An Integrated Mixed-Methods Study of Contract Grading's Impact on Students' Perceptions of Stress, Self-Worth Protection Behaviors, and Academic Performance in High School English Courses.

Emily J. Watson (Ward)

BA, MA

This thesis is submitted in partial fulfillment of the requirements

for the degree of Doctor of Philosophy.

Department of Educational Research,

Lancaster University, UK.

May 2021

I declare that this thesis is 44,244 words and does not exceed the maximum.

This thesis results entirely from my own work and has not been offered previously for any other degree or diploma.

Signature Emily J. Watson (Ward)

Abstract

Background. Contract grading is a holistic assessment approach for learning and grading in which students participate in the assessment process by choosing their desired effort and outcome.

Aims. This thesis sought to understand the impact of contract grading on high schoolers' (grades 9-12) perceptions of stress, self-worth protection behaviors, achievement motivation, and academic performance.

Sample. Participants were 439 young people (ages 13-19), including 284 returning students and 155 first-year students, completing a high-stakes writing assessment in their English course.

Method. The integrated mixed-methods study followed an explanatory sequential design: interviews with 40 adolescents from all grade levels and course types explained the findings of matched-pairs quantitative data generated from four psychometrically sound scales.

Results.

> **Academic Stress**. The contract significantly reduced evaluative threat by clarifying expectations. Consequently, compared to their prior experience with or expectations for the task, adolescents perceived workload demands as significantly less stressful and threatening under the contract.
>
> **Self-Worth Protection Behaviors**. Adolescents reported significantly less fear of failure and social comparison. The qualitative data revealed that the contract oriented them to achieve success, rather than avoid failure.
>
> **Academic Achievement.** 90% (n=390) of participants fulfilled the contract to earn either an A or B, including 94% of returning students and 83% of first-year students. Students with prior experience were 19% more likely to earn an A and 16% more likely to earn a B under the contract compared to conventional grading practices. Overall, 97% (n=421) earned a passing grade on the assessment.

Contribution. The findings of this thesis make a significant contribution, revealing that contract grading reduced secondary students' perceptions of stress and evaluative threat, oriented them toward success, and improved academic performance. This leads to a call to action for secondary teachers to implement contract grading to foster psycho-emotionally healthy learning environments.

Table of Contents

Acknowledgments

Completing a Ph.D., particularly while working full-time as a high school teacher, requires the support, kindness, and encouragement of many whom I am obliged to thank here.

First, I could not have conducted this research without my English Department colleagues' willingness to rethink cherished grading systems and my school administration's full support of this research. I am thankful for each person who engaged in a year-long conversation about secondary writing assessment and then stepped out of their comfort zones to implement a new approach. I am especially grateful for my department chairs, who piloted contract grading in their courses in advance of this research; innovative and open-minded, they served as crucial risk-takers and thought-partners for this work.

I am also beholden to Jesse, who helped me proofread hundreds of pages, sometimes three or four times, comforted me after reviewer #2's painful stabs, engaged in endless conversations about education, and encouraged me to rest. Also, while I wrote this thesis, our beloved black cat, Salem, spent many hours on my lap and keyboard, making many special edits that, unfortunately, had to be deleted.

I must also express deep gratitude to my mom, Cheri, who birthed the fifth living generation of women in our family. A hard-working, single mother, she encouraged my love of reading, always making room in our budget for books from the Scholastic catalog, and my dogged curiosity, always asking me to "do some research" when I had questions. In the '90s, she purchased a Sony computer, a life-changing technology that gave me access to new worlds and the digital literacy skills that allowed me to graduate from college and then complete a Ph.D. from my kitchen table. I recognize, as Rawls observed, that my success is due to my starting position and morally arbitrary factors, particularly the tenacity and strength I inherited from the women who came before me.

I am also thankful for the positive experience I had as an American student studying in the Department of Educational Research at Lancaster University, and I am particularly grateful for the supervision of Carolyn Jackson, who is not only brilliant but kind.

Publications derived from work on the Doctoral Programme

Ward, E. (2021). An integrated mixed-methods study of contract grading's impact on adolescents' perceptions of stress in high school English: A pilot study. *Assessing Writing, 48*, 100508.

Ward, E. (2021). Easing stress: Contract grading's impact on adolescents' perceptions of workload demands, time constraints, and challenge appraisal in high school English. *Assessing Writing, 48*, 100526.

List of abbreviations

BERA	British Educational Research Association
GPA	Grade Point Average
HS	High School (grades 9-12)
MLA	Modern Language Association
ISO	International Organization for Standardization
PASS	Perceptions of Academic Stress Scale
PASA	Primary Appraisal Secondary Appraisal
SAM	Stress Appraisal Measure
SLD	Specific Learning Disability
SWPS	Self-Worth Protection Scale

List of Tables and Figures

Chapter 1: Introduction

1.1 Why research an alternative writing assessment approach?

Recent negative trends in the mental health of adolescents and young adults should concern not only counselors and medical professionals but teachers and administrators. To researcher Twenge (2017), who studies Generation Z or iGen, the cohort born between 1995 and the mid-2000s, "it is not an exaggeration to describe iGen as being on the brink of the worst mental health crisis in decades" (n.p). Between 2007 and 2015, the number of children and teens checking into emergency rooms for suicidal thoughts and attempts doubled in the United States (Burstein, Agostino, & Greenfield, 2019). Gene Beresin of The Clay Center for Young Healthy Minds at Massachusetts General Hospital observed that "kids are feeling more pressure to achieve, more pressure in school, and are more worried about making a living than in previous years" (Bracho-Sanchez, 2019, n.p.). In this observation, the pressures and stresses of school—particularly the neoliberalization of education and the marketization of knowledge, in which schools are engines of social mobility (McArthur, 2016)—sit at the heart of declining mental health and well-being in students.

This pressure is pervasive at Good Shepherd High School (GSHS, anonymized), a high-achieving private, religious institution in an upper-middle-class community county in the Western United States, where I have taught in the English Department since 2013. While Luthar and Latendresse (2005) recognize that adolescents from affluent, suburban neighborhoods are considered low risk for psycho-social challenges, they observe that

achievement pressure is prevalent among affluent families, which intensifies the impact of school demands (Ainslie, Shafer, & Reynolds, 1996), and thus it would be "wise to remain cognizant that those ostensibly most privileged can also confront substantial threats to their psychological well-being" (p. 117). Luthar, Barkin, and Crossman (2013) call affluent youth 'privileged but pressured,' observing that their statistical odds for developing serious maladjustment place them at-risk. Often over-scheduled in competitive lifestyles, these adolescents report higher levels of stress and substance use than peers from modest economic backgrounds (Luthar & Latendresse, 2005). Even when aware that their child is struggling, affluent parents are less likely to seek professional help to "maintain a veneer of well-being, feeling that 'those at the top are supposed to be better able to handle their problems than those further down the scale," according to Wolfe and Fodor (1996) (p. 80). Unfortunately, substance abuse in affluent youth is correlated with depression and anxiety, which suggests an effort to self-medicate, as Luthar and Latendresse (2005) write:

