text{gender}) + b_{12}(\text{race}) + b_{13}(\text{Teacher Eval} \times \text{Stress}) + b_{14}(\text{Admin Support} \times \text{HST Score}) + b_{15}(\text{Teacher Eval} \times \text{HST Score}) + b_{16}(\text{PS-HST, Known Issues}) + b_{17}(\text{PS-HST, Stress}). \]

**Teacher Self-Efficacy: In the weeks leading up to the test**
\[ \beta_0 + b_1(\text{years experience}) + b_2(\text{HST score}) + b_3(\text{non/tested content area}) + b_4(\text{grades taught}) + b_5(\text{highest education}) + b_6(\text{income}) + b_7(\text{neighborhood}) + b_8(\text{free and reduced}) + b_9(\text{class size}) + b_{10}(\text{admin}) + b_{11}(\text{gender}) + b_{12}(\text{race}) + b_{13}(\text{Teacher Eval} \times \text{Stress}) + b_{14}(\text{Admin Support} \times \text{HST Score}) + b_{15}(\text{Teacher Eval} \times \text{HST Score}) + b_{16}(\text{PS-HST, Known Issues}) + b_{17}(\text{PS-HST, Stress}). \]

**Teacher Autonomy-Support/Control**
\[ \beta_0 + b_1(\text{years experience}) + b_2(\text{HST score}) + b_3(\text{non/tested content area}) + b_4(\text{grades taught}) + b_5(\text{highest education}) + b_6(\text{income}) + b_7(\text{neighborhood}) + b_8(\text{free and reduced}) + b_9(\text{class size}) + b_{10}(\text{admin}) + b_{11}(\text{gender}) + b_{12}(\text{race}) + b_{13}(\text{Teacher Eval} \times \text{Stress}) + b_{14}(\text{Admin Support} \times \text{HST Score}) + b_{15}(\text{Teacher Eval} \times \text{HST Score}) + b_{16}(\text{PS-HST, Known Issues}) + b_{17}(\text{PS-HST, Stress}). \]
Chapter Four: Results

The purpose of the present study was to examine the relations of teacher motivational beliefs to perceived stress due to high-stakes tests. Specifically, the study focused on the following research questions:

- What is the relation between teachers’ perceived stress as a result of high-stakes tests and teacher self-efficacy?
- What is the relation between teachers’ perceived stress as a result of high-stakes tests and their autonomy-supportive v. controlling orientation toward teaching?

Quantitative Analyses

The first set of tasks in the analyses included validating the measures that were utilized for the study. There were four separate measures incorporated into this study: the Perceived Stress Due To High-Stakes Test Scale (PS-HST; Dawson, 2011), the Teacher Self-Efficacy Scale (TSES; Tschannen-Moran & Woolfolk Hoy, 2001), the Teacher Autonomy and Control Scale (TACS; Dawson, 2011) and the Problems in Schools Questionnaire (PIS; Deci, Schwartz, Sheinman, & Ryan, 1981). The PS-HST scale consisted of two potential subscales: the items designed to measure the extent to which HST is causing issues in the classroom and the extent to which those issues are causing stress. The TSES was administered twice, once with a qualifier “Generally, throughout the school year...” and once with the qualifier “In the weeks leading up to the test...”
Therefore, in total, six scales were validated. There were several possible rationales for conducting validation exercises with these scales. These scales, even if previously validated for use by other scholars, were being utilized in new and different ways for this study; or, these scales had not been used in some time or with results consistent enough to overlook validation.

**The Perceived Stress Due to High-Stakes Tests (PS-HST); Known Issues Items**

As described in Chapter Three, the *Perceived Stress Due to High-Stakes Tests* measure was constructed in two parts: the first to measure the extent to which teachers perceive actual changes in their day-to-day teaching lives as a result of high-stakes tests; and the second to measure the extent to which teachers perceive these changes are causing stress. The first set of items will be referred to as the *Known Issues* items, and the second set of items will be referred to as the *Stress* items.

The *Known Issues* items were subjected to reliability analyses. Cronbach’s alpha indicated this scale was internally consistent ($\alpha = .79$). In order to determine if the correlation matrices consisted of significant relations among the items, thus warranting further analysis, it is important to test the matrices. This is done with two primary tests conducted in SPSS: *Bartlett’s Test of Sphericity* and the *Kaiser-Meyer-Olkin Test*.

Bartlett’s test examines the extent to which there are relations among the items on a potential scale—without this, there could potentially be as many components as items, and the scale would not be useful to measure the underlying, or latent, construct (Meyers & Gamst, 2005). The Kaiser-Meyer-Olkin test can then be used to determine the strength of the relations between the items using the partial correlation coefficient (Meyers & Gamst, 2005). Bartlett’s test for this scale was significant ($\chi^2 = 1,196.99$, $df = 28$, $p <$
indicating the null hypothesis for this test was void and there are significant relations among the variables warranting further analysis. In order to compare the correlation coefficients to partial correlation coefficients, KMO test for sampling adequacy was utilized and indicated further analysis was justified (.88).

Principal components analysis (PCA) with varimax rotation was then conducted on the items to determine if the items held together as a single underlying construct or if the items seemed to be measuring more than one component. This decision was made in order to account for all of the variance in the items. The PCA extracted two components from the eight items with an eigenvalue above 1.0 (3.32 and 1.22 respectively), accounting for 56.77% of the variance. A scree test indicated one or two components could be extracted. Upon examination of the items, it became clear that the first component consisted of items of a comparative nature, i.e. “There is pressure to improve or maintain our image to the public.” The second set of items indicated beliefs about perceived stress due to tests from the perspective of the days and weeks before the test is administered, i.e. “There is increased pressure in the weeks leading up to the test.” Two subscales were formed from the original eight items reflecting this development, the first labeled Comparative and the second labeled Pre-test. Four items loaded on to the Comparative subscale. These items, their means, standard deviations, and their loadings can be found in Table 4.1. A reliability analysis indicated the Comparative subscale had acceptable internal consistency (α = .69). Four items loaded on to the Pre-Test subscale with a higher internal consistency (α = .75). A Pearson-product moment correlation was conducted to examine the linear relationship (dependence) between the two subscales, revealing a moderately positive correlation (r = .48, p < .001).
Table 4.1
Means, Standard Deviations, and Rotated Component Loadings for the
Perceived Stress due to High-Stakes Tests Scale, Known Issues Items*

<table>
<thead>
<tr>
<th>Item</th>
<th>M</th>
<th>SD</th>
<th>Loading</th>
</tr>
</thead>
<tbody>
<tr>
<td>In my school…</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Comparative (α = .69)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I feel pressured to make sure that my students pass the test.</td>
<td>4.21</td>
<td>.915</td>
<td>.754</td>
</tr>
<tr>
<td>I feel pressure to improve standardized test scores.</td>
<td>4.39</td>
<td>.825</td>
<td>.729</td>
</tr>
<tr>
<td>There is pressure to maintain or improve our image to the public.</td>
<td>4.53</td>
<td>.651</td>
<td>.679</td>
</tr>
<tr>
<td>Tested content areas are often treated as more important than non-tested content areas.</td>
<td>4.37</td>
<td>.850</td>
<td>.605</td>
</tr>
<tr>
<td><strong>Pre-Test (α = .75)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tensions rise as testing time approaches.</td>
<td>4.22</td>
<td>.830</td>
<td>.807</td>
</tr>
<tr>
<td>There is a sense of relief when the tests are over.</td>
<td>4.38</td>
<td>.871</td>
<td>.801</td>
</tr>
<tr>
<td>There is increased pressure in the weeks leading up to the test.</td>
<td>4.29</td>
<td>.877</td>
<td>.720</td>
</tr>
<tr>
<td>Using practice standardized tests is common.</td>
<td>4.16</td>
<td>.959</td>
<td>.571</td>
</tr>
</tbody>
</table>

*Participants were asked to respond to these items on a 5-point Likert type scale, where (1) = Strongly Disagree, (3) = Neutral, and (5) = Strongly Agree.

The Perceived Stress Due to High-Stakes Tests (PS-HST); Stress Items

Initial reliability estimates using Cronbach’s alpha suggest this scale is internally consistent (α = .89). Bartlett’s test of sphericity and a KMO analysis suggest there is significant variability in the responses to warrant further analysis of the component structure ($\chi^2 = 2,561.68, df = 28, p < .001; KMO = .895$). A Principal Components Analysis revealed one component accounting for 60.5% of the variance with an Eigenvalue of 4.84. A scree test suggested one component could be extracted.

Each of the items on this scale was linguistically the same (“This causes me stress”). Conceptually, these items were administered in response to each single item on the Known Issues Scale (see Table 4.1). Therefore, teachers were asked to respond regarding the extent to which they perceived an event occurring as a result of high-stakes tests (described above), and then asked the extent to which that event caused them stress. The results of the PCA suggest the items are tapping into a single construct (perceived...
stress) as a result of the known issues due to high-stakes tests. The actual instrument administered to the sample can be seen in Appendix 4.1. The means, standard deviations, and loadings for the Stress items can be found in Table 4.2.

Table 4.2
Means, Standard Deviations, and Component Loadings for the Perceived Stress due to High-Stakes Tests, Stress Items.

<table>
<thead>
<tr>
<th>Item</th>
<th>M</th>
<th>SD</th>
<th>Loading</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>This causes me stress.</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I feel pressured to make sure that my students pass the test.*</td>
<td>3.93</td>
<td>1.01</td>
<td>.863</td>
</tr>
<tr>
<td>I feel pressure to improve standardized test scores.*</td>
<td>3.94</td>
<td>1.02</td>
<td>.826</td>
</tr>
<tr>
<td>There is pressure to maintain or improve our image to the public.*</td>
<td>3.77</td>
<td>1.00</td>
<td>.721</td>
</tr>
<tr>
<td>Tested content areas are often treated as more important than non-tested content areas.*</td>
<td>3.80</td>
<td>1.04</td>
<td>.799</td>
</tr>
<tr>
<td>Tensions rise as testing time approaches.*</td>
<td>3.85</td>
<td>.981</td>
<td>.866</td>
</tr>
<tr>
<td>There is a sense of relief when the tests are over.*</td>
<td>3.78</td>
<td>1.07</td>
<td>.398</td>
</tr>
<tr>
<td>There is increased pressure in the weeks leading up to the test.*</td>
<td>3.93</td>
<td>1.01</td>
<td>.877</td>
</tr>
<tr>
<td>Using practice standardized tests is common.*</td>
<td>3.34</td>
<td>1.02</td>
<td>.760</td>
</tr>
</tbody>
</table>

*The items from the PS-HST Known Issues scale are listed here for reference. The item means, standard deviations, and loadings in this table refer to the item “This causes me stress” that was asked of each of the Known Issues items, as described in Chapter Three. All of these items were rated on a 5 point Likert-type scale, where (1) = Strongly Disagree, (3) = Neutral, and (5) = Strongly Agree.

The Teacher Self-Efficacy Scale (TSES)

The Teacher Self-Efficacy Scale (TSES) was administered to measure both beliefs throughout the school year and in the weeks leading up to the test. Each item was administered twice; the first item with the following qualifier, “Generally, throughout the school year,” and the second item, “In the weeks leading up to the test...” The TSES is a well-established and highly validated measure (Klassen, Tze, Betts, & Gordon, 2011; Tschannen-Moran & Woolfolk Hoy, 2001). However, due to the unorthodox manner in which this scale was used, the decision was made to validate the psychometric properties of the scale for use in the current study. The two scales are theoretically and conceptually
measuring the same psychological construct; however, the measures did require study participants to think about this construct in two distinct manners—therefore, conducting these analyses provides additional evidence that the measures were sound and valid, and can be used for further analyses as indicated further in this chapter.

Beginning with a reliability analysis, the TSES *Generally, throughout the school year* items indicated the scale was internally consistent (α = .89). In order to determine what these items were tapping into, a Principal Components Analysis with varimax rotation was conducted with both versions of the scale (*Generally, throughout the school year* and *In the weeks leading up to the test*). For the items qualified with the “*Generally, throughout the school year*” prompt, the Bartlett’s test of sphericity indicated there was significant variance among the responses ($\chi^2 = 3,414.9, df = 66, p < .001$). The Kaiser-Meyer-Olkin test indicated further analysis was warranted (.881). The items formed three subscales as indicated in the original validation exercise (Tschannen-Moran & Woolfolk Hoy, 2001), accounting for 72.9% of the variance in responses, all with Eigenvalues over one. The first component extracted yielded an Eigenvalue of 5.46, the second component yielded an Eigenvalue of 1.96, and the third 1.35. These are instruction, management, and engagement. The item means, standard deviations, and loadings for the TSES, *Generally, throughout the school year* can be found in Table 4.3.

For the items qualified with the “*In the weeks leading up to the test*” prompt, the KMO and Bartlett’s tests indicated further analysis was warranted ($\chi^2 = 3,693.7, df = 66, p < .01; \text{KMO} = .891$). A principal components analysis with varimax rotation indicated three components could be extracted, accounting for 74.77% of the variance, with Eigenvalues above one. They were 5.7 for the first component, 1.97 for the second
component, and 1.31 for the third component. These items loaded on to conceptually similar components as in the Generally, throughout the school year scale. Item means, standard deviations, and component loadings for this scale can be found in Table 4.4.

Table 4.3
Means, Standard Deviations, and Loadings for the Teacher Self-Efficacy Scale, Generally, Throughout the School Year

<table>
<thead>
<tr>
<th>Instruction</th>
<th>M</th>
<th>SD</th>
<th>Loading</th>
</tr>
</thead>
<tbody>
<tr>
<td>How much can you do to get children to follow classroom rules?</td>
<td>4.36</td>
<td>.744</td>
<td>.892</td>
</tr>
<tr>
<td>To what extent can you craft good questions for your students?</td>
<td>4.23</td>
<td>.826</td>
<td>.875</td>
</tr>
<tr>
<td>How much can you do to get students to believe they can do well in schoolwork?</td>
<td>4.15</td>
<td>.827</td>
<td>.830</td>
</tr>
<tr>
<td>How well can you establish a classroom management system with each group of students?</td>
<td>4.44</td>
<td>.710</td>
<td>.812</td>
</tr>
<tr>
<td><strong>Instruction</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>How much can you do to motivate students who show low interest in schoolwork?</td>
<td>4.10</td>
<td>.891</td>
<td>.843</td>
</tr>
<tr>
<td>How much can you do to calm a student who is disruptive or noisy?</td>
<td>4.36</td>
<td>.792</td>
<td>.804</td>
</tr>
<tr>
<td>How much can you do to control disruptive behavior in the classroom?</td>
<td>4.00</td>
<td>.907</td>
<td>.797</td>
</tr>
<tr>
<td>How much can you do to help your students value learning?</td>
<td>4.16</td>
<td>.834</td>
<td>.792</td>
</tr>
<tr>
<td><strong>Engagement</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>To what extent can you use a variety of assessment strategies?</td>
<td>3.60</td>
<td>.898</td>
<td>.784</td>
</tr>
<tr>
<td>How well can you implement alternative teaching strategies in your classroom?</td>
<td>3.82</td>
<td>.875</td>
<td>.779</td>
</tr>
<tr>
<td>How much can you assist families in helping their children do well in school?</td>
<td>3.28</td>
<td>.951</td>
<td>.770</td>
</tr>
<tr>
<td>To what extent can you provide an alternative explanation or example when students are confused?</td>
<td>3.98</td>
<td>.823</td>
<td>.770</td>
</tr>
</tbody>
</table>

All items were measured on a 5-pt Likert-type scale, where (1) = I can do nothing, and (5) = I can do much.
Table 4.4
Means, Standard Deviations, and Loadings for the Teacher Self-Efficacy Scale,
*In the weeks leading up to the test...*

<table>
<thead>
<tr>
<th>Instruction</th>
<th>M</th>
<th>SD</th>
<th>Loading</th>
</tr>
</thead>
<tbody>
<tr>
<td>To what extent can you craft good questions for your students?</td>
<td>4.05</td>
<td>.93</td>
<td>.887</td>
</tr>
<tr>
<td>How much can you do to get children to follow classroom rules?</td>
<td>4.18</td>
<td>.85</td>
<td>.886</td>
</tr>
<tr>
<td>How much can you do to get students to believe they can do well in schoolwork?</td>
<td>3.97</td>
<td>.91</td>
<td>.824</td>
</tr>
<tr>
<td>How well can you establish a classroom management system with each group of students?</td>
<td>4.21</td>
<td>.85</td>
<td>.822</td>
</tr>
</tbody>
</table>

**Instruction**

<table>
<thead>
<tr>
<th>Management</th>
<th>M</th>
<th>SD</th>
<th>Loading</th>
</tr>
</thead>
<tbody>
<tr>
<td>How much can you do to motivate students who show low interest in schoolwork?</td>
<td>3.81</td>
<td>1.03</td>
<td>.852</td>
</tr>
<tr>
<td>How much can you do to control disruptive behavior in the classroom?</td>
<td>3.61</td>
<td>1.05</td>
<td>.849</td>
</tr>
<tr>
<td>How much can you do to help your students value learning?</td>
<td>3.73</td>
<td>1.04</td>
<td>.825</td>
</tr>
<tr>
<td>How much can you do to calm a student who is disruptive or noisy?</td>
<td>4.00</td>
<td>1.01</td>
<td>.801</td>
</tr>
</tbody>
</table>

**Engagement**

<table>
<thead>
<tr>
<th>Engagement</th>
<th>M</th>
<th>SD</th>
<th>Loading</th>
</tr>
</thead>
<tbody>
<tr>
<td>To what extent can you use a variety of assessment strategies?</td>
<td>3.25</td>
<td>.98</td>
<td>.805</td>
</tr>
<tr>
<td>How well can you implement alternative teaching strategies in your classroom?</td>
<td>3.46</td>
<td>.96</td>
<td>.779</td>
</tr>
<tr>
<td>To what extent can you provide an alternative explanation or example when students are confused?</td>
<td>3.61</td>
<td>.94</td>
<td>.759</td>
</tr>
<tr>
<td>How much can you assist families in helping their children do well in school?</td>
<td>3.07</td>
<td>.94</td>
<td>.736</td>
</tr>
</tbody>
</table>

All items were measured on a 5-pt Likert-type scale, where (1) = I can do nothing, and (5) = I can do much.