> Practitioners and parents must be alert to the risks potentially attached to wealth and status. The American dream spawns widespread beliefs that Ivy League educations and subsequently lucrative careers are critical for children's long-term happiness. In the sometimes single-minded pursuit of these goals, let us not lose sight of the possible costs to mental health and well-being of all concerned. (p. 52)

In 2015, two GSHS graduates committed suicide during their first year in college. In 2018, 16-year-old Patrick Turner of a nearby school wrote in his suicide letter that the pressure to succeed and subsequent fear of failure drove him to take his own life: "One slipup makes a kid feel like the smallest person in the world. You are looked at as a loser

if you don't go to college or if you get a certain GPA or test score. . . . So much pressure is placed on the students to do well that I couldn't do it anymore" (Jaschik, 2018). To the community, Dr. Bolton, Turner's principal, reflected on the deleterious effects of a "competitive culture" obsessed with grades, grade point averages (GPAs), and rankings:

> Our teachers and District have simply created and maintained a system that our community/country has demanded from us over the past 20 years since college admissions mania went into hyperdrive, since vocational training programs were dismantled, and since earning A's in AP classes became the norm. Our teachers feel the pressure, administration and counseling feel the pressure, and now parents/students are really feeling the pressures. . . . We endlessly discuss test scores, National Merit Scholarships, reading scores, AP scholars, comparisons to other school districts and this is when we start losing our collective souls—and our children. (Jaschik, 2018).

Even twenty years ago, de Anda et al. (2000) found that the greatest stressors for high school students were "expectations about career," "future life plans," "tests," and "grades." Even earlier, Ainslie, Shafer, and Reynolds (1996) found that adolescents who felt it was most important to earn high grades also experienced the highest levels of stress. Recently, Inoue (2019) observed that "grades exert immense pressure on students when made more present in the course" (p. 142). To Elbow (1997), secondary students may experience the most pressure since "grades [are] central to the college admissions process" (p. 6). With their far-reaching consequences, marks on assignments, scores on standardized tests, cumulative GPAs, report cards, and transcripts amplify the stress of school.

Recent qualitative work, however, reveals a therapeutic turn toward individualizing stress as a personal pathology, seeking to help students modulate the effects of stress: a

mindfulness-based stress reduction program (Hjeltnes, Binder, Moltu, & Dundas (2015); emotional freedom technique (Patterson, 2016); and internet- and app-based stress programs (Fleischmann et al., 2018); and cognitive-behavioral stress management (Terp, Bisholt, & Hjärthag, 2019). In 2018, GSHS also added a yoga course with mindfulness meditation, trained a therapy dog for campus, and implemented a social-emotional learning (SEL) small-group to help students manage their emotions; however, ameliorating student stress with yoga and mindfulness training is, as Purser (2019) writes, a "market-friendly palliative" that addresses the symptoms of suffering without striking at the social causes (p. 26). This thesis is concerned with a pervasive force that has bred stress, anxiety, and fear of failure in even the youngest learners: the culture of high-stakes assessment. While work has examined the benefit of optimizing students' responses in the face of stressors (Jamieson, Peters, Greenwood, & Altose, 2016; Jamieson, Crum, Goyer, Marotta, & Akinola, 2018), this thesis research challenges the status quo by tackling the stressors.

To mitigate evaluation pressure by emphasizing learning over performance, minimize the bias that seeps into grading, and help all students reach their highest potential, the GSHS English Department implemented an alternative assessment system called contract grading. Most simply, the contract system contrasts with conventional grading practices by outlining the performance criteria—that is, the specific actions and behaviors that are required to earn each grade—and then invites students to participate in their assessment by choosing the criteria that correspond with their desired goal, effort, and energy

expenditure. In this way, contract grading promotes self-regulated learning, in which students set and monitor goals. To encourage academic risk-taking and frame struggle as central to the learning process, the contract system differs from conventional grading practices by limiting the influence of teacher bias on grades. Contract grading is uncommon in US schools: of the 12 teachers in this study who implemented contract grading in their course, none had learned about it in their teacher education programs. As evidenced by recent research (Spidell & Thelin, 2006; Potts, 2010; Lindemann & Harbke, 2011; Litterio, 2016; Litterio, 2018) and articles in *Inside Higher Ed* ("I Have Seen the Glories of the Grading Contract…" (2016) and "Contract Grading Success, At Last At Last" (2017) by John Warner), contract grading may be used more in US college classrooms where instructors often have more autonomy to implement alternative approaches, but empirical research is needed on contract grading's impact on adolescents at college-preparatory high schools, where they face substantial threats to their well-being (Galloway, Conner, & Pope, 2013; Feld & Shusterman, 2015).

While the conventional grading system often breeds frustration, stress, and writer's block, research with college students reveals that contract grading is an accurate writing assessment tool (Potts, 2010) that can reduce perceptions of academic stress (Fairbanks, 1992). Smith and Lerch (1972) assert that the promise of a good grade upon fulfilling the contract may alleviate the threat and stigma of failure, thereby ameliorating the source of student stress. While recent work points to its efficacy in contemporary college classrooms (Potts, 2010; Lindemann & Harbke, 2011; Litterio, 2016; Litterio, 2018),

most empirical work with contract grading has been relegated to the 1970s, limited in size and scope, and situated in college classrooms. While I was writing this thesis, Cowan (2020) observed that "sadly, no research has been published that directly compares the *anxiety* students feel using a conventional system versus using a grading contract" (p. 6). Fortunately, my pilot study (Ward, 2021) examined this difference in contemporary high school students: compared to the convention grading group, 12th graders with a history of low or failing grades on the assessment who received the contract reported significantly less stress from workload demands and earned significantly higher grades.

Additionally, Cowan also wrote that "most scholarship on grading contracts in composition focuses on individual case studies…with the occasional study of students in those classes," thus the field "could use more comparative, large-scale studies of grading contracts" that reveal "how much grading contracts impact students academically or emotionally compared to other grading schemes" (p. 8). This thesis research, with 439 participants, fills the gap Cowan described, revealing contract grading's benefits, even when used during a short duration. Contract grading is typically used across entire courses; however, this study examines the use of contract grading during a single unit and thus serves as a soft entry point for teachers and departments accustomed to traditional grading practices. In this study, the limited use of contract grading generated buy-in and then success that can be built on for larger institutional changes.

The present study took place in January 2020 when all grade (9-12) and course levels (i.e., regular, AP/honors, and those accommodated for learning disabilities) adopted the grading contract for the high-stakes writing assessment. The following research questions guided the study:

1. How does contract grading affect secondary students' perceptions of their academic stress?
2. How does contract grading affect secondary students' self-worth protection behaviors?
3. How does contract grading impact secondary students' academic performance as measured by their final grade on the project?
4. Which factors (e.g., prior experience, class level, etc.) impact academic performance, motivation, and perceptions of academic stress?

Next, I locate myself in this study before explaining the significance of this work, its contribution to the literature on writing assessment for adolescents, and the structure of the thesis.

1.2 Locating the Researcher: From Shame to Equity in Evaluation

1.2.1 Shame in Evaluation

> In spite of the emotional nature of classrooms, inquiry on emotions in educational contexts, outside of a few notable exceptions...has been slow to emerge.
> —Phye, Schutz, and Pekrun, 2007, p. 3

In the 1990s, I came home from kindergarten in tears after hours of standardized testing for entrance into the Gifted and Talented Education program caused great distress for a six-year-old with then undiagnosed dyslexia from a single-parent, working-class family. After missing the threshold by two points, I internalized that I was neither gifted nor talented. Since grades and test scores send "messages that travel with great volume and

clarity" (Schneider & Hutt, 2014, p. 219), I experienced first-hand what McArthur (2018) observed in *Assessment for Social Justice*: "assessment activities become crucial to our sense of self worth" (p. 88). Self-worth theory, an organizing framework for this study, holds that worth resides in each student inherently, yet so many students—from over-strivers to underachievers—are bedeviled by the "devastating belief...that the sole measure of one's worth is [her] accomplishments" (Covington & Beery, 1976, p. 56).

My work is now concerned with psycho-emotional well-being, recognizing that inequitable assessment practices can harm learners' view of themselves and their feelings toward school by sometimes intentionally but often inadvertently "marginaliz[ing] and even exclud[ing] certain students and groups for whom such an assessment method is both inaccessible and an unfeasible way in which to best evidence their learning" (Schneider & Hutt, 2014, p. 99). While a grade should reveal a student's level of mastery (Wormeli, 2006), conventional grading practices do not accurately reflect what students know and can do (Feldman, 2019). In fact, Inoue (2019) argues that "grade have never equated to students' performances in course" (p. 132). While adolescents who perceived a higher sense of control over their workload also experienced lower levels of stress (Ainslie, Shafer, & Reynolds (1996), many assessment practices harm a student's sense of agency. Unfortunately, as Covington and Beery (1976) observed—

> Once the child enters school he progressively loses control: he is told when to learn, how quickly, and whether or not he can feel a sense of accomplishment. Evaluation becomes official and public, leading to feelings of shame and humiliation. (p. 27)

Citing Boud (2007), McArthur (2018) observes that students have become "cogs in the wheels of the assessment system" (p. 19), which has enforced classist, ableist, and racist standards (Inoue, 2019) that I will explain in Chapter Two. Throughout this thesis, I will argue for an alternative assessment approach that promotes adolescents' well-being and reduces self-worth protection behaviors. I recognize, however, both the powerlessness over the status quo of grading and assessment that adolescents, who have internalized their parents' expectations and societal pressures to avoid failure, and educators, who may feel defenseless against the demands of their administration or higher education. While conventional grading appears "fixed and inevitable—without origin or evolution" (Schneider & Hutt, 2014, p. 202), Elbow (1997) wrote in the article that introduced me to contract grading that "grading is not built into the universe" (p. 6). Grades, however, are deeply entrenched in education, making their eradication unlikely, but the findings of this research reveal that contract grading is a viable alternative tool for teaching and learning that satisfies institutional requirements to assign grades while supporting students' skill development and psycho-emotional well-being.

In the next section, I reflect on my journey as a member of a 13-person English Department at a secondary school and the motivation for this research.

1.2.2 Equitable Writing Assessment

So much of the writing assessment work we do seems complicit in sustaining inequality.
—Poe and Inoue, 2016, p. 119

This thesis adds to a growing body of recent work on equitable assessment and students' psycho-emotional well-being (Feldman, 2019; McArthur, 2018; Inoue, 2019) by examining the impact of contract grading on high school students during high-stakes writing assessment. In my Master's program for English composition, I studied the latest reading, writing, and grammar pedagogy when I first encountered Elbow (1997), who argued that contract grading creates a fairer system that restores students' agency by inviting them into the evaluation process. Contract grading seemed ideal for developing writers because it served as a task-oriented tool to help teach the writing process, not just evaluate it; however, because I did not know any instructors utilizing it, I figured it was likely to be thwarted by department chairs and administrators. After all, "change is regarded as risky or strange, hence the status quo is further reinforced" (McArthur, 2018, p. 4).

In 2013, after teaching first-year composition to college students, I joined the faculty at GSHS to influence younger writers and quickly learned that I could make grading papers alone my full-time work. Even at a private institution with smaller class sizes than public schools, I taught more than 160 students when my department chair advised spending one hour grading each research paper binder, replete with annotated sources, notes, and drafts. The thorough three-page analytic rubric I inherited made it easy to spend such an inordinate amount of time grading. While meticulous and obedient students could earn high grades through strict adherence to the rubric's rules, format errors earned others abysmally low grades on otherwise cohesive papers. If the rubric overwhelmed me (and it

did; however, the workload demands of writing assessment on teachers is a topic for future research), I wondered what impact it had on adolescents under my care.

Department data, analyzed by my department chair, revealed that students who fail as 9th and 10th graders continue to fail as they advance through high school. In 2018, 26% of the approximately 300 first-year 9th graders earned a low or failing grade on the research paper. After three years of experience, 20% of regular 12th graders also earned a D or F in 2018. Submitting an assignment for grading, after all, involves opening up one's self for critique and requires vulnerability (McArthur, 2018), and as Covington (1984) observed, students will go to great lengths to "protect their sense of worth or self-value" (p. 4). Negative academic experiences can lower students' perceptions of self-efficacy and rates of persistence (Bouffard-Bouchard, 1990) and lead them away from future challenges (Shim & Ryan, 2005).

The solution could have been a new rubric—perhaps one that was descriptive, task-specific, and thus more supportive of learning (Brookhart, 2018)—yet others maintain that rubrics are inherently flawed (Wilson, 2007; Kohn, 2006; Borman, 2018). To Anson (1989), rubrics are "guised in the cloak of reliability and efficiency" (cited in Wilson, 2006, p. xix). To Borman (2018), they attempt to standardize a non-standard process; by definition, they contain "rigid, standardized criteria" and "replace the authentic, holistic analysis of writing and reasoning with inauthentic pigeonholing that 'stamps standardization' onto a creative and analytical process, that is, nonstandard" (p.

713, 714). While recognizing that uniform standards are admirable and rubrics can be quick and efficient for teachers to use, Kohn (2006) also called them an inauthentic "tool to promote standardization, to turn teachers into grading machines or at least to allow them to pretend that what they are doing is exact and objective" (p. 12). The issues they create "cannot be solved by a new rubric" (Borman, 2018, p. 741), as I will further explain in Chapter Two.

In January 2019, we not only changed our assessment tool but our philosophy, moving from one focused on evaluation (and its counterparts of numbers and percentage points) to learning and skill acquisition. The English Department gathered weekly in Fall 2018 to re-envision the assessment when our chair described her ideal assessment tool, one focused on learning rather than performance. Stunned, I asked, "Do you mean contract grading?" No one had heard of the approach, but an *Inside Higher Ed* article called "I Have Seen the Glories of the Grading Contract…and I'm never going back" (2016) stimulated their interest:

> This sounds obvious, but for each assignment, [contract grading] forced me to decide what *matters*. With writing, it's tempting to say *everything*, and when I have the full range of numerical grades to consider, to some degree everything does matter, but when it comes to helping student writers develop, I've long known that limiting what they're focusing on at any given time is a more effective approach. . . . By defining that criteria, I gave students a clearer target to aim for. (n.p.)