**The Teacher Autonomy-support and Control Scale (TACS)**

To determine teachers’ tendencies to adopt a controlling orientation toward teaching versus an autonomy-supportive orientation toward teaching, items were developed based upon the literature’s operationalization of these constructs. The analyses began by separating the items designed to measure autonomy-support from those measuring control. To begin, Cronbach’s alpha with the autonomy-support items indicated one item should be removed from the potential scale (*I generally give my students feedback on their work as quickly as possible*), leaving six items remaining. The remaining six items resulted in a scale with marginally acceptable internal consistency.
based upon the reliability analyses using Cronbach’s alpha (α = .68). To determine the validity of these items and demonstrate the internal consistency of the scale, PCA with varimax rotation was conducted. The KMO test of sampling adequacy and Bartlett’s test of sphericity indicated there was significant variance in the items, warranting further analysis ($\chi^2 = 386.89, df = 15, p < .001; \text{KMO} = .775$). Principal components analysis with varimax rotation was conducted with the six items to determine how many components the scale was tapping into. The results yielded one component accounting for 38.8% of the variance in responses (Eigenvalue 2.39), suggesting the items forming this subscale were in fact measuring autonomy-support. The subscale that resulted, including items, means, standard deviations, and component loadings can be found in Table 4.5.

Table 4.5

**Teacher Autonomy-support and Control Scale, Autonomy-Support Items**

All items were measured on a 5-pt, Likert-type rating scale where (1) = Strongly Disagree and (5) = Strongly Agree.

<table>
<thead>
<tr>
<th>Item</th>
<th>M</th>
<th>SD</th>
<th>Loading</th>
</tr>
</thead>
<tbody>
<tr>
<td>I encourage my students to set their own goals and deadlines.</td>
<td>3.06</td>
<td>1.08</td>
<td>.687</td>
</tr>
<tr>
<td>I include my students in the process of designing rules for my classroom.</td>
<td>3.40</td>
<td>1.19</td>
<td>.676</td>
</tr>
<tr>
<td>I include my students in the process of designing lessons.</td>
<td>3.26</td>
<td>1.00</td>
<td>.671</td>
</tr>
<tr>
<td>I usually give my students a choice when designing projects.</td>
<td>3.57</td>
<td>.945</td>
<td>.671</td>
</tr>
<tr>
<td>I am flexible with deadlines for projects.</td>
<td>3.33</td>
<td>1.05</td>
<td>.538</td>
</tr>
<tr>
<td>I always try to understand the students’ perspective when I am working through a difficult problem with a student.</td>
<td>4.40</td>
<td>.596</td>
<td>.458</td>
</tr>
</tbody>
</table>

The control items were subjected to the same analyses. Cronbach’s alpha indicated two items should be removed in order to facilitate the strongest scale, leaving three items. The results yielded a subscale with poor internal consistency (α = .446). Items were reviewed for valence, coding, and other potential complications that could be contributing to these poor results. The performance of these items did not warrant further
analysis. The poor validity and reliability of this measure indicated it was not acceptable for use.

**The Problems in Schools Questionnaire (PIS)**

In order to provide additional validity to these items, the Problems in Schools Questionnaire was administered. This scale was designed to measure the autonomy-supportive and controlling orientations of teachers (Deci, Sheinman, Ryan, & Schwartz, 1981). Due to the age of the scale and the moderate results as published by the authors of the original scale, basic validity exercises were warranted on this already-established scale.

Reliability analyses indicated the scale, as a whole, was internally consistent ($\alpha = .80$). The scale consists of four subscales: Highly Controlling, Moderately Controlling, Moderately Autonomous, and Highly Autonomous. The four subscales did not perform as well as the full scale, with reliability estimates ranging from $\alpha = .61$ to $\alpha = .66$ (see Table 4.7). All items, means, and standard deviations for the *Problems in Schools Questionnaire* can be found in Table 4.7.

Mean scores were calculated for each of the four subscales. The subscales were then weighted according to the original authors’ direction as follows: Highly Controlling * -2; Moderately Controlling *-1; Moderately Autonomous *1; and Highly Autonomous *2. The four subscales were then added together to form one composite score. The resulting weights yielded four scores that ranged from -18 to +18, with the higher scores indicating a more autonomous orientation toward teaching, and a lower score yielding a more controlling orientation toward teaching. The final score for this scale was then standardized for inclusion in the final analyses.
Table 4.6  
*Problems in Schools Questionnaire and subscales*

<table>
<thead>
<tr>
<th>Items</th>
<th>$M$</th>
<th>$SD$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Highly Controlling Subscale ($\alpha = .66$)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Make him stay after school until that day’s assignments are done.</td>
<td>2.4</td>
<td>.64</td>
</tr>
<tr>
<td>Increase her allowance and promise her a ten-speed if she continues to improve.</td>
<td>2.26</td>
<td>1.05</td>
</tr>
<tr>
<td>Put him in a special class which has the structure and reward contingencies which he needs.</td>
<td>2.07</td>
<td>1.02</td>
</tr>
<tr>
<td>Make them drill more and give them special privileges for improvements.</td>
<td>2.66</td>
<td>1.14</td>
</tr>
<tr>
<td>Prod her into interactions and provide her with much praise for any social initiative.</td>
<td>2.65</td>
<td>1.04</td>
</tr>
<tr>
<td>Give him a good scolding; stealing is something which cannot be tolerated and he has to learn that.</td>
<td>2.84</td>
<td>1.16</td>
</tr>
<tr>
<td>Offer a dollar for every A and 50 cents for every B on future report cards.</td>
<td>2.69</td>
<td>1.18</td>
</tr>
<tr>
<td><strong>Moderately Controlling Subscale ($\alpha = .66$)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>She should impress upon him the importance of finishing his assignments since he needs to learn this material for his own good.</td>
<td>3.12</td>
<td>.64</td>
</tr>
<tr>
<td>Continue to emphasize that she has to work hard to get better grades.</td>
<td>3.5</td>
<td>1.11</td>
</tr>
<tr>
<td>Emphasize how important it is for him to control himself in order to succeed in school and in other situations.</td>
<td>3.91</td>
<td>1.11</td>
</tr>
<tr>
<td>Have each child keep a spelling chart and emphasize how important it is to have a good chart.</td>
<td>3.60</td>
<td>1.15</td>
</tr>
<tr>
<td>Talk to her and emphasize that she should make friends so she’ll be happier.</td>
<td>3.14</td>
<td>1.15</td>
</tr>
<tr>
<td>Emphasize that it was wrong and have him apologize to the teacher and promise not to do it again.</td>
<td>1.87</td>
<td>.90</td>
</tr>
<tr>
<td>Stress that she should do better; she’ll never get into college with grades like these.</td>
<td>3.83</td>
<td>1.11</td>
</tr>
<tr>
<td><strong>Moderately Autonomous Subscale ($\alpha = .61$)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Let him see how he compares with other children in terms of his assignments and encourage him to catch up with the others.</td>
<td>3.02</td>
<td>.64</td>
</tr>
<tr>
<td>Tell her that she's now doing as well as many of the other children in her class.</td>
<td>2.10</td>
<td>1.15</td>
</tr>
<tr>
<td>Help him see how other children behave in these various situations and praise him for doing the same.</td>
<td>2.39</td>
<td>1.23</td>
</tr>
<tr>
<td>Have regular spelling bees so that Rangers will be motivated to do as well as the other groups.</td>
<td>3.87</td>
<td>1.01</td>
</tr>
<tr>
<td>Encourage her to observe how other children relate and to join in with them.</td>
<td>2.62</td>
<td>1.13</td>
</tr>
<tr>
<td>Talk to him about the consequences of stealing and what it would mean in relation to the other kids.</td>
<td>3.41</td>
<td>.98</td>
</tr>
<tr>
<td>Go over the report card with her; point out where she stands in the class.</td>
<td>4.06</td>
<td>1.06</td>
</tr>
<tr>
<td><strong>Highly Autonomous Subscale ($\alpha = .65$)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Let him know that he doesn't have to finish all of his work now and see if she can help him work out the cause of the listlessness.</td>
<td>4.01</td>
<td>.56</td>
</tr>
<tr>
<td>Tell her about the report, letting her know that they’re aware of her increased independence in school and at home.</td>
<td>3.2</td>
<td>1.25</td>
</tr>
<tr>
<td>Realize that Donny is probably not getting the attention he needs and start being more responsive to him.</td>
<td>4.56</td>
<td>.81</td>
</tr>
<tr>
<td>Help the group devise ways of learning the words together (skits, games).</td>
<td>3.37</td>
<td>1.01</td>
</tr>
<tr>
<td>Invite her to talk about her relations with the other kids, and encourage her to take small steps when she’s ready.</td>
<td>4.66</td>
<td>.721</td>
</tr>
<tr>
<td>Talk to him about it, expressing her confidence in him and attempting to understand why he did it.</td>
<td>4.24</td>
<td>.96</td>
</tr>
<tr>
<td>Encourage her to talk about her report card and what it means for her.</td>
<td>3.91</td>
<td>1.05</td>
</tr>
</tbody>
</table>

*All items were measured on a 5-pt Likert-type scale where (1) = Strongly Disagree, and (1) = Strongly Agree.*
Summary of Scales

In summary, the final scales utilized for the inferential analyses included the *Perceived Stress Due to High-Stakes Tests Scale*, including the *Known Issues* and *Stress* subscales; the *Teacher Self-Efficacy Scale*, including the *Generally, throughout the school year* and the *In the weeks leading up to the test* items; and the *Problems in Schools Questionnaire*. The Teachers Autonomy/Controlling Orientation Scale, autonomy-support items were included for final analyses, however the TACS controllingness items were not chosen for use with the inferential analyses due to their poor performance psychometrically. These scales and their reliability coefficients can be found in Table 4.7.

Table 4.7
*Means, Standard Deviations, and Cronbach’s α for scales included in final analyses*

<table>
<thead>
<tr>
<th>Measure</th>
<th>M</th>
<th>SD</th>
<th>α</th>
</tr>
</thead>
<tbody>
<tr>
<td>PS-HST, Known Issues</td>
<td>4.32</td>
<td>.542</td>
<td>.79</td>
</tr>
<tr>
<td>PS-HST, Stress</td>
<td>3.79</td>
<td>.777</td>
<td>.89</td>
</tr>
<tr>
<td>TSES, <em>Generally...</em> (Full)</td>
<td>4.04</td>
<td>.564</td>
<td>.89</td>
</tr>
<tr>
<td>TSES-G (Instruction)</td>
<td>4.16</td>
<td>.717</td>
<td>.86</td>
</tr>
<tr>
<td>TSES-G (Management)</td>
<td>4.29</td>
<td>.701</td>
<td>.92</td>
</tr>
<tr>
<td>TSES-G (Engagement)</td>
<td>3.67</td>
<td>.725</td>
<td>.84</td>
</tr>
<tr>
<td>TSES-In the weeks... (Full)</td>
<td>3.74</td>
<td>.663</td>
<td>.90</td>
</tr>
<tr>
<td>TSES-I (Instruction)</td>
<td>3.79</td>
<td>.899</td>
<td>.89</td>
</tr>
<tr>
<td>TSES-I (Management)</td>
<td>4.10</td>
<td>.800</td>
<td>.92</td>
</tr>
<tr>
<td>TSES-I (Engagement)</td>
<td>3.35</td>
<td>.788</td>
<td>.84</td>
</tr>
<tr>
<td>TACS (Autonomy Support)</td>
<td>3.50</td>
<td>.619</td>
<td>.68</td>
</tr>
<tr>
<td>PIS (Full)</td>
<td>2.84</td>
<td>.518</td>
<td>.79</td>
</tr>
</tbody>
</table>
Preliminary Analyses

**Relations between demographics.** Bivariate correlations were calculated in order to examine the relations between the variables. Pearson-product moment correlations were utilized for this purpose, and can be found in Table 4.8. Significant correlations existed between several of the demographic variables; these results are outlined in Table 4.8. Many of these correlations were theoretically sound—for example, teachers’ perceptions of administrative support were positively, and significantly, correlated with perceptions of the school’s likelihood of passing the test ($r = .19, p = .01$). The strength of the relation was not strong, but it was still significant, suggesting that teachers who teach in schools that are more likely to pass the test are also more likely to perceive the administration in their school as supportive of their teaching. Teachers’ perceptions of administrative support were also significantly negatively related to the frequency of free/reduced lunch status in the school in which they taught ($r = -.14, p = .01$). This suggests the more students that receive free/reduced lunch in a teacher’s school, the less likely they are to perceive administrative support, and vice versa. This relation is also small, but significant.

Free/reduced lunch status was negatively related to the school’s likelihood of passing the test ($r = -.5, p = .01$). This result suggests that schools that are less likely to pass the test are more likely to have higher numbers of students receiving free/reduced lunch, according to teacher self-reports. Free/reduced lunch was also modestly negatively related to teacher salary ($r = -.18, p = .01$). This suggests teachers make less money in schools that have greater numbers of students receiving free/reduced lunch. The correlation between free/reduced lunch status and teaching experience was not
statistically significant ($r = -.08$). Finally, free/reduced lunch status was positively, significantly correlated to a small degree with teacher evaluations (How much of your evaluation as a teacher depends on your students’ test scores?), suggesting that administrators in schools with higher levels of free/reduced lunch status hold their teachers accountable for their students’ test scores to a greater degree than schools that do not, and vice versa ($r = .19, p = .01$). These results can be found in Table 4.8.

Teacher salary was positively related to the likelihood that the school will pass the test ($r = .15, p = .01$), and to teacher experience and teacher education level (as mentioned above). Salary was negatively related to free/reduced lunch status (see above), and to teacher evaluations ($r = -.14, p = .01$). These results can be found in Table 4.8.
Table 4.8

Bivariate correlations between demographic variable

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Class Size</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Elementary*</td>
<td>-.101*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Jr. High*</td>
<td>-.016</td>
<td>-.160**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>High School*</td>
<td>.096*</td>
<td>-.325**</td>
<td>-.196**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Tested</td>
<td>.127**</td>
<td>-.327**</td>
<td>.300**</td>
<td>.038</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>HST Scores</td>
<td>.009</td>
<td>-.010</td>
<td>-.017</td>
<td>.055</td>
<td>-.032</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Free/Reduced</td>
<td>-.052</td>
<td>.124**</td>
<td>.056</td>
<td>-.173**</td>
<td>.065</td>
<td>-.5**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Highest Education</td>
<td>.077</td>
<td>-.062</td>
<td>-.014</td>
<td>.093*</td>
<td>.020</td>
<td>.075</td>
<td>-.041</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Salary</td>
<td>.033</td>
<td>-.048</td>
<td>-.053</td>
<td>.122**</td>
<td>-.079</td>
<td>.151**</td>
<td>-.179**</td>
<td>.394**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Teacher Eval</td>
<td>-.048</td>
<td>.082</td>
<td>-.001</td>
<td>-.202**</td>
<td>.001</td>
<td>-.057</td>
<td>.187**</td>
<td>-.061</td>
<td>-.135**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Teacher Exp</td>
<td>-.015</td>
<td>-.031</td>
<td>-.037</td>
<td>.094*</td>
<td>-.064</td>
<td>.019</td>
<td>-.077</td>
<td>.256**</td>
<td>.566**</td>
<td>-.048</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Admin Support</td>
<td>-.021</td>
<td>-.061</td>
<td>.028</td>
<td>.001</td>
<td>.009</td>
<td>.185**</td>
<td>-.140**</td>
<td>.018</td>
<td>-.094*</td>
<td>-.094*</td>
<td>-.039</td>
</tr>
<tr>
<td>13</td>
<td>Pol Persuasion</td>
<td>-.01</td>
<td>.132**</td>
<td>.010</td>
<td>-.126**</td>
<td>-.048</td>
<td>.092*</td>
<td>.078</td>
<td>-.043</td>
<td>.046</td>
<td>.046</td>
<td>.035</td>
</tr>
</tbody>
</table>

Codes for grade level grouping were Elementary (1) and Other (0); Jr High (1) and Other (0), and High School (1) and Other (0)
Relations between variables. Bivariate correlations were calculated in order to examine the relations between outcome variables. Pearson-product moment correlations were utilized for this purpose, and scale means, standard deviations, and Pearson’s $r$ can be found in Table 4.9.