Indeed, our problem was that we were valuing everything. *S*ome teachers maintained that we ought to value everything—that it was *fair* to value everything because we had taught everything—but the majority of the department recognized that our high schoolers

needed a clearer target to reach the standard of excellence we desired. McArthur (2018) cites Carless (2015) who includes five forms of fairness related to assessment, nearly all of which were violated under the old assessment approach. I argue that contract grading promotes three types of social justice in assessment that are often lacking in conventional practices: procedural fairness in the form of clear assessment standards that were perceived as challenging but understandable and interpersonal justice by improving the relationship between student and teacher by upsetting the unequal power dynamic and giving students more control through choice. In this way, students Consequently, this strikes at the cause of many grading issues, as Danielewicz and Elbow (2009) argue:

> The contract helps strip away the mystification of institutional and cultural power in the everyday grades we give in our writing courses. Using the contract method over time has allowed us to see the root of our discontent: conventional grading rests on two principles that are patently false: that professions in our field have common standards for grading, and that the 'quality' of a multidimensional product can be fairly or accurately represented with a conventional one-dimensional grade. In the absence of genuinely common standards or a valid way to represent quality, every grade masks the play of hidden biases inherent in readers and a host of other a priori power differentials. (p. 249)

Beyond just a rubric of another name, the grading contract fundamentally changes the conditions of the learning environment in critical ways that allow students to focus on labor, which is fundamental to learning, rather than grades and numbers. First, in the absence of point-values associated with each contract item, each part of the contract carries equal importance, thereby emphasizing the meaningfulness of each task and its contribution to the learning process, the ultimate goal of the project. Furthermore, central to the contract philosophy is the unusual arrangement to minimize bias that seeps into evaluation: grades (at least up to B, or 89%) are guaranteed (that is, given without the

teacher's subjective evaluation of writing quality) upon completion of the learning tasks. In this way, while a rubric artificially quantifies learning, the contract grading operates under a different philosophy, serving as a holistic and metacognitive tool for learning

The next section further outlines the philosophical underpinnings of contract grading.

1.2.3 Contract Philosophy

> Not everyone is highly talented, but most are capable of attaining an adequate level of achievement.
> —Covington and Beery, 1976, p. 143

> Adopting a contract grading model [...] is not simply changing your model of grading. It is a total upending of your assessment philosophy and of your students' expectations
> —Parks and Zimmaro, 2017, n.p.

To introduce contract grading, Elbow (2017) asks his students to imagine that they were freely attending his home studio for a cooking or painting course. Naturally, in such a class, they would receive praise and feedback aimed at improvement, rather than grades, on their omelets or paintings. The home-studio environment, he writes, is more conducive to learning than academic courses, which obligate students to attend and then require an official grade to define their performance. To approximate the beneficial working conditions of the home studio, he then argues for using "a kind of contract for grading" that fosters a culture of support that allows students to take risks (n.p.).

While satisfying institutional requirements to issue grades, contract grading works to orient students toward learning while limiting the stress and pressure of grades. Unlike conventional grading, in which students can cut corners in pursuit of the highest grade,

the contract system values labor, which is why Inoue (2019) advocates for a labor-based contract system; though essential to learning, labor is often taken for granted in conventional grading practices, which is unfair to diverse students. Fortunately, as Inoue argues, the area most within students' control—how much time they spend laboring—is most essential to developing their writing skills. As Danielewicz and Elbow (2009) observed—

> The contract creates a link between grades and work that is more frank and less roundabout: 'To get a B you must do x, y, and z.' As a result, we find the contract yielding more *total* work: more tasks done thoroughly by more students (p. 256).

While some teachers craft the contract with students (Litterio, 2018), most offer a unilateral (Danielewicz & Elbow, 2009) or "blanket" (Potts, 2010) contract that allows them to maintain full control of the course requirements and the final grade decision while surrendering much of the power over grading. Twelve GSHS English teachers utilized the unilateral hybrid approach (also used by Litterio, 2016, 2018), which was first advocated by Danielewicz and Elbow (2009), in which grades up to B (89%) are guaranteed for completing learning tasks, while grades higher than B rest on the teacher's subjective estimate of exceptional writing. Unlike Inoue's (2019) labor-based system, which resists the dominant White language standard and promotes an antiracist writing assessment ecology, the hybrid model may not go far enough in offering a socially just way to produce grades but does greatly limit the teacher's judgment for most grades.

Figure 1.1

Grading Contracts for Regular 12th Grade English

Contract for an A

The following contract is intended to detail the major elements of a senior-level research paper.

High (98%) **Medium** (95%) **Low** (92%)

You will earn an A if you meet all of the following:

1. Format the paper and Works Cited Page in **MLA format** with no errors.
2. Write a minimum of 100 paraphrased **notes** and use those only to write your paper. Cite each one with an **MLA parenthetical citation**.
3. Find **eight-ten sources** from the library databases. No more than ten sources.
4. Write a well-crafted **thesis statement and corresponding topic sentences that** includes an argument regarding a world problem that has not yet been solved.
5. **Analysis** is consistently present, compelling, and well explained within each paragraph.
6. Dissect **cause** and **effect** related to the issue as well as a well-constructed and supported **solution** (both failed and proposed). Includes well-written introduction, body paragraphs, and conclusion.
7. Cite at least **two different sources** in each body paragraph without repeating citations showing seamless and balanced use of sources.
8. Meet **minimum word** count of 2100 without exceeding 2400 words (-1% per 1% over or under).
9. Submit a printed essay, and submit to TurnItIn **on time** and receive a **clean report**.
10. Writing must be edited and pruned for **sophisticated use of grammar, mechanics, diction, and sentence structure**.

Contract for a B

The following contract is intended to detail the major elements of a senior-level research paper.

High (88%) **Medium** (85%) **Low** (82%)

You are guaranteed a B if you meet all of the following:

1. Format the paper and Works Cited Page in **MLA format** with no major errors
2. Write a minimum of 100 paraphrased **notes** to support your argument and use only those notes to write your paper. Cite each one with an **MLA parenthetical citation**.
3. Find and use **eight sources** from the library databases. No more than ten sources.
4. Write a **thesis statement and corresponding topic sentences** that include an argument regarding a world problem that has not yet been solved.
5. **The analysis** is consistently present and reasonably explained within each paragraph.
6. Write **cause**, **effect**, and **solution** (both failed and proposed) as well as an introduction, body paragraphs (that follow proper academic structure), and conclusion.
7. Cite at least **two different sources** in each body paragraph without repeating citations. Balanced use of sources.
8. Meet **minimum word** count of 2100 without exceeding 2400 words (-1% per 1% over or under).
9. Submit a printed essay, and submit to TurnItIn.com **on time** and receive a **clean report**.
10. Writing must be edited and pruned for **proper grammar and mechanics**.

Note. The A and B contracts above were used in the pilot study. Each grade level had its own contracts, which increased in difficulty from 9th to 12th grade.