Examining the bivariate correlations in Table 4.9, it is clear that teacher-efficacy is an important construct in this study. Significant relations exist between the full scales and the subscales and the other variables in the study, providing important information about the possible role teacher-efficacy plays in teachers’ perceptions of stress due to high-stakes tests. Using Pearson product-moment correlation tests, these relations were examined with several interesting findings emerging.

The relations between teacher self-efficacy and the PS-HST Known Issues and Stress subscales were both negative and small ($r = -.08, p = .07$ for Known Issues and $r = -.16, p < .001$ for Stress). Teachers with higher levels of teacher self-efficacy are less likely to perceive stress or to perceive less stress than teachers with lower levels of self-efficacy, regarding the TSES items framed with the In the weeks leading up to the test qualifier.

Generally, throughout the school year, teachers’ self-efficacy is strongly, positively correlated to their teacher self-efficacy in the weeks leading up to the test ($r = .69, p < .001$). This indicates that teachers that report higher levels of teacher efficacy throughout the school year also report higher levels of teacher-efficacy in the weeks leading up to the test. There is also a small, positive correlation between teacher self-efficacy generally, throughout the school year and their self-reported autonomy-supportive orientation in the classroom ($r = .16, p < .001$ for the TACS scale and $r = .15,$
for the PIS Autonomy subscale), suggesting that teachers with higher levels of teacher-self efficacy generally throughout the school year are slightly more likely to self-report an autonomy-supportive orientation in the classroom. Only one other significant negative relation existed between teacher self-efficacy generally throughout the school year and the other variables: teachers with higher levels of teacher self-efficacy report slightly less likelihood of adopting a controlling-orientation in the classroom ($r = -.10, p = .01$).

Teacher self-efficacy in the weeks leading up to the test was related significantly to several other variables (see Table 4.9). Teachers who perceive a higher level of known issues on the PS-HST scale reported a moderately lower level of teacher-efficacy than teachers who did not perceive these issues ($r = -.25, p < .001$). This suggests that teachers who perceive that high-stakes tests are interfering with their teaching also experience lower teacher-efficacy, or vice versa.

Teacher autonomy support, as measured by the TACS subscale, was positively and significantly related to several of the TSES subscales, suggesting that there is an important relation between teacher self-efficacy and teacher autonomy-supportive orientation toward teaching. Interestingly, teacher autonomy as measured by the TACS Autonomy subscale was significantly and positively related to the PS-HST, Stress subscale. These results indicate that teachers who perceive greater levels of stress because of the known issues related to high-stakes tests also adopt a greater autonomy-supportive orientation for teaching. These results cannot be interpreted as valid at this time due to the conflicting results indicated by the PIS scale; an opposite, although not
statistically significant, relation between the PIS and PS-HST, Stress scales offer mixed findings at best.
Table 4.9  
Means, Standard Deviations, and Correlations

<table>
<thead>
<tr>
<th>Variable</th>
<th>M</th>
<th>SD</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 PS-HST I</td>
<td>4.32</td>
<td>.54</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 PS-HST II</td>
<td>3.80</td>
<td>.78</td>
<td>.75</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 TSES, G</td>
<td>4.04</td>
<td>.56</td>
<td>-.08</td>
<td>-.16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 TSES-G (I)</td>
<td>4.16</td>
<td>.72</td>
<td>-.11</td>
<td>-.18</td>
<td>.74</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 TSES-G (M)</td>
<td>4.29</td>
<td>.70</td>
<td>-.09</td>
<td>-.15</td>
<td>.81</td>
<td>.38</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 TSES-G (E)</td>
<td>3.67</td>
<td>.72</td>
<td>.02</td>
<td>-.03</td>
<td>.81</td>
<td>.37</td>
<td>.55</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7 TSES, I</td>
<td>3.74</td>
<td>.66</td>
<td>-.25</td>
<td>-.36</td>
<td>.69</td>
<td>.59</td>
<td>.55</td>
<td>.49</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8 TSES-I (I)</td>
<td>3.79</td>
<td>.89</td>
<td>-.26</td>
<td>-.34</td>
<td>.44</td>
<td>.71</td>
<td>.18</td>
<td>.14</td>
<td>.79</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9 TSES-I (M)</td>
<td>4.10</td>
<td>.80</td>
<td>-.18</td>
<td>-.28</td>
<td>.66</td>
<td>.37</td>
<td>.80</td>
<td>.40</td>
<td>.40</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 TSES-I (E)</td>
<td>3.35</td>
<td>.62</td>
<td>-.15</td>
<td>-.22</td>
<td>.58</td>
<td>.31</td>
<td>.37</td>
<td>.67</td>
<td>.82</td>
<td>.45</td>
<td>.56</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11 TACS, A-S</td>
<td>3.50</td>
<td>.62</td>
<td>.09</td>
<td>.15</td>
<td>.16</td>
<td>.08</td>
<td>.08</td>
<td>.23</td>
<td>.05</td>
<td>.01</td>
<td>-.00</td>
<td>.13</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12 PIS (Full)</td>
<td>2.81</td>
<td>.57</td>
<td>-.02</td>
<td>-.07</td>
<td>-.08</td>
<td>-.06</td>
<td>-.07</td>
<td>-.02</td>
<td>-.02</td>
<td>-.02</td>
<td>.00</td>
<td>-.09</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13 PIS-Autonomy</td>
<td>3.10</td>
<td>.64</td>
<td>.06</td>
<td>-.05</td>
<td>-.02</td>
<td>-.03</td>
<td>-.02</td>
<td>-.01</td>
<td>.01</td>
<td>.50</td>
<td>.03</td>
<td>.04</td>
<td>-.05</td>
<td>.92</td>
<td></td>
</tr>
<tr>
<td>14 PIS-Control</td>
<td>2.30</td>
<td>.66</td>
<td>-.06</td>
<td>-.07</td>
<td>-.17</td>
<td>-.11</td>
<td>-.12</td>
<td>-.17</td>
<td>-.07</td>
<td>-.02</td>
<td>-.09</td>
<td>-.07</td>
<td>-.12</td>
<td>.75</td>
<td>.41</td>
</tr>
</tbody>
</table>

Figures in boldface indicate the correlation is significant ($p < .01$). Figures with an asterisk* indicate the correlation is significant ($p = .01$).
Stress Items. As the results indicate, the means for the PS-HST scales were moderately high to high, indicating that teachers agree to strongly agree that known issues due to high-stakes tests are occurring in their schools. They also agree that these issues are causing stress. The scale means for the Known Issues subscale and the Perceived Stress subscale were 4.32 and 3.80, respectively. A closer examination of the individual items revealed means ranging from 4.16 at the lowest end (“In my school, using practice standardized tests is common”) to 4.53 (“In my school, there is pressure to maintain or improve our image to the public”). The extent to which teachers perceive these events as stressful ranged from 3.34 (“In my school, using practice standardized tests is common”) to 3.94 (“I feel pressure to improve standardized test scores”) with two items receiving a mean score of 3.93 (“In my school, there is increased pressure in the weeks leading up to the test” and “In my school, I feel pressured to make sure that my students pass the test”). Results for the two PS-HST subscales are detailed in Table 4.10.
Table 4.10  
Means, standard deviations, and ANOVAs, Known Issues scale

<table>
<thead>
<tr>
<th></th>
<th>M</th>
<th>SD</th>
<th>F</th>
<th>df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Teaching Grade Level</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--Elementary</td>
<td>4.45</td>
<td>.45</td>
<td>2.77</td>
<td>1</td>
<td>.097</td>
</tr>
<tr>
<td>--Jr High</td>
<td>4.43</td>
<td>.46</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--High</td>
<td>4.15</td>
<td>.59</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Teacher report of school HST scores</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--Below (“We are a failing school”)</td>
<td>4.42</td>
<td>.48</td>
<td>2.45</td>
<td>1</td>
<td>.01</td>
</tr>
<tr>
<td>--At Level</td>
<td>4.36</td>
<td>.51</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--Performs Well</td>
<td>4.19</td>
<td>.60</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Teacher report of neighborhood</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--Rural</td>
<td>4.32</td>
<td>.51</td>
<td>.032</td>
<td>1</td>
<td>.969</td>
</tr>
<tr>
<td>--Suburban</td>
<td>4.32</td>
<td>.57</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--Urban</td>
<td>4.33</td>
<td>.54</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Political Persuasion</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--Very Liberal</td>
<td>4.23</td>
<td>.70</td>
<td>.613</td>
<td>4</td>
<td>.653</td>
</tr>
<tr>
<td>--Liberal</td>
<td>4.28</td>
<td>.59</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--Moderate</td>
<td>4.34</td>
<td>.51</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--Conservative</td>
<td>4.36</td>
<td>.52</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--Very Conservative</td>
<td>4.35</td>
<td>.54</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Teacher Experience</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--Five years or less</td>
<td>4.31</td>
<td>.53</td>
<td>1.19</td>
<td>2</td>
<td>.273</td>
</tr>
<tr>
<td>--Six years to 20 years</td>
<td>4.35</td>
<td>.53</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--Over 20 years</td>
<td>4.39</td>
<td>.54</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Gender</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--Male</td>
<td>4.16</td>
<td>.59</td>
<td>15.9</td>
<td>1</td>
<td>.01</td>
</tr>
<tr>
<td>--Female</td>
<td>4.37</td>
<td>.52</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Teacher Evaluations</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--None</td>
<td>4.13</td>
<td>.57</td>
<td>23.6</td>
<td>3</td>
<td>.01</td>
</tr>
<tr>
<td>--Some</td>
<td>4.41</td>
<td>.48</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--Nearly Half</td>
<td>4.55</td>
<td>.43</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--Over Half</td>
<td>4.78</td>
<td>.26</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

As indicated in Figure 4.1, the mean responses on the two stress subscales (Known Issues and Perceived Stress) increased consistently according to the self-reported
incidence of free/reduced lunch (FRL). Teachers report more stress due to high-stakes tests according to the amount of students they teach who qualify for free or reduced lunch—students who qualify for this support typically come from low-income families. These differences were statistically significant ($F = 4.39, df = 9, p < .001$ for the Known Issues subscale; $F = 2.98, df = 9, p = .002$ for the Stress subscale). These findings suggest that teachers perceive greater amounts of stress due to high-stakes tests in schools where students are more frequently reliant on FRL than in schools where they are less so.

![Figure 4.1](image)

**Figure 4.1**
*Means on the PS-HST, Known Issues and Stress scales according to self-reported free/reduced lunch status. Figures on the x-axis indicate the percentage of students that qualify for FRL.*
Elementary teachers report slightly higher levels of perceived stress as a result of high-stakes tests than junior high or high school teachers as demonstrated in Figure 4.2. Elementary teachers agree to strongly agree with the issues occurring as a result of high-stakes tests \((N = 73, M = 4.47, SD = .45; \text{See Table 4.10})\). They perceive higher levels of stress as a result of this \((M = 4.04, SD = .79; \text{Table 4.11})\) than the teachers in the junior high or high school groups. Junior high teachers \((N = 117)\) reported agreement to strong agreement on the Known Issues subscale \((M = 4.40, SD = .49; \text{Table 4.10})\) and perceived stress as a result of these issues \((M = 3.88, SD = .73; \text{Table 4.11})\). For high school teachers, the mean on the Known Issues subscale was 4.17 \((N = 251)\) and 3.55 on the Perceived Stress subscale. These means are plotted in Figure 4.2 and reported in Table 4.10 (Known Issues) and 4.11 (Perceived Stress).
Teachers were asked to report how their schools fare on standardized tests. For teachers who reported that their school performed below average, the mean on the Known Issues subscale was 4.42 (SD = .49) and 3.90 (SD = .73) on the Perceived Stress subscale. For teachers in schools that perform at or around the appropriate level, these means were slightly lower (M = 4.36, SD = .51 for Known Issues and M = 3.91, SD = .69 for the Perceived Stress subscale). For teachers in schools that usually pass these tests, the means and standard deviations were lower still; the Known Issues subscale for this group is M = 4.19, SD = .60 and the Perceived Stress subscale is M = 3.57, SD = .86. The difference between the groups is statistically significantly different according to a one-way ANOVA (F = 8.59, df = 2, p < .001 for Known Issues, F = 11.17, df = 2, p < .001 for Perceived Stress). These means are shown in Figure 4.3.

![Figure 4.3](image)

**Figure 4.3**
Means on the Known Issues and Stress subscales for teachers according to the relative success of their school on HST.
The means for these subscales were virtually equal for teachers according to self-reported neighborhood. That is, teachers reporting they teach in rural schools reported very similar beliefs about the Known Issues and Perceived Stress as those in suburban and urban districts. There were no significant differences between these groups. Along these lines, self-reported political persuasion was not related to teachers’ perceptions of stress due to high-stakes tests. Perhaps the most surprising non-finding lies in the amount of experience teachers report—teachers that are new to the field as well as teachers with more than 20 years of experience in the classroom report similar levels on these subscales. Additionally it is important to note there were no significant differences in perceived stress or issues according to gender or race. Data for these findings are reported in Table 4.10.

Teachers were asked to report how much of their own evaluations as professionals was dependent upon high-stakes test scores. Teachers who are evaluated based upon their students’ high-stakes test scores report higher perceived stress due to high-stakes tests than do teachers who are not evaluated based on high-stakes test scores. The difference between these groups was statistically significant ($F = 23.61, df = 3, p < .001$ for Known Issues; $F = 24.00, df = 3, p < .001$ for Perceived Stress). These means are plotted in Figure 4.4, and reported in Table 4.10.
Figure 4.4
Means for Known Issues and Stress subscales according to teacher evaluations.
Table 4.11  
*Means, standard deviations, and ANOVAs, Perceived Stress scale*

<table>
<thead>
<tr>
<th></th>
<th>M</th>
<th>SD</th>
<th>F</th>
<th>df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Teaching Grade Level</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--Elementary</td>
<td>4.45</td>
<td>.45</td>
<td>2.40</td>
<td>2</td>
<td>.122</td>
</tr>
<tr>
<td>--Jr High</td>
<td>4.43</td>
<td>.46</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--High</td>
<td>4.15</td>
<td>.59</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Teacher report of school HST scores</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--Below (“We are a failing school”)</td>
<td>3.90</td>
<td>.73</td>
<td>11.2</td>
<td>2</td>
<td>.01</td>
</tr>
<tr>
<td>--At Level</td>
<td>3.91</td>
<td>.69</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--Performs Well</td>
<td>3.57</td>
<td>.86</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Teacher report of neighborhood</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--Rural</td>
<td>3.86</td>
<td>.74</td>
<td>1.48</td>
<td>2</td>
<td>.228</td>
</tr>
<tr>
<td>--Suburban</td>
<td>3.73</td>
<td>.84</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--Urban</td>
<td>3.79</td>
<td>.72</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Political Persuasion</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--Very Liberal</td>
<td>3.79</td>
<td>.77</td>
<td>.802</td>
<td>4</td>
<td>.524</td>
</tr>
<tr>
<td>--Liberal</td>
<td>3.80</td>
<td>.82</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--Moderate</td>
<td>3.77</td>
<td>.78</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--Conservative</td>
<td>3.87</td>
<td>.73</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--Very Conservative</td>
<td>3.58</td>
<td>.89</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Teacher Experience</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--Five years or less</td>
<td>3.83</td>
<td>.69</td>
<td>.113</td>
<td>2</td>
<td>.894</td>
</tr>
<tr>
<td>--Six years to 20 years</td>
<td>3.79</td>
<td>.75</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--Over 20 years</td>
<td>3.83</td>
<td>.82</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Gender</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--Male</td>
<td>3.45</td>
<td>.89</td>
<td>34.7</td>
<td>1</td>
<td>.01</td>
</tr>
<tr>
<td>--Female</td>
<td>3.90</td>
<td>.71</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Teacher Evaluations</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--None</td>
<td>3.52</td>
<td>.79</td>
<td>24.0</td>
<td>3</td>
<td>.01</td>
</tr>
<tr>
<td>--Some</td>
<td>3.91</td>
<td>.70</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--Nearly Half</td>
<td>4.17</td>
<td>.67</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--Over Half</td>
<td>4.46</td>
<td>.48</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Teacher Self-Efficacy.** A paired-samples t-test conducted on the TSES

*Generally, throughout the school year and the TSES In the weeks leading up to the test*
indicated a statistically significant difference between the means for these scales \((t = 13.77, df = 528, p < .001)\). These results indicate that teachers perceive significantly lower levels of teacher self-efficacy in the weeks before high-stakes tests than they do throughout the remainder of the school year. Refer to Table 4.12 for detailed findings.