Compared to conventional practices, contract grading also invites students to take a more active role in the learning and assessment process. During the first week of the high-stakes assessment, the students in this study contracted for either an A (mastery) or a B (proficiency) (see Figure 1.1). A busy student could contract for a B to reduce stress, while another could push themselves to strive for an A. The B contract, opposed to C, D, or F, served as the minimum threshold for performance to buoy all students up to their highest potential—or, as Danielewicz and Elbow (2009) suggested, "badger and cajole

every student into getting a B—that is, into doing everything we specified in the contract" (p. 254). Additionally, the English Department was influenced by Johnston and O'Neill (1973) who found that even as the criteria changed, students adjusted their behavior to meet the minimum threshold. Their conclusion urged teachers to set high standards for the minimum pass criteria, outline them as precisely as possible at the beginning of the course, and *not* to successively lower criteria for lower course grades. Some students would *not* fulfill the B contract and earn grades Cs, Ds, or Fs, but the goal was to ask all students to strive for proficiency (B) or mastery (A), not lower. In this way, the goal was the highest possible fulfillment of potential to help each student achieve the best possible outcome (Nichols, 1979).

Unlike subjective points-based rubrics or marking criteria, each contract item on the B contract is as concrete and objective as possible to yield a crude yes-or-no response, thereby limiting teacher bias. The A contract built upon the B contract by articulating what Danielewicz and Elbow (2009) call 'fuzzy criteria,' the subjective qualities of exceptional writing. In an early piece that alludes to the principles of contract grading without explicitly mentioning it, Elbow (1983) aptly described embracing the tension between a teacher's commitment to their students—who "need an ally, not a judge" (p. 56)—and to the standards of their subject. He embraces this "paradoxical coherence in teaching" by showing students the "tough standards" of his course and emphasizing his role as a gatekeeper, saying—

> Those are the specific criteria I will use in grading; that's what you are up against, that's really me. But now we have most of the semester for me to help you attain those standards, do well on those tests and papers. They are high standards but I

> suspect all of you can attain them *if you work hard.* I will function as your ally. I'll be a kind of lawyer for the defense, helping you bring out your best in your battles with the other me, the prosecuting-attorney me when he emerges in the end. (p. 59; emphasis added)

In this view, contract grading can produce a more supportive classroom environment that values labor. Next, I describe the thesis's contribution to knowledge and its structure.

1.3 Contribution to Knowledge and Structure of Thesis

The findings of this thesis reveal that contract grading is a promising alternative for secondary classrooms that uses the letters and numbers of the conventional grading system (and thus satisfies institutional requirements to assign grades) while also restoring students' agency by allowing them to meaningfully participate in their assessment. Using a social justice lens focused on reform, this study is the first study to focus on the intimate relationship between subjective experiences of stress and self-worth protection behaviors under the contract grading system. Using qualitative data from 40 interviews to explain the pre-post quantitative data from 439 participants, this study offers an in-depth understanding of how contract grading impacts students' psycho-emotional experiences to create a more just learning environment that facilitates high academic achievement for diverse learners. This work contributes the knowledge that contract grading reduces evaluative threat for high schoolers, even when employed for a single unit, to create a psycho-emotionally healthy learning environment that reduces perceptions of stress and self-worth protection behaviors to facilitate high academic achievement.

Chapter Two examines the root cause of student stress and the subject of the investigation: stress from high-stakes assessment. I synthesize the key research on contract grading and its relationship to evaluative threat, which negatively impacts well-being and performance (Hancock, 2001; Thompson & Dinnel, 2007) and ties together both organizing frameworks: the transactional model of stress and coping (Lazarus & Folkman, 1984) and self-worth theory of achievement motivation (Covington, 1992). Using stress theory, this work understands that a student's appraisal of a stressor called 'research paper' determines their reality. I then offer a critical interpretation and synthesis of the troubling history of conventional grading and where this work is situated in the field of equitable writing assessment.

Chapter Three outlines the study's methodology and explores my positionality as a teacher at the school I am researching. I have worked to thoroughly describe my decisions and methods to provide full transparency about my behavior toward participants. I also outline my ontological post-positivism and epistemological critical realism, which led to a mixed-methods approach with an explanatory sequential design that included the voices of a diverse group of adolescents. This work adds to a small body of literature on academic stress and contract grading (Fairbanks, 1992) by including an important dimension that drives students' stress surrounding high-stakes assessment: fear of failure and self-worth protection behaviors. After a call to use qualitative methods for stress research (Putwain, 2007), this work contributes 40 semi-structured interviews with

adolescents of all grades and course types to understand how and why their perceptions of stress and behaviors shifted under the grading contract.

The findings of this thesis are split into two chapters, which combine analysis with discussion, each with a distinct focus related to the research questions. First, Chapter Four examines the findings of the primary goal of the study and the first three phases of data generation (i.e., pre- and post-survey; in-depth interviews; and grade collection): the impact of contract grading on adolescents' psycho-emotional well-being, particularly their perceptions of stress, fear of failure, and approach to the task. Among the high schoolers I interviewed, evaluation stress was stronger when the teacher emphasized their control over the outcome and undermined the contract; however, for the majority of adolescents, the contract fostered a sense of control and certainty about a successful final outcome that mitigated the cause of evaluation stress, subsequently reducing academic and examination stress as well. Additionally, adolescents revealed that key aspects of the contract led them to appraise workload demands as challenging, not threatening, and thus the grading contract significantly reduced evaluative threat in comparison to conventional grading practices. The chapter concludes with an exploration of the adolescent's shift in achievement motivation, revealing that the grading contract minimized students' fear of failure and social comparison, which the students I interviewed noted are exacerbated by doubts and worries about standards and expectations.

Chapter Five focuses on the secondary focus of this thesis study: the impact of the grading contract on academic performance, as measured by the final grade on the

assessment. The quantitative analysis found that 90% of adolescents in the study fulfilled the contract by reaching mastery or proficiency, including 94% of students with prior experience (grades 10-12) and 84% of first-time students (grades 9-12). Unlike previous studies that have relied on quantitative data, this chapter also includes students' perspectives and experiences to understand how and why academic performance improved. My analysis of the qualitative data found that offering adolescents clear expectations for a desirable goal increased task manageability, inspired confidence, and reduced fear of failure, all of which made it safe to increase task-oriented effort. The chapter also examines why the academic performance of students enrolled in regular-level courses improved most significantly.

Chapter Six concludes the thesis and details the study's strengths and limitations before offering recommendations for implementing contract grading to promote achievement equity and help students meet their full potential. Including 13 different English courses taught by 12 different instructors, this thesis study can serve as an empirical foundation to inspire positive change in English Departments beyond GSHS. This is important since college instructors in the US often have considerable freedom in their course design and assessment approach, yet this same autonomy is not extended to many secondary school teachers who need to be in alignment, using the same resources, prompts, and assessments, with other teachers of the same section. Fortunately, the English Department's cohesion and the administration's support of alternative practices allowed this study to show the impact of contract grading on a diverse set of adolescents with

teachers, both new and veteran, who adopted the approach for the first time. The findings of this thesis are original and significant, providing empirical support for teachers, department chairs, and administrators to implement contract grading in secondary schools to create psycho-emotionally healthy learning environments and ultimately improve students' feelings about themselves and academic writing.

1.4 Summary

My argument in Chapter One has been that contemporary adolescents face substantial threats to their well-being, in part due to pervasive academic pressure in a neoliberalized education system focused on performance outcomes, like grades and GPAs, which have a considerable impact on an individual's life. At its worst, this achievement pressure has led to self-harm, but more often it leads to other forms of violence: damaged self-concept, fear of failure, and lack of interest in school from the constant evaluation. Unfortunately, the grading system has hurt those who most need success. While no grading system can completely remove bias that leads to discrimination and inequality, contract grading works to dismantle some of the injustices built into conventional practices. In this thesis, I argue that contract grading can help students of all levels and course types meet high standards by creating a psycho-emotionally healthy environment that supports learning while minimizing stress and fear of failure.

Chapter 2: Theorizing Stress and Self-Worth Protection Behaviors from Grading and Assessment: Concepts and Interpretations

2.1 Introduction

The primary motivation for implementing contract grading was to improve well-being, alleviate evaluative threat and fear of failure, and encourage each student to reach their full potential. This chapter explores the foundational literature on both stress (Research Question One) and self-worth protection behaviors (Research Question Two). Both theoretical and organizing frameworks—the transactional model of stress and coping (Lazarus & Folkman, 1984) and self-worth theory of achievement motivation (Covington, 1992)—are united by a cognitive focus on evaluative threat. After outlining each theory, I offer a critical interpretation and synthesis of the studies that show where this work sits in the field of equitable assessment, by first critiquing conventional assessment practices before examining the prior contract grading research. The chapter ends with what is known about the impact of contract grading on students' perceptions of stress and task-oriented motivation.

The next section examines the concept of stress, the initial impetus for implementing an alternative assessment tool.

2.2 What type of 'stress' is the subject of this investigation?

People are disturbed not by a thing, but by their perception of a thing.
— Epictetus

Central to this thesis is the concept of stress and the neuroendocrinological view that the brain, the home of perception and appraisal, is its central organ (McEwen, 2012).

Recently, academic articles examining the effects of stress have grown exponentially (Becker, 2013). While high stress serves as a risk factor for poor academic performance (Aborn, 1953; Akgun & Ciarrochi, 2003; Arsenio & Loria, 2014; Shankar & Park, 2016; Morazes, 2016) and negative health outcomes (Wilks, 2008; Galloway, Conner, & Pope, 2013; Feld and Shusterman, 2015; Shankar & Park, 2016), 'stress' still lacks a common definition among researchers and theorists despite its widespread, multidisciplinary use (Romano, 1993; American Institute of Stress, 2019), and the American Institute of Stress (2019) reported that "there has been no definition of stress that everyone accepts" (n.p.). Throughout this chapter, I will make clear the subject of this study by first outlining the psychological understanding of stress that underlies this thesis and defining important concepts related to stress appraisal.

In 1936, Selye (1974), a neuroendocrinologist deemed the 'father of stress,' laid the conceptual groundwork for this thesis when he defined stress as the body's nonspecific response, whether negative ("distress") or positive ("eustress"), to a stimulus; however, as the term became synonymous with distress, its positive benefits were ignored by lexicographers and the masses, which caused Selye to distinguish the stimulus, or stressor, from the response (American Institute of Stress, 2018). Growth in the field led researchers to recognize the importance of cognitive appraisal, the "the ongoing and moment-by-moment way in which [an] appeal is perceived and interpreted by the recipient" (Putwain et al., 2016, p. 22). This work examines three appraisal processes (i.e., primary, secondary, and reappraisal) explained in the next paragraph.

In their seminal work, Lazarus and Folkman (1984) outlined a psychological theory of stress called the transactional model of stress and coping, which understands stress as the interplay of two factors: a stimulus and an individual's interpretation of the potential stressor. Primary appraisal is the process of discerning the potential 'threat' of a stimulus, like a high-stakes assessment. Secondary appraisal follows when an individual takes stock of their available resources for managing the perceived threat, which is the imminent anticipation of potential 'harm,' or psychological damage (Lazarus & Folkman, 1984). Perceiving that one's abilities are not well-matched with the task, or that poor performance will harm self-image, can increase evaluative threat and maladaptive coping strategies, like avoidance orientation or procrastination, to protect self-worth (Thompson & Parker, 2007) and avoid harm (Lazarus & Folkman, 1984).

Perception determines whether the experience was positive or negative, challenging or threatening, which is well illustrated in studies with first-time and experienced skydivers. Both groups had statistically similar physiological reactivity, including cortisol activation and heart rate (Allison et al., 2012; Hare, Wetherell, & Smith, 2013); however, only first-time skydivers reported significantly higher psychological distress (Hare, Wetherell, & Smith, 2013). This helps to understand other work that has revealed that for some, increased perceptions of stress lowered academic performance (Ng, Koh, & Chia, 2003; Spivey, Havrda, Stallworth, Renfro, & Chisholm-Burns, 2020), while perceptions of stress did not impact the performance of others, such as those training to be medical

doctors (Sanders & Lushington, 2002), a self-selecting program that admits high-ability students who may be better prepared to cope with stress for a variety of reasons. Notably, the teenage participants in this study are more diverse in skill and ability and more vulnerable to the impacts of stress, which can impact brain development, particularly neural maturation, and increase morbidities, such as anxiety and depression (Eiland & Romeo, 2012). Additionally, girls who earned higher grades and reported higher levels of workload demands also had higher cortisol levels (Östberg, Plenty, Låftman, Modin, & Lindfors, 2018). For the majority of adolescents, high stress is related to negative academic affect and disengaged coping, which results in lower academic performance (Arsenio & Loria, 2014). Chambel and Curral (2005) found that a student's levels of anxiety and depression, as well as their satisfaction with academic life, depends upon their perception of academic work characteristics.

Putwain (2007) recommended identifying the subject (i.e., academic or examination) of investigation in educational stress studies. Academic stress occurs when a student perceives scholastic demands as taxing or exceeding their resources (Wilks, 2008), such as time, energy, or abilities. While two scales focused on stress appraisal and one on academic stress (e.g., workload demands, time constraints, etc.), many of the survey statements were adapted to reference the 'research paper,' thus evoking 'examination stress.' As Connor (2001, 2003) and Hall (2004) (cited in Putwain, 2007) correctly observe, much of the stress of scholastic demands (i.e., academic stress) actually stems from their relationship to the assessment (i.e., examination stress). For example, the students in my study were asked to take notes, which as a classroom activity is generally

low-stress; however, taking notes with the understanding that the notes are for a major assessment, worth 20% of their final grade, can produce examination stress, the true focus of this research.