Looking at specific subgroups, teacher efficacy was related to HST scores and perceptions of administrative support. Teachers who reported that their students did not generally pass the standardized test also reported lower levels of teaching efficacy compared to their counterparts in higher-performing schools. Interestingly, teachers in schools hovering around the appropriate passing rate for HST also reported lower teaching efficacy than teachers in higher-performing schools, and nearly identical teaching efficacy as those in lower-performing schools. The means for these results are compared in figure 4.5. These findings were significant \((F = 6.38, df = 2, p < .001 \text{ for TSES-G}, F = 14.80, df = 2, p < .001 \text{ for TSES-I})\) and are detailed in Table 4.12.

![Figure 4.5](image)

*Figure 4.5*

*Means for the TSES, Generally throughout the school year and in the weeks leading up to the test according to HST performance*
Administrative support was also related to teacher efficacy, with teachers reporting that their administration is not supportive of them also reporting lower levels of teacher self-efficacy, and teachers reporting their administration is supportive reporting higher teacher self-efficacy, with the differences between the groups statistically significant ($F = 6.17, df = 4, p < .001$ for TSES-G, and $F = 3.19, df = 4, p < .001$ for TSES-I). Results for these findings are reported in Table 4.12.

![Graph showing means on the TSES scales according to the item “My administration is supportive of my work as a teacher.”](image)

**Figure 4.6**
*Means on the TSES scales according to the item “My administration is supportive of my work as a teacher.”*

Teacher self-efficacy varied significantly for teachers of differing races on both subscales. Asian teachers ($N = 3$) reported the lowest teacher self-efficacy in the weeks leading up to the test ($M = 3.08, SD = .43$). Black teachers report the highest levels of teacher self-efficacy in the weeks leading up to the test ($N = 9, M = 4.02, SD = .74$). The ANOVA was conducted on white teachers compared to teachers of all other races due to
the small N of each of the minority subgroups. These results are presented in Figure 4.6, and the ANOVAs can be found in Table 4.12.

Figure 4.7
*Means for teacher self-efficacy on both scales according to teacher race.*

Teacher self-efficacy did not vary significantly according to grade level taught, free/reduced lunch status, self-reported neighborhood of the school, teaching experience, gender, teacher evaluation policy, or class size in this sample.
Table 4.12
*Means, Standard Deviations, and results on ANOVA and t-test for TSES*

<table>
<thead>
<tr>
<th></th>
<th>M</th>
<th>SD</th>
<th>F or t</th>
<th>df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>TSES Generally, throughout year</td>
<td>4.04</td>
<td>.56</td>
<td>13.77</td>
<td>528</td>
<td>.01</td>
</tr>
<tr>
<td>TSES In the weeks leading up to the test</td>
<td>3.74</td>
<td>.66</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TSES <em>Generally, throughout year</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HST Scores</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--Below “we are a failing school”</td>
<td>3.98</td>
<td>.57</td>
<td>6.38</td>
<td>2</td>
<td>.002</td>
</tr>
<tr>
<td>--At level</td>
<td>3.98</td>
<td>.56</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--Well performing school</td>
<td>4.17</td>
<td>.53</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perceptions of Administrative Support</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--Strongly disagree</td>
<td>3.81</td>
<td>.64</td>
<td>6.17</td>
<td>4</td>
<td>.01</td>
</tr>
<tr>
<td>--Disagree</td>
<td>4.02</td>
<td>.64</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--Neutral</td>
<td>3.82</td>
<td>.57</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--Agree</td>
<td>4.01</td>
<td>.56</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--Strongly agree</td>
<td>4.19</td>
<td>.51</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--White</td>
<td>4.04</td>
<td>.55</td>
<td>.118</td>
<td>1</td>
<td>.732</td>
</tr>
<tr>
<td>--Other</td>
<td>4.00</td>
<td>.76</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TSES <em>In the weeks leading up to the test</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HST Scores</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--Below “we are a failing school”</td>
<td>3.61</td>
<td>.67</td>
<td>6.38</td>
<td>2</td>
<td>.002</td>
</tr>
<tr>
<td>--At level</td>
<td>3.67</td>
<td>.65</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--Well performing school</td>
<td>3.96</td>
<td>.61</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perceptions of Administrative Support</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--Strongly disagree</td>
<td>3.65</td>
<td>.72</td>
<td>7.62</td>
<td>4</td>
<td>.01</td>
</tr>
<tr>
<td>--Disagree</td>
<td>3.64</td>
<td>.69</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--Neutral</td>
<td>3.46</td>
<td>.70</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--Agree</td>
<td>3.70</td>
<td>.63</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Autonomy-supportiveness and controllingness.** Elementary school teachers reported a slightly higher autonomy-supportive orientation than middle school or high school teachers, as indicated in Figure 4.7.
Table 4.13
Means and Standard Deviations for the Problems in Schools Questionnaire, according to Teacher Evaluation and Grade Level.

<table>
<thead>
<tr>
<th>Grade Level Taught</th>
<th>M</th>
<th>SD</th>
<th>F</th>
<th>df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elementary</td>
<td>3.42</td>
<td>.39</td>
<td>4.35*</td>
<td>1</td>
<td>.037</td>
</tr>
<tr>
<td>Jr High</td>
<td>3.37</td>
<td>.43</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High School</td>
<td>3.38</td>
<td>.39</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Results of the ANOVA indicate significant differences between elementary and the other grades only.

Autonomy-support, as measured by the PIS subscale, was positively and significantly related to the Known Issues subscale ($r = .14, p < .001$), indicating teachers who are oriented toward adopting an autonomy-supportive orientation in the classroom are slightly more likely to perceive disruptions in their classroom due to high-stakes tests. These results were also significant with the Stress subscale, indicating that teachers with an autonomy-supportive orientation perceived the disruptions due to high-stakes tests as stressful ($r = .09, p < .01$). These findings are reported in Table 4.9.

As mentioned above, the PIS Autonomy subscale was positively and significantly related to several of the TSES subscales. Autonomy-supportive orientations were positively, and significantly, related to the TSES Generally, throughout the school year ($r = .15, p < .001$), as well as the three Generally subscales, Instruction ($r = .13, p < .001$), Classroom Management ($r = .10, p = .01$), and Engagement ($r = .12, p < .001$).

Autonomy-support as measured by the PIS Autonomy was not significantly related to any of the TSES In the weeks leading up to the test measures. This suggests that the extent to which a teacher adopts an autonomy-supportive orientation toward teaching is related to teacher self-efficacy throughout the school year, however, in the weeks before testing this relation is disrupted.
The PIS was not significantly related to any other variables in the study, nor were there any other significant differences in autonomy-supportiveness and controllingness evident. This may be a result of the challenges in measurement that this tool presented; these issues will be discussed further in Chapter 5.

**Multiple Regression Analyses**

In order to test the hypotheses of the research study, a series of hierarchical multiple regression (HMR) analyses were conducted. The purpose of the HMR analyses was to model a parsimonious relation between teacher perceptions of stress due to high-stakes tests and teacher motivation (defined as teacher self-efficacy).

**Prediction of teacher self-efficacy, generally, throughout the school year.**

Standardized demographic variables were entered into the regression analyses at the first step in order to control for variance (Years teaching experience; teacher evaluations; high-stakes test scores; grades taught; highest education; income; free/reduced lunch; class size; perceptions of administrative support) and dummy coded variables were entered at the first step (gender, race, teacher report of neighborhood of school, tested/non-tested content area).

At Step 2 of the regression, the two variables used to measure perceived stress as a result of high-stakes tests were entered into the equation (PS-HST; Known Issues and Perceived Stress), and teacher evaluations, high-stakes test scores, tested/non-tested content area, grade level, highest education, neighborhood, free/reduced lunch status, class size, gender, and race were dropped from the model, as they failed to predict TE. The significant predictors that remained were teacher income, perceptions of
administrative support, and level of teaching experience. Dropped predictors did not affect the significance of the model or significant predictors.

Step 3 of the regression model included entering the three hypothesized interaction variables after centering them by using the standardized score. The interaction variables were not significant predictors of TE, and therefore were dropped for the final model. The results indicated multicollinearity was not an issue for this model (Tolerance .664 - .983; VIF 1.0198 – 1.505). Dropped predictors did not affect the significance of the model or significant predictors.

The model accounted for 6.4% of the variance in teacher self-efficacy generally throughout the school year, a significant amount ($F = 6.03; df 1, 510; p = .014$). The variables accounted for a significant amount of the variance at each step, with the exception of PS-HST, Known Issues, which was not a significant predictor of teacher self-efficacy, and therefore eliminated from the model.

The model accounted for a significant amount of the variance at Step 1 (5.4%, $p < .001$) and at Step 2 (6.4%, $p < .01$). Adding the stress variables into the model resulted in an additional 2.1% of variance accounted for—a small, but statistically significant amount. The variables included at each step of the model predicted a significant amount of the variance at each step ($\Delta R^2 = 0.054, 0.064; p < .01$ and .014, Table 4.14). The final model, including significant predictors, was as follows:

$$
TSESGen = 3.62 + 0.01(teacher\ income) + .10(\text{perception\ of\ administrative\ support}) + -0.01(teacher\ experience) + -0.08(\text{Perceived\ Stress}).
$$

Only significant predictors were included in Table 4.16. The results indicate, at Step 1, teacher income, perceptions of administrative support, and level of teaching
experience were all significant predictors of teacher self-efficacy. These also continued to be significant predictors at Step 2, when Perceived Stress was entered ($\beta = -.079$, $p = .014$ at Step 2). This suggests that teachers who perceive higher levels of stress due to the test also report lower levels of teacher self-efficacy than those who perceive less stress.

The greater income a teacher has, the greater the teacher self-efficacy beliefs they report. Perceptions of administrative support also significantly predict teacher self-efficacy beliefs; teachers who perceive their administrators as supportive are more likely to have higher levels of teacher self-efficacy than teachers who do not. Additionally, the results indicate that greater teaching experience is a significant negative predictor of teacher self-efficacy, generally. This means that teachers’ self-efficacy beliefs are related inversely to teaching experience in this sample. The dropped predictors did not affect the significance of the model or the predictors.

Table 4.14 (continued)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Dist.</th>
<th>SD</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teacher Income</td>
<td>50</td>
<td>14.32</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Perceptions of Admin Support</td>
<td>3.98</td>
<td>.97</td>
<td>.97</td>
<td>.01</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Teaching Experience</td>
<td>15</td>
<td>10.33</td>
<td>.566**</td>
<td>-.039</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Perceived Stress</td>
<td>3.79</td>
<td>.78</td>
<td>-.163**</td>
<td>-.116**</td>
<td>-.057</td>
<td>3</td>
</tr>
</tbody>
</table>

**Correlation significant (.001)

Table 4.15

Regression results, including change in $R^2$ for the TSESGen Model

<table>
<thead>
<tr>
<th>Step</th>
<th>$R$</th>
<th>Adjusted $R^2$</th>
<th>$SE$</th>
<th>$R^2$ change</th>
<th>$F$ change</th>
<th>$df1$</th>
<th>$df2$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.245</td>
<td>.054</td>
<td>.55</td>
<td>.06</td>
<td>10.84</td>
<td>3</td>
<td>511</td>
<td>&lt; .01</td>
</tr>
<tr>
<td>2</td>
<td>.266</td>
<td>.064</td>
<td>.55</td>
<td>.011</td>
<td>6.03</td>
<td>1</td>
<td>510</td>
<td>.014</td>
</tr>
</tbody>
</table>

TSES generally throughout the year is the outcome variable for this model. Entered at Step 1, demographic variables; Step 2, Perceived Stress
Table 4.16
Regression results; coefficients at Step 2 of the TSESGen model

<table>
<thead>
<tr>
<th></th>
<th>B</th>
<th>SE</th>
<th>ß</th>
<th>t</th>
<th>Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Step 1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teacher Income</td>
<td>.008</td>
<td>.002</td>
<td>.210</td>
<td>4.031</td>
<td>.01</td>
</tr>
<tr>
<td>Perceptions of Admin Support</td>
<td>.100</td>
<td>.025</td>
<td>.170</td>
<td>3.954</td>
<td>.01</td>
</tr>
<tr>
<td>Teaching Experience</td>
<td>-.007</td>
<td>.003</td>
<td>-.120</td>
<td>-2.305</td>
<td>.022</td>
</tr>
<tr>
<td><strong>Step 2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teacher Income</td>
<td>.008</td>
<td>.002</td>
<td>.190</td>
<td>3.629</td>
<td>.01</td>
</tr>
<tr>
<td>Perceptions of Admin Support</td>
<td>.092</td>
<td>.025</td>
<td>.156</td>
<td>3.633</td>
<td>.01</td>
</tr>
<tr>
<td>Teaching Experience</td>
<td>-.006</td>
<td>.003</td>
<td>-.115</td>
<td>-2.214</td>
<td>.027</td>
</tr>
<tr>
<td>Perceived Stress</td>
<td>-.079</td>
<td>.032</td>
<td>-.107</td>
<td>-2.455</td>
<td>.014</td>
</tr>
</tbody>
</table>

**Prediction of teacher self-efficacy in the weeks leading up to the test.**

Standardized demographic variables were entered into the regression analyses at the first step in order to control for variance (Years teaching experience; teacher evaluations; high-stakes test scores; tested/non-tested content area; grades taught; highest education; income; neighborhood; free/reduced lunch; class size; administrative support; gender; race).

At Step 2 of the regression, the two variables used to measure perceived stress as a result of high-stakes tests were entered into the equation (PS-HST; Known Issues and Perceived Stress), and teacher evaluations, level of teaching experience, tested/non-tested content area, grade level, highest education, neighborhood, free/reduced lunch status, class size, gender, and race were dropped from the model, as they failed to predict TE. The significant predictors that remained were teacher income, perceptions of administrative support, and school history of high-stakes test scores.

Step 3 of the regression model included entering the three hypothesized interaction variables. The interaction variables were not significant predictors of TE, and therefore were dropped for the final model. The results indicated multicollinearity was
not an issue for this model (Tolerance .929 - .977; VIF 1.024 – 1.076). Dropped predictors did not affect the significance of the model or significant predictors.

The model accounted for 16.8% of the variance in teacher self-efficacy in the weeks leading up to the test, a significant amount \( (F = 50.93; df\ 1, 508; p < .01) \). The variables accounted for a significant amount of the variance at each step, with the exception of PS-HST, Known Issues, which was not a significant predictor of teacher self-efficacy, and therefore eliminated from the model.

The model accounted for a significant amount of the variance at Step 1 (8.7%, \( p < .001 \)) and at Step 2 (16.8%, \( p < .01 \)). Adding the stress variables into the model resulted in an additional 8.1% of variance accounted for—a significant amount. The variables included at each step of the model predicted a significant amount of the variance at each step (\( \Delta R^2 = 0.087, 0.168; p < .01 \) and .01, Table 4.16). The final model, including significant predictors, was as follows:

\[
\text{TSESTest} = 3.972 + 0.005(\text{teacher income}) + 0.088(\text{perception of administrative support}) + 0.103(\text{high-stakes test score}) - 0.251(\text{Perceived Stress}).
\]

Only significant predictors were included in Table 4.14. The results indicate, at Step 1, teacher income, perceptions of administrative support, and high-stakes test scores were all significant predictors of teacher self-efficacy in the weeks before the test. These also continued to be significant predictors at Step 2, when Perceived Stress was entered (\( \beta = -0.297, p < .01 \) at Step 2). This suggests that teachers who perceive higher levels of stress due to the test also report lower levels of teacher self-efficacy in the weeks before the test than those who perceive less stress. The greater income a teacher has, the greater the teacher self-efficacy beliefs they report. Perceptions of administrative support also
significantly predict teacher self-efficacy beliefs; teachers who perceive their administrators as supportive are more likely to have higher levels of teacher self-efficacy before the test than teachers who do not. Additionally, the results indicate that high-stakes test scores are a significant predictor of teacher self-efficacy before the test. This means that teachers’ self-efficacy beliefs are higher for teachers in schools that typically perform better on the HST.

Table 4.17
Means, Standard Deviations, and Correlations for TSES, in the weeks before the test

<table>
<thead>
<tr>
<th>Variable</th>
<th>Dist</th>
<th>SD</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teacher Income</td>
<td>50</td>
<td>14.32</td>
<td>.97</td>
<td>.01</td>
<td></td>
</tr>
<tr>
<td>Perceptions of Admin Support</td>
<td>3.98</td>
<td>.97</td>
<td>.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HST Scores</td>
<td>2</td>
<td>.79</td>
<td>.151**</td>
<td>.179**</td>
<td></td>
</tr>
<tr>
<td>Perceived Stress</td>
<td>3.79</td>
<td>.78</td>
<td>-.163**</td>
<td>-.116**</td>
<td>-.167**</td>
</tr>
</tbody>
</table>

**Correlation significant (.001)

Table 4.18
Regression results, including change in $R^2$ for the TSESTest

<table>
<thead>
<tr>
<th>Step</th>
<th>$R$</th>
<th>Adjusted $R^2$</th>
<th>$SE$</th>
<th>$R^2$ change</th>
<th>$F$ change</th>
<th>df1</th>
<th>df2</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.304</td>
<td>.092</td>
<td>.63</td>
<td>.092</td>
<td>17.22</td>
<td>3</td>
<td>509</td>
<td>&lt; .01</td>
</tr>
<tr>
<td>2</td>
<td>.418</td>
<td>.175</td>
<td>.60</td>
<td>.083</td>
<td>50.93</td>
<td>1</td>
<td>508</td>
<td>&lt; .01</td>
</tr>
</tbody>
</table>

TSES generally throughout the year is the outcome variable for this model. Entered at Step 1, demographic variables; Step 2, Perceived Stress

Table 4.19
Regression results; coefficients at Step 2 of the TSESTest

<table>
<thead>
<tr>
<th>B</th>
<th>SE</th>
<th>$\beta$</th>
<th>$t$</th>
<th>Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step 1 Teacher Income</td>
<td>.007</td>
<td>.002</td>
<td>.151</td>
<td>3.531</td>
</tr>
<tr>
<td>Perceptions of Admin Support</td>
<td>.110</td>
<td>.029</td>
<td>.161</td>
<td>3.753</td>
</tr>
<tr>
<td>HST Score</td>
<td>.135</td>
<td>.036</td>
<td>.163</td>
<td>3.753</td>
</tr>
<tr>
<td>Step 2 Teacher Income</td>
<td>.005</td>
<td>.002</td>
<td>.108</td>
<td>2.611</td>
</tr>
<tr>
<td>Perceptions of Admin Support</td>
<td>.088</td>
<td>.028</td>
<td>.129</td>
<td>3.135</td>
</tr>
<tr>
<td>HST Score</td>
<td>.103</td>
<td>.035</td>
<td>.124</td>
<td>2.967</td>
</tr>
<tr>
<td>Perceived Stress</td>
<td>-.251</td>
<td>.035</td>
<td>-.297</td>
<td>-7.136</td>
</tr>
</tbody>
</table>
Prediction of teacher autonomy-supportiveness and controllingness.

Standardized demographic variables were entered into the regression analyses at the first step in order to control for variance (Years teaching experience; teacher evaluations; high-stakes test scores; tested/non-tested content area; grades taught; highest education; income; neighborhood; free/reduced lunch; class size; administrative support; gender; race).

At Step 2 of the regression, the two variables used to measure perceived stress as a result of high-stakes tests were entered into the equation (PS-HST; Known Issues and Perceived Stress), and all predictor variables except level of teaching experience and extent which teachers are evaluated based upon student test scores were removed from the model, as they failed to predict autonomy-supportiveness and controllingness.

Step 3 of the regression model included entering the three hypothesized interaction variables. The interaction variables were not significant predictors of TE, and therefore were dropped for the final model. The results indicated multicollinearity was not an issue for this model (Tolerance .664 - .983; VIF 1.0198 – 1.505). Additionally, neither of the stress variables predicted the outcome variable, so these were removed as well. The only variable to predict teacher autonomy-support and controllingness was teacher evaluations.

The model accounted for 1% of the variance in teacher autonomy-supportiveness and controllingness, a significant amount \((F = 5.36; df \ 1, 509; \ p = .021)\). The model accounted for a significant amount of the variance at Step 1 (8%, \(p = .021\)). Adding the stress variables into the model did not result in any additional variance accounted for. The variables included at each step of the model predicted a significant amount of the
variance at each step ($R^2 = 0.1; p = .02$). The final model, including the only significant predictor (teacher evaluations), was as follows:

$$PIS = 34.64 + .509 \text{ (teacher evaluations)}.$$  

<table>
<thead>
<tr>
<th>Step</th>
<th>$R$</th>
<th>Adjusted $R^2$</th>
<th>$SE$</th>
<th>$R^2$ change</th>
<th>$F$ change</th>
<th>$df_1$</th>
<th>$df_2$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.102</td>
<td>.010</td>
<td>4.02</td>
<td>.010</td>
<td>5.363</td>
<td>1</td>
<td>509</td>
<td>.021</td>
</tr>
</tbody>
</table>

**Summary.** The data from this study revealed several new and unique findings. The first of these is the level of stress that teachers perceive as a result of high-stakes testing. Means on the two stress subscales were very high, suggesting teachers in all states and demographic groups are under stress as a result of the requirements placed on them due to tests. The second major finding is that teachers who report more stress also report lower levels of teacher self-efficacy. This indicates that teacher self-efficacy may be suffering as a direct result of high-stakes tests. These findings were more salient for various subgroups of teachers, including minority teachers and teachers who work in schools that do not typically perform well on the test. The third major finding of this study is slightly nuanced; this study revealed the difficulty in measuring the autonomy-support/control construct while simultaneously showing that teachers who are evaluated based upon their students’ test scores report being less autonomy-supportive in the classroom. Overall, these findings contribute to the body of knowledge on the relation of high-stakes testing policies to a variety of classroom outcomes; of greatest interest is the unique contribution that these findings have for the study of teacher motivation. These will be discussed further in Chapter 5.
Chapter Five: Discussion

The purposes of this study were to examine teachers’ perceptions of stress due to high-stakes tests and the relations of this to teacher motivation, operationalized as teacher self-efficacy and autonomy-supportiveness and controllingness. Specifically, this study examined two research questions:

• What is the relation between teachers’ perceived stress as a result of high-stakes tests and teacher self-efficacy?

• What is the relation between teachers’ perceived stress as a result of high-stakes tests and their autonomy-supportive versus controlling orientation toward teaching?

Measuring Teacher Stress and High-Stakes Tests

As described in Chapters 3 and 4, measuring teachers’ perceived stress due to high-stakes testing proved to be a complex task that ultimately required separating the measures into two interdependent tools. In order to avoid confounded findings, thereby invalidating the results of the study, I opted to measure two test-related constructs: the extent to which teachers perceive tests to be disruptive, and the stress that they perceive related to each disruptive event. This method represents the output approach to studying stress, indicating that stress, in this study, is treated as a reaction to a particular condition, and measurement in this case is being directed at the subjective state that results (Lazarus, 1990). For the purposes of this study, the method chosen represents a starting
point for determining whether stress is present for teachers due to the testing context; now that stress has been determined to be a salient issue for teachers, future studies can be designed that begin to analyze the mechanisms and interactions of relevant and related variables.

**Teacher Stress Due to High-Stakes Tests**

Teachers in this study reported several disruptions in their work as a result of HST, and moderate to high levels of stress as a result. Teachers report that they feel high levels of pressure to make certain that their students are passing the test, bolstering findings in many other studies (e.g., Barrett, 2009; Willis & Haymore-Sandholtz, 2009). Fifty-three percent of teachers in this sample report that they “strongly agree” that they feel pressure to make certain that their students are passing the test, and 34% report that they “agree” with the statement ($M = 4.2, SD = .91$). The literature indicates that pressure to pass high-stakes tests causes teachers to focus intently on the material that will be tested, often reducing the curriculum to a series of rote tasks such as memorization of facts (Zastrow & Zanc, 2004). The struggle with testing pressure caused one teacher to “abandon her personal and professional philosophies gleaned from years of experience for a testing curriculum” that reduced the previously rich and authentic discussions about books to mastery of low-level skills that were isolated from real reading, in spite of knowing these changes were limited and not what was best for the students—this teacher specifically felt she was pushed to consider the test above all else (Asaaf, 2006, pp. 164-165).

The resulting perceived stress as a result of the perceived pressure to make certain that their students are passing the test was moderately-high to high for the majority of
teachers in this sample. When presented with the sub-item “This causes me stress,” 69.5% of teachers selected “agree” or “strongly agree,” with a mean response of 3.93 on a 5-point scale. This pressure was related to the adoption of a controlling-orientation toward teaching (Flink, Boggiano, & Barrett, 1990). Additionally, schools that endure these kinds of pressures to improve test scores have been found to frequently rid themselves of students that may pose a threat to the school’s test scores (Haney, 2000). These incidents cannot be concluded from the findings in this data; nevertheless, the evidence from this data suggest the mechanisms of stress due to testing need further examination, and that stress because of the test is a salient, and significant, issue for teachers.

Teachers reported that they feel pressure to maintain or improve the school’s image publicly, and that the pressure seems to increase as the testing window approaches. When presented with the item “In my school there is pressure to maintain or improve our image to the public,” teachers responded with a mean of 4.52 (SD = .65), and subsequently responded with a mean of 3.8 (SD = 1.0) for “This causes me stress.” In previous studies, teachers reported frustration, feeling out of control, feelings of anxiety, shame, embarrassment, and pressure as a result of Arizona’s decision to publicize test scores through the media in 1991, a decision that has been replicated many times across the United States since then (Cimbric, 2002; Smith, 1991). Consequently, many teachers react to these decisions by focusing intensely on the test, preparing their students for test questions in order to avoid future embarrassment, ridicule, and feelings of shame (Cimbric, 2002). From the perspective of the stress framework, these emotions represent a reaction to an external stimulus that is perceived as a threat rather than a
challenge, when the individual perceives that he does not have sufficient resources to effectively cope with the stressor (Lazarus & Folkman, 1984).

Additionally, teachers report relief after the testing window has passed, indicating a possible reduction in stress levels due to testing. Lazarus defines stress as a transaction between an individual and his environment (1990), and that the stress itself lies solely in neither, but is the transaction between the two. Specifically, stress is the “conjunction of a person with certain motives and beliefs with an environment whose characteristics pose harm, threats, or challenges depending on these person characteristics” (Lazarus, 1990, p. 3). High-stakes tests represent the environmental threat that teachers perceive, and upon completion of the test, this transaction is complete, and stress is replaced with relief.

Similarly, teachers were asked to rate the item “In my school, tensions rise as testing time approaches” resulting in a mean of 4.38, with 85.4% of teachers selecting “agree” or “strongly agree.” Additionally, the perceived stress item in direct response to this received a mean response of 3.77, with 61.4% of teachers selecting “agree” or “strongly agree.”

Teachers also report that they commonly use practice standardized tests to prepare their students, suggesting that time in the classroom is being spent toward preparing students to take the test, rather than on instruction. When presented with the item “In my school, using practice standardized tests is common,” 78% of teachers selected “agree” or “strongly agree.” The consequences of this are clear in the literature: test preparation limits high-quality instruction, narrows curriculum, and increases the rote skills of memorization and reduced deeper level processing (Abrams, Madaus, & Pedulla, 2005).
It is undetermined from these results whether this occurs as a result of a decision made by the teacher or if it is a decision handed down from the administration.

Male teachers reported fewer disruptions due to tests, and also reported significantly less test-related stress than did female teachers in this sample (see Table 4.10). The pattern of response for male teachers was similar across all of the Known Issues and Perceived Stress items—that is, males reported lower means on both, for all items. This finding echoes the findings in several other studies on teacher stress, burnout, and gender—for example, Klassen (2010) found slight but persistent higher stress levels for female teachers persisting across domains; that is, female teachers perceived greater stress due to workload as well as student behavior. One potential explanation for this may be that female teachers may have greater stress in the home than male teachers, resulting in a greater overall stress experience (Kokkinos, 2007). In addition, married female teachers reported greater stress than unmarried female teachers in a study of 447 Greek teachers, further supporting this argument (Kokkinos, 2007).

According to the findings in this study, tested content areas are reportedly treated as more important than non-tested content areas, similar to findings in other studies (e.g., Berliner 2009; Diamond, 2007). On the item, “In my school, tested content areas are often treated as more important than non-tested content areas,” teachers responded with a mean score of 4.37 on a 5-point scale. Fifty-three percent of the sample indicated that they “strongly agreed” with this item, while thirty-four percent selected “agree.” In other studies, the consequences of tested content areas being treated as more important than the rest of the curriculum lead to further narrowing of the curriculum, as subject areas that are not tested are frequently less taught, or not at all (Berliner, 2009; Diamond, 2007).
This unintended consequence of high-stakes tests was found to occur more frequently in schools with more minority students (Berliner, 2009). Other effects were noted by Berliner, including lower performance on the SAT for students in states with high-stakes exit exams, less creativity for students, reduction in higher-order cognitive skills for students, and loss of professional identity for teachers (Berliner, 2009).

Teachers report experiencing more stress and more disruptions due to HST in elementary grades than in high school grades, according to the findings in this study (see Table 4.10.). Students in many states are required to test through the 10th grade, so it is possible that the reduced levels of stress reported in this sample reflect the responses of teachers of grades that are no longer tested. Teachers of elementary grade levels reported a mean of 4.47 ($SD = .45$) for the Known Issues items, and a mean of 4.04 ($SD = .61$) for the Perceived Stress items. For the middle/junior high levels, the results were slightly lower ($Known Issues: M = 4.4; SD = .49$, $Perceived Stress: M = 3.88; SD = .73$). For high school teachers, these results were different yet again ($Known Issues: M = 4.17; SD = .46$, $Perceived Stress: M = 3.55; SD = .83$).

Teachers reported more test-related issues and experienced more stress as a result according to the level of free/reduced lunch status in the school. In other words, the more students that were eligible for free/reduced lunch in the school, the more likely the teacher was to report that HST caused disruptions, and the more likely the teacher was to report perceived stress as a result of this (See Tables 4.10 and 4.11). There are many possible explanations for this. Free/reduced lunch status is often used as a crude proxy measure for poverty or socioeconomic status (Lubienski & Crane, 2010). These results indicate that stress may be a more salient issue for teachers in lower-income school
districts than for teachers in more affluent areas. Given the findings in this study indicate that teachers of lower SES schools and lower grade levels experience more test-related stress than other groups, examining this subgroup of teachers for the effects of test-related stress on teaching effectiveness, burnout, and other classroom-related issues is warranted and will be carried out in future studies.

Disruptions due to tests and stress as a result were more salient for teachers working in schools that generally do not pass the test. This finding was expected; schools that generally underperform on tests are more frequently threatened with sanctions such as funding cuts, staffing cuts or reorganization, or takeover from state agencies (Boardman & Woodruff, 2004; Darling-Hammond, 2007; 2009; United States Congress, 2001, p. 2; Welner, 2005). For teachers in schools that generally perform well on standardized tests, perception of disruptions as well as stress were significantly lower, however still quite high. These findings suggest that teachers in schools performing at all levels perceive HST as disruptive to their work and perceive these disruptions as stressful. While the stress experience may be qualitatively different for teachers depending on their specific contextual situation, these findings suggest that all teachers are experiencing the stress in some capacity. In future research, I hope to examine how the stress experience varies from school to school, and from teacher to teacher.

Teachers in schools that hover right around the standard for passing or failing on the test reported just as much stress due to disruptions as did the teachers in underperforming schools (See Table 4.10). Teachers in schools that are close to the criteria for failing appear to be experiencing the same levels of stress as teachers in lower performing schools. From the stress framework, this is logical, as many school districts
(e.g., Denver, Houston) tie teacher pay to student performance on tests. This is an event that could be perceived as threatening to an individual, particularly for individuals in lower-performing schools, resulting in an increased fear of potential loss, ultimately resulting in a host of consequences, such as reduced resilience, maladaptive coping, or reduced function (in this case, in the classroom) (Weinstein & Ryan, 2010). There are many implications of this, particularly for schools in states where the standards are constantly changing. For teachers it is not only being employed in a low-performing school that causes perceived stress; the threat of possibly falling into the failing range or being labeled a failing school is related to these levels of stress for teachers. The results from this study are clear from a motivational perspective: test-related stress does not appear to foster motivated teaching in the classroom.

For teachers whose evaluation depends on their students’ test scores, known issues reported were much greater than for teachers who are not evaluated based upon student test scores (How much of your evaluation as a teacher depends on your students’ test scores? None; Some; Nearly Half; Over Half). Teachers who report that none of their teacher evaluation depends upon their students’ test scores reported a mean of 4.13 on the Known Issues scale and 3.52 on the Perceived Stress scale, whereas teachers who report that over half of their teacher evaluation depends upon their students’ test scores reported a mean of 4.78 on the Known Issues scale and 4.46 on the Perceived Stress scale (See Figure 4.4; Table 4.10). For teachers whose evaluations depend heavily on student test scores, the results on the Known Issues scale were very high, with a mean score of 4.78 on a 5-point scale. For this same subgroup of teachers, results on the Perceived Stress scale were also quite high, with a mean score of 4.46 on a 5-point scale. These findings,
while critically important, were not unexpected. Interestingly, for teachers whose evaluations are not based on student test scores, these results were also quite high, suggesting that the presence of such pressures are salient for all teachers, and to a greater extent for those teachers who are subjected to these evaluation policies.

The differences for both scales were statistically significant, indicating a real pattern of stress for teachers whose evaluations are based on student scores on high-stakes tests. As this policy continues to gain momentum in state legislatures around the United States, further research is needed to understand how teacher evaluation decisions impact teacher motivation, teacher effectiveness, and teacher job satisfaction and job turnover. High-stakes testing is not just high-stakes for students; the stakes of evaluation policies such as these extend beyond students to teachers as well. As described above, this policy has great potential to be perceived as a threat to individuals, and in many cases the individual will also perceive a lack of resources to handle the perceived threat, resulting in a host of negative outcomes.