2.3 How does evaluative threat impact students' self-worth protection behaviors?

This study also utilizes Covington's (1992) self-worth theory of achievement motivation, which states that individuals desire to uphold a positive self-image and will go to great lengths to do so. Self-worth protection also stems from a situation's perceived evaluative threat, or the "forecast of failure which is likely to impact negatively on the self, threatening self-estimates of ability and challenging the already fragile estimates of global self-worth" (Thompson & Parker, 2007, p. 131). Students employ self-protective mechanisms, such as self-handicapping, the intentional withdrawal of effort to avoid showing low ability and "protect the already uncertain self-images against further assail" (Thompson & Parker, 2007, p. 131). Students who are high in the trait of self-handicapping experience greater anxiety (Thompson & Richardson, 2001; Thompson, 2004), higher rates of helpless thinking and behavior (De Castella, Byrne, & Covington, 2013), greater negative affect (Thompson & Richardson, 2001), and lower overall academic performance (Martin, Marsh, & Debus, 2001).

Also important to this thesis is Thompson and Dinnel's (2007) finding that self-worth protection was best understood as "an outcome of choking under pressure, fueled by evaluative threat" (p. 509), which GSHS teachers and counselors (personal

correspondence) observed prior to this study. Unfortunately, rather than work to reduce fear of failure, teachers often use it as a motivator, relying on its presence by frightening children into work (Jackson, 2010). In fact, GSHS teachers often used what Putwain et al. (2016) call fear appeals, which can elicit either a challenge or threat appraisal. How students perceive the appeal can vary according to the message's severity (Putwain & Roberts, 2012; Putwain & Remedios, 2014). Depending upon the perceived evaluative threat, students high in self-worth protection perform differently (Thompson, 1999): in situations with limited risk to self-worth, a student will generally perform well (Thompson et al., 1995), yet high-risk situations can lead to greater evaluative threat and anxiety, which can impair performance and lead to underachievement (Thompson & Dinnel, 2007). Hancock (2001) found that classrooms with higher evaluative threat led all students, including those *without* a propensity toward 'test anxiety,' to earn lower grades. They concluded that "all students are more motivated to learn in classrooms deemed less evaluative" (289).

The next section examines the studies that warranted this research on adolescents' psycho-emotional well-being during high-stakes writing assessment.

2.4 Which studies warranted this research?

> A noncompetitive task-oriented society might equal the intellectual achievements of our competitive system. We cannot run experiments at the level of society to examine this question, but we can and should do this at the level of the classroom.
> —Nicholls, 1979, p. 1081

One call for this research came from Putwain (2007), who asserted that "serious consideration should also be given to qualitative research strategies, especially appropriate for investigating academic stress in the open system of school" (p. 217). Additionally, as explained in Chapter One, this thesis study sought to strike at one cause of adolescents' stress: grades, rather than symptoms. Recent qualitative studies examined medical students' top stressors, which included perceived unfair grading practices (Sarkar, Menon, & Kumar, 2020); fear of failure in university students taking an eight-week mindfulness-based stress reduction course (Hjeltnes, Binder, Moltu, & Dundas, 2015); Arab-American students' perceived stress and bullying experiences (Albdour, Lewin, Kavanaugh, Hong, & Wilson, 2017); and app-based stress reduction in college students (Fleischmann, Harrer, Zarski, Baumeister, Lehr, & Ebert, 2018). This work, however, focuses on the impact of a pedagogical change.

In the area of self-worth protection, Thompson and Dinnel (2007) encouraged teachers to minimize the evaluative threat in classrooms, which can lead students to choke under pressure. In their pedagogical recommendations, Thompson and Dinnel (2003) encourage educators to minimize the implications of failure and evaluative threat in classrooms by reducing uncertainty and the adverse effects of failure and incorporating practices that "de-emphasize individualistic, competitive orientations...to ameliorate the concern over ability status" (p. 104). In 1973, Smith and Lerch suggested that contract grading may alleviate the threat and stigma of failure, yet my research is the first empirical study to examine fear of failure and evaluative threat under the contract.

In the area of contract grading, Lindemann and Harbke (2011) studied its impact on the academic performance of 40 first-year psychology students, noting that the approach is "rarely used in classrooms today" (p. 1) but urging teachers to utilize contract grading to impact effort and motivation. Under the contract, students rated themselves higher in working hard for their grades, enjoying the course format, and developing independent thinking compared to their peers in a conventional grading course. Their other results support previous studies with college students: under the contract, students perceived greater control over their grade (Grau, 1999; Kirschenbaum & Wetter Riechmann, 1977; Polczynski & Shirland, 1977); reported high course evaluation scores (Kirschenbaum & Wetter Riechmann, 1977); and earned higher grades than their peers with the conventional grading course, which corroborates the findings of my pilot study (Ward, 2021). For further study, they suggest an "analysis of individual differences in the effect of contract grading and its impact on individual student performance" (p. 5). Using the self-worth protection theory of achievement motivation, the qualitative phase of this thesis sought to understand the impact of contract grading on adolescents' effort and motivation. Lindemann and Harbke (2011) suggest further study with upper-division college students, on whom there is a scarcity of studies, because "it is possible that contract grading is even more beneficial among older students, given the growth that often occurs during college" (p. 5). This belief may be another reason why this thesis is the first to examine the impact of contract grading on adolescents (ages 13-19), whose developing brains are more vulnerable to the negative impact of stress.

Most recently, Litterio (2018) studied the impact of contract grading in an upper-level technical writing classroom with 18 students who participated in the contract's development, as Lindemann and Harbke (2011) suggested. Gathering exploratory data from the teacher-as-researcher method and anonymous, open-ended surveys, the findings revealed that students embraced being involved in all aspects of their assessment and enjoyed the clear expectations and autonomy. While Litterio (2018) found that "students are still fixated on traditional grades," she concluded that instructors should "recognize and return to contract grading as a powerful system not only to break through barriers of conventional assessment but also to reinforce the collaborative and individual practices involved with writing" (p. 9). For further research, she suggested examining students with more experience with the subject material. Before this study, however, the English Department worried that the contract would be less effective with first-year students who may have less understanding of the contract's terms, a point covered in Chapter Four. Finally, as Litterio observed, the final grade was the strongest source of motivation for success for the majority of the adolescents I interviewed, a topic also covered in Chapters Four and Five.

The next section examines which studies this thesis research speaks to, outlining important gaps and why filling such a gap with adolescents during high-stakes assessment is important.

2.5 What studies does this research speak to?

As mentioned previously, this work adds to a growing body of recent work on equitable assessment and students' psycho-emotional well-being, including McArthur's *Assessment for Social Justice* (2018), Feldman's *Grading for Equity* (2019), and Inoue's (2019) *Labor-Based Grading Contracts: Building Equity and Inclusion in the Compassionate Writing Classroom*. In a society that views education as a means of social mobility, conventional grading practices fuel evaluative threat by yielding a symbol of academic achievement—that is, a one-dimensional number or letter known as a grade—that has substantial power over students' learning, self-concept, and lives beyond school. The far-reaching effects of assessment impact students' learning and self-belief, and ironically, assessment often distracts students from what teachers are trying to assess by moving attention from learning to performance. Conversely, social justice in assessment is concerned with equitable practices that provide the resources to allow all students to meet their potential (Nieto & Bode, 2007), minimize attainment gaps (Hanesworth, Bracken, & Elkington, 2019), and develop student voice and agency (Chapman, Hobbel, & Alvarado, 2011).