**Summary of stress-related findings.** To briefly summarize the findings from the stress-related portion of the study, teachers perceive a great amount of test-related disruptions in their classrooms, and consider these interruptions to be considerably stressful to them professionally. Specifically, teachers perceive pressure to increase student test scores, pressure to ensure their students pass the test, and pressure to improve or maintain the school’s image in the eyes of the public. Teachers perceive a greater level of both test-related disruptions and subsequent stress depending upon the extent to which their performance as teachers is evaluated based on student test scores, and whether the school in which they teach historically passes the test or not. Additionally, female
teachers perceive greater levels of test-related stress than male teachers. Finally, teachers in schools with higher percentages of students receiving free and/or reduced lunch perceive greater levels of test-related stress than teachers in schools with fewer students receiving this support.

**General Teacher Self-Efficacy and High-Stakes Tests**

One of the main purposes of this study was to examine the relations between high-stakes testing and teacher self-efficacy. There were significant findings related to high-stakes testing and teacher self-efficacy throughout the school year. In this model, test-related disruptions (*Known Issues*), teacher income level, teachers’ perceptions of administrative support, and teaching experience all predicted teacher self-efficacy throughout the school year. The model in total accounted for a small amount of variance (6.4%; see Table 4.15); however, the model was significant, suggesting an important pattern exists. These findings establish knowledge that lays the groundwork for future research on these issues. These findings will be described in detail next.

**Test-related disruptions (*Known Issues*)**. Test-related disruptions significantly predicted teacher self-efficacy generally throughout the school year after controlling for gender, race, school neighborhood, tested/non-tested content area, teaching experience, teacher evaluations, school high-stakes test scores, grades taught, highest education earned, teacher income, free/reduced lunch status, class size, and perceptions of administrative support. Interestingly, parsing out the distinction between perceived disruptions due to the test and the stress those disruptions caused was useful for the prediction of teacher self-efficacy, as the disruptions were significant predictors of teacher self-efficacy, whereas the perceived stress as a result of them was not. For
teachers, general self-efficacy for teaching (throughout the school year) is slightly related to the disruptions they perceive as a result of the test; however, the stress they perceived as a result was not shown to predict this particular form of teacher self-efficacy.

The magnitude of the relation that perceived disruptions due to tests had to teacher self-efficacy generally throughout the school year was small but significant ($\beta = -0.11$). The results illustrate that teachers that perceive greater levels of test-related disruptions also report lower levels of teacher self-efficacy throughout the school year, after controlling for other variables. For the Generally, throughout the year model, however, these disruptions were less salient for teachers than their level of income, their perceptions of administrative support, or their level of teaching experience. The greatest predictor for this sample, and in this model, was teacher level of income, with small increases in teacher income equating to increased teacher self-efficacy. This was followed by perceptions of administrative support, with perceptions of greater administrative support equating to higher levels of teacher self-efficacy. Next was teaching experience, which revealed a negative relationship—with greater levels of teaching experience related to lower levels of teacher self-efficacy for the model. Finally, test-related disruptions, with greater perceptions of test-related disruptions related to lower levels of teaching self-efficacy (see Table 4.16). Delving into this finding, it appears that teacher income level is a more salient predictor of teacher’s self-efficacy beliefs than test-related stress; another possible explanation is that teachers with greater levels of self-efficacy are selected into higher-paying school districts.

**Perceived stress due to tests.** The stress that teachers perceive due to test-related disruptions was a significant predictor of teacher self-efficacy generally throughout the
school year. This suggests that the test-related stress teachers perceive may become more salient as the testing window approaches; in other words, teachers may be perceiving greater levels of stress depending on how proximal the event is. It cannot be determined from this data whether teacher stress levels are changing; a longitudinal study would need to be conducted to begin exploring the patterns of stress and/or perceived stress in teachers, and this kind of study is warranted.

Teacher income. Teacher income also significantly predicted general teacher self-efficacy in this sample, with small increases in salary equating to higher levels of self-efficacy for teaching ($\beta = .19$). Education level was not a significant predictor, although the two variables were moderately correlated ($r = .39$), suggesting teacher income accounts for some unique variance in teacher self-efficacy, independent of education level. Teaching experience is another variable that is generally highly correlated to teacher salary, and in this study it was, although it was not a perfect correlation ($r = .57$). In sum, it appears that teacher income, while correlated with teacher experience level and teacher education level, did contribute a small but significant amount to the prediction of teacher self-efficacy throughout the school year.

In a study examining teacher attrition among English teachers, Hancock and Scherff (2010) found that teacher salary was a significant predictor of likelihood to leave the field—that is, the lower the salary, the greater the likelihood the teacher would leave the profession. It is unclear how the relation between teacher self-efficacy and teacher salary takes place; this relation will be explored in depth in future research. Teacher income, in this sample, was the strongest of the significant predictors of teacher self-efficacy in the weeks leading up to the test ($\beta = .19$). For this sample under these
conditions (generally throughout the school year), teacher income has a relation almost twice as large as test-related disruptions to teacher self-efficacy (See Table 4.16). This indicates that teachers’ self-efficacy beliefs in particular are much more susceptible to issues of compensation throughout the school year than their perceptions of test-related disruptions.

Additionally, of particular interest was the finding that teacher self-efficacy throughout the year was predicted positively by teacher income, and negatively by teacher experience level (see Table 4.16). This finding is contrary to my expectations as a scholar and also contrary to the literature (e.g., Tickle, Chang, & Kim, 2011); however, there are many possible explanations for this. Teachers that are employed in school districts that are more affluent likely receive higher levels of pay than teachers in less affluent districts; additionally, they likely experience more frequent efficacy-building events as a result of many privilege-related factors such as opportunities for professional development, additional resources in the school building, or fewer students in the classroom. This will be discussed in more depth below.

**Years of experience teaching.** Teaching experience, as mentioned above, was a significant negative predictor of teacher self-efficacy in this model. In this sample, teachers with more experience in the classroom (How many years of teaching experience do you have?) reported lower levels of teacher self-efficacy generally throughout the year ($\beta = -.12$; see table 4.16). The size of the effect is not large, as it is the third largest in this model; however, it is a stronger predictor than test-related disruptions.

Why would teaching experience negatively predict teacher self-efficacy when teacher salary and teacher education level (often highly correlated) are positive
predictors? The three variables were tested for possible interactions; none existed. It is possible, as described above, that the current climate in schools, that is, the high-stakes testing climate, causes affluent school districts to seek out the best teachers that they can afford, resulting in an accumulation of teachers with higher levels of self-efficacy and higher pay, regardless of level of experience in the classroom or education attained. In spite of the fact that most school districts pay teachers on a salary stepladder, some school districts may be paying more than other school districts. This, if accurate, would also imply that school districts that are less affluent are also more likely to employ teachers with lower levels of teacher self-efficacy.

There is another possible explanation for this finding. As the high-stakes testing era ages, it is possible that new teachers are entering the teaching field that are more accepting of the testing policies as a result of being educated in the system as it is today. Younger teachers may not oppose testing policies, as it is all they have known both professionally and in their own academic development as students. These teachers may perceive testing policies as effective teaching, whereas teachers whose professional career outlives the era of testing may have greater resistance to the policy, resulting in a reduction in teacher efficacy levels. If this phenomenon is occurring, it could explain how teachers with more experience in the classroom are reporting lower levels of teacher self-efficacy in spite of education and salary levels.

**Administrative support.** Perceptions of administrative support significantly predicted teacher self-efficacy for this sample, as expected. In studies of perceptions of administrative support and teaching, lack of administrative support was cited as one of the leading causes of teacher attrition (Tickle et al. 2011; Loeb et al. 2005; Luekens 2004;
Worthy 2005). Teacher self-efficacy was related to individual characteristics such as persistence (Gibson & Dembo, 1984) and professional commitment (Coladarci, 1992; Evans & Trimble, 1986). The direct relations between testing, administrative support, and teacher self-efficacy warrants further study that was beyond the scope of the current study; however these findings will provide the grounding framework for future studies in this area. In this sample, perceptions of administrative support predicted teacher self-efficacy generally throughout the year. Whereas the effect was not large ($\beta = .16$), it was approximately one-and-a-half times larger than the effect size for test-related disruptions ($\beta = -.11$). The magnitude of the effect is large enough to warrant further study and large enough to offer support of the related work on the topic.

**Teacher Self-Efficacy in the Weeks Before the Test and HST**

Aligning with the original research question, teacher self-efficacy in the weeks leading up to the high-stakes test was examined in order to determine if teacher motivational beliefs were related to test-related stress. In this case, teacher self-efficacy in the weeks before the test was significantly predicted by teachers’ perceived stress as a result of the test. Additionally, teacher self-efficacy in the weeks leading up to the test was also significantly predicted by teacher income, teachers’ perceptions of administrative support, and the likelihood of the school to pass the high-stakes test. This model accounted for approximately 18% of the variance in teacher self-efficacy in the weeks leading up to the test—a considerable amount in relation to the first model, and a considerable amount for a social sciences investigation. The results will be described in detail below.
**Test-related disruptions (Known Issues).** Known issues due to testing, or test-related disruptions, did not significantly predict teacher self-efficacy in the weeks leading up to the test, after controlling for gender, race, school neighborhood, tested/non-tested content area, teaching experience, teacher evaluations, school high-stakes test scores, grades taught, highest education earned, teacher income, free/reduced lunch status, class size, and perceptions of administrative support. These results suggest that teachers still perceive test-related disruptions in the weeks before testing; however it is the resulting stress that is more strongly related to teacher self-efficacy at this time.

**Perceived stress due to tests.** The stress that teachers reported as a result of the test-related disruptions was a significant predictor of teachers’ self-efficacy in the weeks before the test. The magnitude of this prediction is large ($\beta = -.30$; Keith, 2006), approximately 2.5 times larger than the next largest significant predictor in the model, perceptions of administrative support (See Table 4.18). The model itself accounted for nearly 18% of the variance in teacher self-efficacy before the test, and adding perceived stress doubled the amount of variance accounted for (from approximately 9% to approximately 18%). A shift in the pattern for teacher self-efficacy related to perceived test disruptions and resulting stress occurs for teachers in terms of their general efficacy beliefs and their efficacy beliefs in the weeks before testing. Stress, directly resulting from high-stakes tests, seems to take precedence over all other predictors of teacher self-efficacy as the test window approaches, becoming a strong factor in the teacher self-efficacy beliefs of teachers.

**Teacher income.** The level of income that teachers earn again significantly predicted their teacher self-efficacy, this time in the weeks before high-stakes tests, for
this sample. To explore this a bit further, it appears from this sample that teacher income level is somewhat related to their education and teaching experiences levels; however, teacher income does predict variance in teacher self-efficacy in the weeks before testing that cannot be explained by education or experience. This variance is significantly related to teacher self-efficacy for both measures—in the weeks leading up to the test, and generally throughout the school year.

As described above in the *Generally throughout the year* section, it is possible that teachers working in higher performing districts are paid higher salaries, regardless of experience level or education. Higher paying school districts may be able to afford supports and resources that lower paying districts are not able to, possibly contributing to higher teacher self-efficacy. Additionally, higher-performing districts likely seek out higher-performing teachers for their workforce through recruitment efforts and by being more attractive as a workplace, possibly resulting in a teaching staff that is inclined to have a higher level of teacher self-efficacy than other school districts.

In the weeks leading up to the test, teacher income was the smallest of the significant predictors of teacher self-efficacy ($\beta = .11$); however, this is still considered a moderate effect according to Keith (2006). Therefore the relation of teacher income to teacher self-efficacy in these weeks is small; however, the relation is significant, suggesting an important pattern is present. Interestingly, the effect size of this predictor shifted from being the most important of the predictors in the *Generally, throughout the school year* model to being the least important in the *Weeks before the test model*. It appears that teacher income level is a small, but significant, predictor of teacher self-
efficacy for both time periods; perceived stress due to tests becomes situated as the
prominent predictor for teachers in the weeks before testing.

In future studies, the relation between teacher self-efficacy and teacher income
must be explored in depth. Is teacher self-efficacy subject to increases just by increasing
teacher pay? This is an important, relevant, and timely question that warrants further
investigation and will likely take up residence in my body of research in the future.

Administrative support. Perceptions of administrative support contributed to the
prediction of teacher self-efficacy in the weeks leading up to the test as well as generally
throughout the school year, as described above (β = .13; See Table 4.19). As with teacher
income, the relative importance of the predictor changed as perceived stress due to the
test shifted into the role of greatest predictor of teacher self-efficacy in the weeks before
testing.

School high-stakes test score. The extent to which the school was likely to pass
the high-stakes test significantly predicted teacher self-efficacy, but only in the weeks
leading up to the test (β = 12.9; p < .01). The score of the school on previous years’ tests
was not a significant predictor of teacher self-efficacy beliefs generally throughout the
school year. This finding indicates that teachers may be experiencing a shift in the source
of self-efficacy for teaching, as this indicator was not present as a significant predictor of
teacher self-efficacy generally throughout the school year.

When considering the four sources of self-efficacy according to Bandura (1997),
the school’s previous test scoring patterns could be considered part of all of these. For
example, for a teacher who was part of the previous year’s teaching staff, the previous
year’s test score is considered a mastery experience, the single most salient source of
self-efficacy for the teacher (Bandura, 1997). The teacher who is new to the teaching staff experiences the previous year’s score as a vicarious experience, the second most salient source of self-efficacy (Bandura, 1997). For the teaching staff at large, there is likely much increased discussion, both in meetings and more informal conversation about the test, potentially acting as a source of verbal persuasion—for instance, in many schools there is a rally to get the school geared up for the test in the weeks before, including announcements made to the building, emails to teaching staff, and professional development sessions geared toward raising scores. All of this acts as verbal persuasion to inform self-efficacy beliefs. Lastly, in light of all of this, physical and affective responses to both the test and the increased attention on the test are likely in teachers, informing self-efficacy beliefs yet again. Examples of this exist throughout the literature, as teachers report they are unable to sleep and experience illness due to the test (Dawson, 2011). Verbal persuasion and physical/affective responses comprise the third and fourth most salient sources of self-efficacy (Bandura, 1997). The school’s likelihood of passing the test is relatively important; for this model the effect size was approximately the same size as teachers’ perceptions of administrative support, and larger than teacher income.

**Summary of teacher efficacy-related findings.** To briefly summarize the Teacher Self-Efficacy related findings, teacher self-efficacy throughout the year is predicted by teacher income, the extent that teachers perceive their administration as supportive, teaching experience, and perceived stress as a result of the test. Teacher self-efficacy in the weeks leading up to the test is predicted by perceived stress due to the test to a great extent. Teacher income, perceptions of administrative support, and the
likelihood of the school to pass the test also were other significant predictors for teacher self-efficacy in the weeks leading up to the test.

Teacher self-efficacy in the weeks before the test is significantly lower than teacher self-efficacy throughout the year (See Table 4.12). Additionally, teacher self-efficacy varied significantly for teachers in low-performing schools and for teachers who do not perceive their administrators as supportive in both the throughout the year and before the test time frames.

**Autonomy-supportiveness and controllingness**

Test-related disruptions and associated stress did not predict teacher autonomy-supportiveness and controllingness for this sample. This finding was unexpected; I hypothesized early on that perceived stress would be related to teachers’ likelihood of developing an autonomy-supportive orientation in the classroom versus a controlling orientation in the classroom. However, this could not be determined from these findings. There are several potential reasons for this; the first and most salient possible explanation for the failure of perceived stress to predict teacher autonomy-supportiveness and controllingness rests in the measurement of the construct. This will be explored in detail throughout the remainder of the discussion.

Demographic predictors that were entered into the regression model at Step 1 were all subsequently dropped from the model at Step 2 because they failed to account for any of the variance in teacher autonomy-support and control, with the exception of teacher evaluations (*How much of your evaluation as a teacher depends on your students’ test scores?*). This predictor negatively predicted teacher autonomy-supportiveness and controllingness (β = -.51), meaning that teachers who perceived that
their evaluations were based on student test scores also reported a less autonomy-supportive orientation and more controlling orientation in the classroom. The size of the relation of teacher evaluation policies to teachers’ autonomy-supportiveness and controllingness was quite large (Keith, 2006). The implications of this are unique, yet unsurprising. Research in the area of self-determination theory is clear that situating individuals under controlling circumstances is likely to instigate controlling behavior as a result (Ryan & Weinstein, 2009).

Teachers that are evaluated based upon student test scores experience an event that is, by definition, autonomy suppressing. In work examining autonomous and controlled individuals and stress in the workplace, Weinstein and Ryan argue that individuals that perceive an event as threatening typically perceive the danger as exceeding their resources, whereas individuals that perceive the event as a challenge will perceive their resources for handling the situation as adequate (Weinstein & Ryan 2009a). It is possible that teachers who are told that they will be evaluated based upon their students’ test scores will appraise the situation as a threat that exceeds their resources, leading to the development of a more controlling-orientation in the classroom. Additionally, teacher evaluations that are based upon student test scores could be perceived as controlling events, which have been shown to elicit a more controlling-orientation in the classroom (Ryan & Weinstein, 2009b).

**Scale validation and reliability.** Measuring these constructs proved a much greater challenge than I anticipated. Several issues arose with the use of the *Problems in Schools Questionnaire* (PISQ) beginning with validation of the scale, and continuing through use of the scale to measure the autonomy/control constructs. Additionally,
attempting to validate my own new scale proved unfruitful, as the reliability of the tool never reached acceptable levels, regardless of the variety of efforts I employed in attempt to increase the effectiveness of the scale.

Measuring autonomy-supportiveness and controllingness raised several concerns. The PISQ (Deci, Schwartz, Sheinman, & Ryan 1981) is designed to measure the two constructs on a spectrum from low to high; low scores indicate a controlling orientation toward teaching whereas high scores indicate an autonomy-supportive orientation toward teaching. It may be that the two constructs are orthogonal, and must be measured on separate scales. Although world-renowned scholars in the area of self-determination validated the scale in the original research, reviewing the original research did reveal some concerns with the internal consistency of the measure. The reliabilities presented in the original research ranged from .63 to .8. Additionally, the researchers used an unorthodox method of reliability analysis by reporting correlations between each item and the item’s scale score (1981, p. 645). This scale was selected after considerable searching for a better tool, and also after constructing a new scale to measure the construct, as a backup measure was desired. Seeking such a tool included broad literature reviews, internet searches, and reaching out to other scholars, including the original authors on the study. After a correspondence with Richard Ryan, it became clear that few alternatives existed to measure these constructs, and also that the reliability of the tool was questionable (Richard Ryan, personal contact, June 2011). Ultimately, the decision to use this scale was made with reservations, and those reservations turned out to be warranted.
Perhaps the greatest revelation of the entire dissertation project is that there is a desperate need for an accurate, reliable, and relevant tool to measure teachers’ autonomy-supportiveness and controllingness. This work has revealed itself to be complex and will require effortful, careful, and thoughtful planning and execution. Moving forward with this task, it will be important to examine the factor structure of any measure that is tested to determine if autonomy and control are parallel (as assumed for this study) or orthogonal. That is, if they are parallel, the individual who scores highly on one will score low on the other, and vice versa. If they are orthogonal, it will be possible to score highly on both, and vice versa. Challenges to measuring this construct must be considered and handled from a theoretical perspective, as well as a methodological and psychometric perspective.

Additionally, one of the limitations of the PISQ is that it is a third-person measure. In other words, the items within the measure ask participants what someone else should do in a given situation. A better approach to use might be to utilize a first-person approach, writing items that reflect what a person would do if they were in the situation.

**Limitations**

There were several limitations to the current study that must be addressed. This study was a self-report examination of teacher perceptions of stress as a result of high-stakes tests, and their motivational beliefs as teachers. Methodologically, self-report research is an important aspect of all social sciences; however, it has limited power to inform. Self-report measures are restricted to theory-bound measurement; that is, in order for a measure to have construct validity it must adhere to a theory—a researcher must know what they intend to measure before they measure it (Haeffel & Howard, 2010).
There are many researchers who argue that theory-driven research is limiting, and as researchers we are responsible for wrestling with the objective nature of such data (e.g., Lather, 2006). Additionally, there also exist issues related to the accuracy and relevance of self-report measures that are generated from the perspective of the researcher and “reflect the researcher’s notion of what classroom life is about” (Kaplan, Katz, & Flum, 2011, p. 172). The implication here is evident—methods must be varied if they are to accurately reflect what is really happening in the classroom, in order to prevent the researcher holding the predominant voice and empowering a proactive, collaborative research experience.

Additionally, there exists some intellectual debate about the usefulness of self-reports. Some scholarly work argues that individuals are capable of reporting accurately on cognitive content but not cognitive processes (Haefel & Howard, 2010; Nisbett & Wilson, 1977), and that self-report measures and interviews are sufficient for studies that examine content, including perceptions, beliefs, motivations, and behaviors because individuals are capable of identifying and reporting on these cognitive content. According to Nisbett and Wilson (1977), individuals have greater access to privately held knowledge and can more accurately report on these things than an observer, if it concerns his emotions, plans, or evaluations. Nevertheless, the self-report is not a perfect method of inquiry and its limitations must be acknowledged.

As with any study, these findings must be extended with research that examines these variables, the relations between them, and their relations to other variables, using a variety of methods and designs. Self-report methods are useful tools that inform the body of knowledge on this topic, but in order to capture the phenomena with greater precision,
depth, and integrity to the actual processes that occur in the classroom, methods must be varied and extended. It is important to choose methods that align with the research questions, and incorporate studies that utilize methods, which will allow the researcher to bridge paradigms and deepen understanding of a phenomenon from multiple perspectives.

A second limitation of the current study lies in the methodological decision to measure teachers’ perceptions of teacher self-efficacy at two different time points—obvious weaknesses exist with the accuracy of individual self-reports, particularly when asking individuals to report on their own beliefs referring back to two different times. A more precise method would be to conduct a longitudinal study across the school year and determine how teacher self-efficacy beliefs change as the testing time approaches and then passes. Ideally, an experimental design would bolster the strength of the findings, assuming one could create conditions that allow for measures to be taken in similar settings with and without high-stakes testing conditions.

Additionally, measuring teacher self-efficacy has proven to be a challenge in its own right. Debates continue regarding the specificity of the task that is measured; Bandura (2001) argues that effective self-efficacy scales must be precise and the tasks that are measured must have impact on the domain of functioning. Additionally, self-efficacy items must accurately reflect the construct that is to be measured. It could be argued that teacher self-efficacy scales are not accurate measures of the construct of teaching in today’s hyper-testing classroom environment, and therefore measures of teacher self-efficacy are not as valid as they once were—most well-validated measures of teacher self-efficacy were written to reflect true teaching (e.g., Tschannen-Moran &
Woolfolk Hoy, 2001), while much research examining the effects of high-stakes testing is clear that teaching has been reduced to test-preparation and does not resemble genuine teaching whatsoever.

**Future directions**

As with all good research studies, this dissertation project produced more questions than answers. Moving forward with this work, stress will be included as an important variable in relation to high-stakes tests. Incorporating the mechanisms of stress, examining stress as a process, and framing this work in relevant theories such as Appraisal Theory will build on the knowledge that this study has generated, allowing these findings to serve as a basis for future, evolved, and more complex work. Additionally, stress that teachers experience due to tests, as well as stress on administrators and students, and the interactions of these will be important factors to examine and will bolster this line of research, with important outcomes. The benefits of this work extend beyond the current situated context; that is, it is useful to study these issues to know how current policies are manifesting in classrooms in order to promote policymaking, ideally to improve the education process as it currently stands. However, researching stress in the classroom can also inform important theoretical knowledge and therefore has applicable as well as theoretical validity.

Additionally, high-stakes testing is not the only policy that warrants examination. Teacher evaluation policies became salient throughout the process of this dissertation study as a potentially harmful event that teachers are increasingly facing in their work, and the results of this study have convinced me that research into these policies is not only important, but also urgent. Other relevant policies include anti-collective bargaining
movements, teacher merit pay policies, charter schooling, and numerous other related policies. There is currently no shortage of policies that have great potential to influence teacher and student motivation in the classroom, and this study, while informative, has barely scratched the surface of the work that needs to be completed. While on job interview at Virginia Commonwealth University one professor asked me, “What will happen to your work if the testing policies and No Child Left Behind are overturned?” to which I responded, “There is always going to be a policy, and the policy is always going to manifest in the classroom.” This is the nature of my work—examining policy at the intersection of motivation in the classroom. The absence of one bad policy will not be a career-ender for me.

The results from this study clearly indicated that stress due to tests is having some influence on teachers’ self-efficacy leading up to the test. This finding raises questions about how teacher stress and teacher self-efficacy really behave in relation to one another, questions about teachers’ self-efficacy over the course of a school year, and questions about how stress influences teachers’ self-efficacy under varying conditions. Future work will include longitudinal studies that examine teachers’ motivational beliefs over the course of a school year to look for patterns of trajectory—how soon do teachers’ self-efficacy beliefs begin to decline in the time before testing? How does the reduction in teacher’ self-efficacy influence students? What does this say about quality teaching? How important is administrative support in relation to teacher self-efficacy in the context of high-stakes tests? What is the condition of administrative self-efficacy in light of high-stakes testing policies? How do we prepare teachers for conditions that are hostile to teacher self-efficacy, such as a teacher who will be employed in a school district that
bases 51% of its’ teacher evaluations on student test scores, and hasn’t passed a high-stakes test in five years?

Of particular interest (and annoyance) to me moving forward is designing an adequate measurement tool to assess teachers’ autonomy-supportiveness and controllingness. The implications of autonomy-support and control in the classroom are significant; it seems that there ought to exist a quality measure that is less than 30 years old. In my immediate future, I plan to examine the construct and/or constructs of autonomy-support and control, and their measurement.

From there, I intend to examine autonomy-support and control in relation to education policies; of particular importance are teacher evaluation policies that seem to have a large effect on teacher autonomy-support and control. It may be necessary to examine these issues from a comparative perspective, as current conditions in the United States seem to limit the opportunity to study the issues in classrooms that are not controlled. Possible study designs include comparative work outside the U.S.; or comparative studies with schools in the U.S. that do not utilize such conditions, such as private or Montessori schools. Comparing institutions such as these in differing countries or societies presents challenges; however, I consider SDT to be a challenge itself, so it will be expected.

**Conclusion**

The purpose of this study was to examine the relations between teachers’ perceptions of stress as a result of high-stakes tests and teachers’ motivational beliefs. Motivational beliefs were operationalized using two prolific motivational theories: Social Cognitive Theory (teacher self-efficacy) and Self-Determination Theory (autonomy-
Over five-hundred teachers were given a series of measures to complete during the Summer and Autumn of 2011; these measures included a two-part stress measure, the Teacher Self-Efficacy Scale in which each item was administered twice; the Problems in Schools Questionnaire, and the Teacher Autonomy and Control Scale, in addition to a series of demographic questions. Multiple regression analyses were conducted to determine if the stress perceived by teachers predicted either of the motivational constructs.

Results indicate that teachers perceive great levels of stress as a result of high-stakes tests, and that this stress is related to their motivation as defined by Social Cognitive Theory (teacher self-efficacy). Stress was a strong predictor of teacher self-efficacy in the weeks leading up to the testing window; throughout the school year the strongest predictors of teacher self-efficacy were teacher income, teachers’ perceptions of administrative support, and teacher level of experience in the classroom. Teachers’ self-efficacy across the two time points also seems to shift from source to source, as teacher experience level was a significant predictor of teacher self-efficacy throughout the school year, but not in the weeks leading up to the test; additionally, the likelihood that the school would pass the test became a significant predictor of teacher self-efficacy in the weeks leading up to the test. Stress did not predict teacher autonomy-supportiveness or controllingness. Teacher evaluation policies were found to be the only predictor of teacher autonomy-supportiveness or controllingness, and the effect size was quite large.

Measuring teacher autonomy-supportiveness and controllingness was challenging; there is much to be explored in this area. Teacher autonomy-supportiveness and controllingness within the context of teacher evaluation policies appears to be an area that
is ripe for further study; extrapolating further, the relationships between teacher
autonomy and student autonomy as a result of these policies is an area that warrants
further examination as well. Teacher self-efficacy appears to be suffering as a result of
the high-stakes testing context; future work will attempt to investigate how teacher self-
efficacy can be bolstered or preserved in the classroom in the face of extenuating
challenges that are out of teachers’ control.
References


146
Foucault, M. (1979). Michel Foucault’s lecture at the College de France on neo-liberal

Gable, S.L. (2006). Approach and avoidance social motives and goals. *Journal of
Personality, 74*, 175-222.

Geving, A.M. (2007). Identifying the types of student and teacher behaviours associated

Educational Psychology, 76*(4), 569.

teachers on efficacy, ego development, and problem solving. *Psychology in
Schools, 19*, 558-562.

Relationships to motivation and achievement. *Journal of Early Adolescence, 13*,
21-43.

efficacy belief patterns and selected teacher characteristics. *Journal of Research


cites-test-score-errors.html.


Hazel, C. (2010). Interactions between bullying and high-stakes testing at the elementary

learning of urban minority students in a high-stakes testing context. *Educational


McMillan, 2000 http://www.cepi.vcu.edu/policy_issues/saa/high_stakes.html#


Reeve, J. (2009). Why teachers adopt a controlling motivating style toward students and how they can become more autonomy supportive. *Educational Psychologist, 44*(3), 159-175.


Ryan, 2011 (personal communication)


United States Department of Education, Office of the Secretary, 2003


Weinstein, N., & Hodgins, H.S. (2009). The moderating role of motivation for written


Appendix A
Dear Teachers,

My name is Heather Dawson. I’m a Ph.D. Candidate at The Ohio State University. You are receiving this email because you are a teacher employed in a public school, and I would like to ask for your participation in an important research study.

No Child Left Behind (2002) ushered in the era of high-stakes testing in American schools. Although the ’02 NCLB law did not introduce high-stakes testing to schools, it mandated the use of them, proliferating their use around the country. In many schools, graduation decisions, teacher employment and salary decisions, and school-wide or district-wide funding decisions are made based on student scores on these tests.

My work seeks to examine the effect of these tests on teacher and student motivation for learning. It is important to find out if these tests have any effect on teachers’ or students’ motivation in order to improve the conditions in the classroom for all. This is why I request your participation—I would like your help to begin answering these questions by clicking on the survey link here:


Your participation is confidential. You do not need to provide your name or the name of your school district—your answers will not be connected to you in any way. This survey will take a short amount of time (less than 20 minutes). In exchange for your participation, I would invite you to submit your contact information to highstakestests@gmail.com to be entered for a chance to win one of 5 Barnes and Noble gift cards OR one Apple iPad. You need only enter your contact information one time to be entered for all the prizes. I have separated the survey from the incentive email register to give added confidentiality.

If you are interested in participating, have any questions about the study, or if you decide to participate in the study and feel harmed as a result of participation, please contact me (Heather Dawson, dawson.282@osu.edu). For questions about your rights as a participant in this study, or to discuss other study-related concerns or complaints with someone who is not part of the research team, please contact Ms. Sandra Meadows in the Office of Responsible Research Practices at 1-800-678-6251.

Additionally, if you would like to assist with this study in addition to your own participation, please forward this message to your peers and colleagues that are public school teachers.

Thank you for your time!

--

Heather S. Dawson
Ph.D. Candidate
Educational Psychology and Philosophy
The Ohio State University
Department of Educational Policy and Leadership
29 W. Woodruff Ave.
Columbus, OH 43210
dawson.282@osu.edu
(00.1) 419.280.1069
Appendix B
High Stakes Testing

First, about you...

1) With which gender do you most closely identify?
   ( ) Male
   ( ) Female

2) With which race or ethnicity do you most closely identify?
   ( ) White
   ( ) Black
   ( ) Asian
   ( ) Native American/Pacific Islander
   ( ) Hispanic
   ( ) Arab

3) How many years of teaching experience do you have?
   ( ) 1
   ( ) 2
   ( ) 3
   ( ) 4
   ( ) 5
   ( ) 6
   ( ) 7
   ( ) 8
   ( ) 9
   ( ) 10
   ( ) 11
   ( ) 12
   ( ) 13
   ( ) 14
   ( ) 15
   ( ) 16-20
   ( ) 21-30
   ( ) Over 30

4) How does your school usually score on high-stakes tests?
   ( ) Below the appropriate level (e.g., we are a "failing" school or in "continuous improvement").
   ( ) Right about at the appropriate level.
   ( ) My school generally performs well on the standardized tests.
5) Which content areas do you primarily teach?
[ ] Tested content areas (math, reading).
[ ] Non-tested content areas.
[ ] Both.
[ ] Neither.

6) What grades do you teach?
[ ] K
[ ] 1
[ ] 2
[ ] 3
[ ] 4
[ ] 5
[ ] 6
[ ] 7
[ ] 8
[ ] 9
[ ] 10
[ ] 11
[ ] 12
[ ] Special Education/Resource

7) What is the highest level of education you have achieved?
( ) Associate's degree
( ) Bachelor's degree
( ) Master's degree
( ) Ph.D.

8) In what state do you currently teach?
( ) Alabama
( ) Alaska
( ) Arizona
( ) Arkansas
( ) California
( ) Colorado
( ) Connecticut
( ) Delaware
( ) District of Columbia
( ) Florida
( ) Georgia
( ) Hawaii
( ) Idaho
( ) Illinois
( ) Indiana
( ) Iowa
Kansas
Kentucky
Louisiana
Maine
Maryland
Massachusetts
Michigan
Minnesota
Mississippi
Missouri
Montana
Nebraska
Nevada
New Hampshire
New Jersey
New Mexico
New York
North Carolina
North Dakota
Ohio
Oklahoma
Oregon
Pennsylvania
Rhode Island
South Carolina
South Dakota
Tennessee
Texas
Utah
Vermont
Virginia
Washington
West Virginia
Wisconsin
Wyoming

9) What kind of school do you teach in?
[ ] Public
[ ] Private
[ ] Charter
10) How much are you paid for teaching?
( ) <$30,000
( ) $30,001-$40,000
( ) $40,001-$50,000
( ) $50,001-$60,000
( ) $60,001-$70,000
( ) $70,001-$80,000
( ) $80,001-$90,000
( ) $90,001-$100,000
( ) > $100,000

11) In what kind of neighborhood is the school that you teach in located?
( ) Urban/inner city
( ) Suburban
( ) Rural

12) If you had to guess, how many students in your school receive free and/or reduced lunch?
( ) >10%
( ) 11%-20%
( ) 21%-30%
( ) 31%-40%
( ) 41%-50%
( ) 51%-60%
( ) 61%-70%
( ) 71%-80%
( ) 81%-90%
( ) 91%-100%

13) What is your average class size?
[] Under 20
[] 21-25
[] 26-35
[] Above 35

14) With what political persuasion do you most closely identify?
( ) Very liberal
( ) Liberal
( ) Moderate
( ) Conservative
( ) Very Conservative
15) How much of your evaluation as a teacher depends on your students' test scores?
( ) None
( ) Some of my evaluation depends on student performance on tests, but not a lot.
( ) Nearly half of my evaluation depends on student performance on tests.
( ) Over half of my evaluation depends on student performance on tests.

16) The administration in my school is supportive of us as teachers.
( ) Strongly agree
( ) Agree
( ) Neutral
( ) Disagree
( ) Strongly disagree

How much do you agree with the following statements?

For each odd-numbered question, consider the extent to which an event is happening in your school or classroom. For each even-numbered item, consider the extent to which the event causes you stress.

17) In my school, I feel pressure to improve standardized test scores.
( ) Strongly agree
( ) Agree somewhat
( ) Neutral
( ) Disagree
( ) Strongly Disagree

18) This causes me stress.
( ) Strongly agree
( ) Agree
( ) Neutral
( ) Disagree
( ) Strongly disagree

19) In my school, tested content areas are often treated as more important than non-tested content areas.
( ) Strongly agree
( ) Agree somewhat
( ) Neutral
( ) Disagree
( ) Strongly Disagree
20) This causes me stress.
   ( ) Strongly agree
   ( ) Agree
   ( ) Neutral
   ( ) Disagree
   ( ) Strongly disagree

21) In my school, I feel pressured to make sure that my students pass the test.
   ( ) Strongly agree
   ( ) Agree somewhat
   ( ) Neutral
   ( ) Disagree
   ( ) Strongly Disagree

22) This causes me stress.
   ( ) Strongly agree
   ( ) Agree
   ( ) Neutral
   ( ) Disagree
   ( ) Strongly disagree

23) In my school, there is increased pressure in the weeks leading up to the test.
   ( ) Strongly agree
   ( ) Agree somewhat
   ( ) Neutral
   ( ) Disagree
   ( ) Strongly Disagree

24) This causes me stress.
   ( ) Strongly agree
   ( ) Agree
   ( ) Neutral
   ( ) Disagree
   ( ) Strongly disagree

25) In my school, using practice standardized tests is common.
   ( ) Strongly agree
   ( ) Agree somewhat
   ( ) Neutral
   ( ) Disagree
   ( ) Strongly Disagree
26) This causes me stress.
   ( ) Strongly agree
   ( ) Agree
   ( ) Neutral
   ( ) Disagree
   ( ) Strongly disagree

27) In my school, there is pressure to maintain or improve our image to the public.
   ( ) Strongly agree
   ( ) Agree somewhat
   ( ) Neutral
   ( ) Disagree
   ( ) Strongly Disagree

28) This causes me stress.
   ( ) Strongly agree
   ( ) Agree
   ( ) Neutral
   ( ) Disagree
   ( ) Strongly disagree

29) In my school, there is a sense of relief when the tests are over.
   ( ) Strongly agree
   ( ) Agree somewhat
   ( ) Neutral
   ( ) Disagree
   ( ) Strongly Disagree

30) This reduces my stress.
   ( ) Strongly agree
   ( ) Agree
   ( ) Neutral
   ( ) Disagree
   ( ) Strongly disagree

31) In my school, tensions rise as testing time approaches.
   ( ) Strongly agree
   ( ) Agree somewhat
   ( ) Neutral
   ( ) Disagree
   ( ) Strongly Disagree

32) This causes me stress.
   ( ) Strongly agree
   ( ) Agree
33) Is there anything you would like to add or say about these statements?
____________________________________________

These questions ask about your ability to complete teaching-related tasks. Thanks for your patience--I know this is a lot of questions..

34) To what extent can you use a variety of assessment strategies **generally during the school year**?
( ) I can do nothing
( ) I can do little
( ) I can do some
( ) I can do much
( ) I can do a lot

35) To what extent can you use a variety of assessment strategies **in the weeks leading up to the high-stakes test**?
( ) I can do nothing
( ) I can do little
( ) I can do some
( ) I can do much
( ) I can do a lot

36) To what extent can you craft good questions for your students **generally throughout the school year**?
( ) I can do nothing
( ) I can do little
( ) I can do some
( ) I can do much
( ) I can do a lot

37) To what extent can you craft good questions for your students **in the weeks leading up to the high-stakes test**?
( ) I can do nothing
( ) I can do little
( ) I can do some
( ) I can do much
( ) I can do a lot
38) To what extent can you provide an alternative explanation or example when students are confused, generally throughout the school year?
( ) I can do nothing
( ) I can do little
( ) I can do some
( ) I can do much
( ) I can do a lot

39) To what extent can you provide an alternative explanation or example when students are confused, in the weeks leading up to the high stakes test?
( ) I can do nothing
( ) I can do little
( ) I can do some
( ) I can do much
( ) I can do a lot

40) How well can you implement alternative strategies in your classroom generally throughout the school year?
( ) I can do nothing
( ) I can do little
( ) I can do some
( ) I can do much
( ) I can do a lot

41) How well can you implement alternative strategies in your classroom in the weeks leading up to the high-stakes test?
( ) I can do nothing
( ) I can do little
( ) I can do some
( ) I can do much
( ) I can do a lot

42) How much can you do to control disruptive behavior in your classroom, generally throughout the school year?
( ) I can do nothing
( ) I can do little
( ) I can do some
( ) I can do much
( ) I can do a lot

43) How much can you do to control disruptive behavior in your classroom, in the weeks leading up to the high-stakes test?
( ) I can do nothing
( ) I can do little
( ) I can do some
( ) I can do much
( ) I can do a lot

44) How much can you do to get children to follow classroom rules, generally throughout the school year?
( ) I can do nothing
( ) I can do little
( ) I can do some
( ) I can do much
( ) I can do a lot

45) How much can you do to get children to follow classroom rules in the weeks leading up to the high-stakes test?
( ) I can do nothing
( ) I can do little
( ) I can do some
( ) I can do much
( ) I can do a lot

46) How much can you do to calm a student who is disruptive or noisy, generally throughout the school year?
( ) I can do nothing
( ) I can do little
( ) I can do some
( ) I can do much
( ) I can do a lot

47) How much can you do to calm a student who is disruptive or noisy in the weeks leading up to the high-stakes test?
( ) I can do nothing
( ) I can do little
( ) I can do some
( ) I can do much
( ) I can do a lot

48) How well can you establish a classroom management system with your students, generally, throughout the school year?
( ) I can do nothing
( ) I can do little
( ) I can do some
( ) I can do much
( ) I can do a lot
49) How well can you establish a classroom management system with your students in the weeks leading up to the high-stakes test?
( ) I can do nothing
( ) I can do little
( ) I can do some
( ) I can do much
( ) I can do a lot

50) How much can you do to motivate students who show low interest in schoolwork generally throughout the school year?
( ) I can do nothing
( ) I can do little
( ) I can do some
( ) I can do much
( ) I can do a lot

51) How much can you do to motivate students who show low interest in schoolwork in the weeks leading up to the high-stakes test?
( ) I can do nothing
( ) I can do little
( ) I can do some
( ) I can do much
( ) I can do a lot

52) How much can you do to get students to believe they can do well in schoolwork generally throughout the school year?
( ) I can do nothing
( ) I can do little
( ) I can do some
( ) I can do much
( ) I can do a lot

53) How much can you do to get students to believe they can do well in schoolwork in the weeks leading up to the high-stakes test?
( ) I can do nothing
( ) I can do little
( ) I can do some
( ) I can do much
( ) I can do a lot

54) How much can you assist families in helping their children do well in school generally throughout the school year?
( ) I can do nothing
( ) I can do little
( ) I can do some
( ) I can do much
( ) I can do a lot

55) How much can you assist families in helping their children do well in school in the weeks leading up to the high-stakes test?
( ) I can do nothing
( ) I can do little
( ) I can do some
( ) I can do much
( ) I can do a lot

56) How much can you do to help your students value learning generally throughout the school year?
( ) I can do nothing
( ) I can do little
( ) I can do some
( ) I can do much
( ) I can do a lot

57) How much can you do to help your students value learning in the weeks leading up to the high-stakes test?
( ) I can do nothing
( ) I can do little
( ) I can do some
( ) I can do much
( ) I can do a lot

58) Is there anything you would like to add or say about these questions?

What do your daily classroom practices entail?

If you just hang tight, the end is near. These questions are so important to teachers, and haven't really been answered before in the literature. You're making a huge contribution here!

59) I offer my students rewards for good behavior.
( ) Strongly agree
( ) Agree
( ) Neutral
( ) Disagree
( ) Strongly disagree
60) I generally give my students feedback on their work as quickly as possible.
( ) Strongly agree
( ) Agree
( ) Neutral
( ) Disagree
( ) Strongly disagree

61) I prefer to have my students learn to solve problems the way I have taught them.
( ) Strongly agree
( ) Agree
( ) Neutral
( ) Disagree
( ) Strongly disagree

62) I expect my students to follow the directions that I explain to them.
( ) Strongly agree
( ) Agree
( ) Neutral
( ) Disagree
( ) Strongly disagree

63) I always try to understand the students' perspectives when I am working through difficult problems with them.
( ) Strongly agree
( ) Agree
( ) Neutral
( ) Disagree
( ) Strongly disagree

64) I make it clear to my students that there are negative consequences for failing in school.
( ) Strongly agree
( ) Agree
( ) Neutral
( ) Disagree
( ) Strongly disagree

65) I make my presence visible to my students as they work.
( ) Strongly agree
( ) Agree
( ) Neutral
( ) Disagree
( ) Strongly disagree

66) I include my students in the process of designing lessons.
67) I am flexible with deadlines for projects.
( ) Strongly agree
( ) Agree
( ) Neutral
( ) Disagree
( ) Strongly disagree

68) I usually give my students a choice when assigning projects.
( ) Strongly agree
( ) Agree
( ) Neutral
( ) Disagree
( ) Strongly disagree

69) I encourage my students to set their own goals and deadlines.
( ) Strongly agree
( ) Agree
( ) Neutral
( ) Disagree
( ) Strongly disagree

70) I include my students in the process of designing rules for my classroom.
( ) Strongly agree
( ) Agree
( ) Neutral
( ) Disagree
( ) Strongly disagree

71) Jim is an average student who has been working at grade level. During the past two weeks he has appeared listless and has not been participating during reading group. The work he does is accurate but he has not been completing assignments. A phone conversation with his mother revealed no useful information. The most appropriate thing for Jim's teacher to do is:

<table>
<thead>
<tr>
<th>Very inappropriate</th>
<th>Moderately appropriate</th>
<th>Very appropriate</th>
</tr>
</thead>
</table>
She should impress upon ( ) ( ) ( ) ( ) him the importance of
finishing his assignments since he needs to learn this material for his own good. Let him know that he doesn't have to finish all of his work now and see if she can help him work out the cause of the listlessness. Make him stay after school until that day's assignments are done. Let him see how he compares with the other children in terms of his assignments and encourage him to catch up with the others.

72) At a parent conference last night, Mr. and Mrs. Greene were told that their daughter Sarah has made more progress than expected since the time of the last conference. All agree that they hope she continues to improve so that she does not have to repeat the grade (which the Greene's have been kind of expecting since the last report card). As a result of the conference, the Greenes decide to:

<table>
<thead>
<tr>
<th>Activity</th>
<th>Very inappropriate</th>
<th>Moderately appropriate</th>
<th>Very appropriate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Increase her allowance and promise her a ten-speed if she continues to improve.</td>
<td>( )</td>
<td>( )</td>
<td>( )</td>
</tr>
<tr>
<td>Tell her that she's now doing as well as many of the other children in her class.</td>
<td>( )</td>
<td>( )</td>
<td>( )</td>
</tr>
<tr>
<td>Tell her about the report, ( ) letting her know that they're aware of her increased independence in school and at home.</td>
<td>( )</td>
<td>( )</td>
<td>( )</td>
</tr>
<tr>
<td>Continue to emphasize that she has to work hard to get better grades.</td>
<td>( )</td>
<td>( )</td>
<td>( )</td>
</tr>
</tbody>
</table>
73) Dylan loses his temper a lot and has a way of agitating other children. He doesn't respond well to what you tell him to do and you're concerned that he won't learn the social skills he needs. The best thing for you to do with him is:

<table>
<thead>
<tr>
<th></th>
<th>Very inappropriate</th>
<th>Moderately appropriate</th>
<th>Very appropriate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Emphasize how important it is for him to &quot;control himself&quot; in order to succeed in school and in other situations.</td>
<td>()</td>
<td>()</td>
<td>()</td>
</tr>
<tr>
<td>Put him in a special class which has the structure and reward contingencies which he needs.</td>
<td>()</td>
<td>()</td>
<td>()</td>
</tr>
<tr>
<td>Help him see how other children behave in these various situations and praise him for doing the same.</td>
<td>()</td>
<td>()</td>
<td>()</td>
</tr>
<tr>
<td>Realize that Dylan is probably not getting the attention he needs and start being more responsive to him.</td>
<td>()</td>
<td>()</td>
<td>()</td>
</tr>
</tbody>
</table>

74) The Rangers spelling group has been having trouble all year. How could Miss Wilson best help the Rangers?

<table>
<thead>
<tr>
<th></th>
<th>Very inappropriate</th>
<th>Moderately appropriate</th>
<th>Very appropriate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Have regular spelling bees so that Rangers will be motivated to do as well as the other groups.</td>
<td>()</td>
<td>()</td>
<td>()</td>
</tr>
<tr>
<td>Make them drill more and give them special privileges for improvements.</td>
<td>()</td>
<td>()</td>
<td>()</td>
</tr>
<tr>
<td>Have each child keep a spelling chart and emphasize how important it is to have a good chart.</td>
<td>()</td>
<td>()</td>
<td>()</td>
</tr>
</tbody>
</table>
Help the group devise ways of learning the words together (skits, games, and so on).

75) In your class is a girl named Maddie who has been the butt of jokes for years. She is quiet and usually alone. In spite of the efforts of previous teachers, Maddie has not been accepted by the other children. Your wisdom would guide you to:

<table>
<thead>
<tr>
<th></th>
<th>Very inappropriate</th>
<th>Moderately appropriate</th>
<th>Very appropriate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prod her into interactions and provide her with much praise for any social initiative.</td>
<td>()</td>
<td>()</td>
<td>()</td>
</tr>
<tr>
<td>Talk to her and emphasize that she should make friends so she'll be happier.</td>
<td>()</td>
<td>()</td>
<td>()</td>
</tr>
<tr>
<td>Invite her to talk about her relations with the other kids, and encourage her to take small steps when she's ready.</td>
<td>()</td>
<td>()</td>
<td>()</td>
</tr>
<tr>
<td>Encourage her to observe how other children relate and to join in with them.</td>
<td>()</td>
<td>()</td>
<td>()</td>
</tr>
</tbody>
</table>

76) For the past few weeks things have been disappearing from the teacher's desk and lunch money has been taken from some of the children's desks. Today, Rachel was seen by the teacher taking a silver dollar paperweight from her desk. The teacher phoned Rachel's mother and spoke to her about this incident. Although the teacher suspects that Rachel has been responsible for the other thefts, she mentioned only the one and assured the mother that she'll keep a close eye on Rachel. The best thing for the mother to do is:

<table>
<thead>
<tr>
<th></th>
<th>Very inappropriate</th>
<th>Moderately appropriate</th>
<th>Very appropriate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Talk to her about the consequences of stealing and what it would mean in relation to the other kids.</td>
<td>()</td>
<td>()</td>
<td>()</td>
</tr>
<tr>
<td>Talk to her about it, expressing her</td>
<td>()</td>
<td>()</td>
<td>()</td>
</tr>
</tbody>
</table>
confidence in her and attempting to understand why she did it. Give her a good scolding; stealing is something which cannot be tolerated and she has to learn that. Emphasize that it was wrong and have her apologize to the teacher and promise not to do it again.

77) Your student has been getting average grades, and you'd like to see her improve. A useful approach, in your opinion, might be to:

<table>
<thead>
<tr>
<th>Very inappropriate</th>
<th>Moderately appropriate</th>
<th>Very appropriate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Encourage her to talk about her report card and what it means for her.</td>
<td>()</td>
<td>()</td>
</tr>
<tr>
<td>Go over the report card with her; point out where she stands in the class.</td>
<td>()</td>
<td>()</td>
</tr>
<tr>
<td>Stress that she should do better; she'll never get into college with grades like these.</td>
<td>()</td>
<td>()</td>
</tr>
<tr>
<td>Suggest her parents offer her a dollar for every A and 50 cents for every B on future report cards.</td>
<td>()</td>
<td>()</td>
</tr>
</tbody>
</table>

Thank You!
Thank you for taking our survey. Your response is very important to us.