This work builds upon research that critiques the inequity, bias, and inaccuracy that seeps into conventional writing assessment (Breland, 1983; Rachal, 1984; Brimi, 2011), and the next section leads to the call for educators concerned with equity and social justice to consider alternative practices that challenge the status quo.

2.5.1 Critiquing Conventional Writing Assessment

When we consider the practically universal use in all educational institutions of a system of marks, whether numbers or letters, to indicate scholastic attainment of pupils or students in these institutions, and when we remember how very great stress is laid by teachers and pupils alike upon these marks as real measures or indicators of attainment, we can but be astonished at the blind faith that has been felt in the reliability of the marking system.
—Finkelstein, 1913, cited in Schinske and Tanner, 2014, p. 159

While many changes have taken place to adapt educational practices for the 21st-century beliefs that all students can meet demanding standards and deserve the opportunity to succeed (Feldman, 2019), McArthur (2016) observed that "assessment has proven a sticky practice, reluctant to change, immune to innovation" (p. 6). Just fifty years ago, however, in 1971, only 67% of primary and secondary schools issued grades (Feldman, 2019), yet this included 80% US students (Schneider & Hutt, 2014), likely because larger schools moved from written feedback to one-dimensional letter grades earlier to meet the increased demands. Before grades, teachers offered narrative descriptions that explained a student's competency and areas for improvement, yet this was a time-consuming endeavor (Brookhart et. al, 2016). In 1971, only 35% of US secondary teachers believed that letter grades were the best reporting method (Schneider & Hutt, 2014), revealing that grades were more organizational than pedagogical, yet teachers felt stuck in a widely embraced method, as I did early in my career. Alternatives, like contract grading, I believed, lacked empirical support for secondary classrooms and thus would get struck down by department chairs and administrators.

Ironically, however, the current system that "has been hardwired into our conception of schools" is inequitable (Feldman, 2019, p. 12) and rife with validity issues, as I will explain later in this section. Observing the unhealthy relationship between teachers and students and symbols of learning, Inoue (2019) wrote that "the yearning for grades, the yearning to be graded, has become so habitual to many students, especially those who have been rewarded by them, that it feels natural" (p. 72). Grades have become an unhealthy obsession in contemporary schooling and for this reason, I believe McArthur (2018) noted that they have become 'fetishized.' Perhaps, then, it is their seductiveness that has blinded educators to their deleterious effects on learning and why "teachers regularly use and leaders tolerate grading systems that may appear to be accurate but are devoid of the most basic elements of mathematical reasoning and are neither fair nor effective" (Reeves, 2006, quoted on O'Connor, 2011, p. 138). With 100 distinct levels, the US grading system offers only 30 levels of "passing" with a C or higher (70-100) and 60 ways to "fail," thus it is skewed toward failure (see Table 2.1) and often marginalizes and excludes certain students, perpetuating opportunity and achievement gaps, as Feldman (2019) observed:

> To promote equity, we implement restorative justice discipline policies, learn culturally responsive instructional strategies, teach more diverse authors and perspectives, and expand our repertoire of assignments and assessments to address the different ways students learn. Yet our grading system remains virtually unchanged. By continuing to use century-old grading practices, we inadvertently perpetuate achievement and opportunity gaps, rewarding our most privileged students and pushing those who are not. (p. 53)

Table 2.1

Grading Scale in the US

Grade	Percentage	Description	Grade Point
A	90-100	Excellent	4.0
B	80-89	Good	3.0
C	70-79	Average	2.0
D	60-69	Below average	1.0
F	<59	Failing	0.0

Note. The grading scale in the United States is tilted toward low and failing grades.

The use of number grades, for example, was concurrent with the belief in an objective measure of achievement that resulted in the mental testing movement (i.e., IQ tests) of the early 1900s. The standardized grading system began to serve as "a key technology of educational bureaucracy, a primary means of quantification, and the principal mechanism for sorting students" (Schneider & Hutt, 2014, p. 202) into "academic tracks that best reflected their supposedly fixed intellectual capacity" (Feldman 2019, p. 53). As a result, standardized grading served as a quantitative justification to deny equal opportunities for students based on race or class. Similarly, for writing assessment, rubrics were built on exclusionary practices, beginning at Ivy Leagues and other prestigious institutions who tested students via writing on their understanding of ancient geography and classical languages. This knowledge, a sign of wealth and privilege, allowed these institutions to

exclude immigrants who attended public schools on the basis of their essay, not their race or class. This "exclusionary motivation," Borman (2018) argued, was their unstated goal:

> [Rubrics were] developed initially to create barriers to access to higher education for lower-income and immigrant applicants. Later, rubrics materialized in the form of standardized tests, ostensibly designed to reflect objectivity in assessment but unavoidably subjective. The federal government glommed onto standardized testing as a method of controlling funding to public schools, exclusionary motivations became legislation: students are assessed on standardized scales, educators are assessed for effectiveness, and learning takes a back seat to 'teaching to the test.' School systems with better means and students with higher family income to surpass students of lesser means in federal 'scoring systems,' and educational institutions continue to rank students according to inherent knowledge or ability. (p. 738)

Kohn (2006) observed that many contemporary teachers are against standardized tests and curricula that provide the illusion of objectivity, yet their opposition "mysteriously fails to extend to standardized in-class assessments" (p. 12-13), which involves biased writing assessment. For example, Starch and Elliott's (1912) landmark study asked high school English teachers (n=142) to evaluate two papers according to their school's standards. The results varied widely, with the first paper earning scores between 64-98 and the second from 50-97–one paper received 30 different scores. They concluded that "the promotion or retardation of a pupil depends to a considerable extent upon the subjective estimate of his teacher" (p. 454). More recently, Brimi (2011) trained teachers to use a 100-point rubric and then replicated Starch and Elliott's study with almost identical results, finding a 46-point variation in scores.

For writing assessment, as Guskey (2013) observed, "even if one accepts the idea that there are truly 100 discernible levels of student writing performance, it's clear that even

well-trained teachers cannot distinguish among those different levels with much accuracy or consistency" (p. 70). As Inoue (2019) observed, "The more distinctions that are made in a grading system the less consistent grades can be" (p. 133). To Danielewicz and Elbow (2009), teachers are "experts about writing, but individuals, nevertheless, who cannot pretend to be wholly impersonal or fair" (p. 247). In fact, in addition to time, pressure, and fatigue, Schinske and Tanner (2014) cite the following studies that reveal factors that can influence an instructor's score of an essay:

- the penmanship of the author (Bull and Stevens, 1979);
- sex of the author (Spear, 1984);
- ethnicity of the author (Fajardo, 1985);
- level of experience of the instructor (Weigle, 1999);
- order in which the papers are reviewed (Farrell and Gilbert, 1960 ; Spear, 1996); and
- the attractiveness of the author (Bull and Stevens, 1979).

Unfortunately, bias in writing assessment still hurts working-class students, students of color, and multilingual students most. For example, when examining a US college writing program's grade and survey data, Inoue (2014) found that students of color suffered from failure in writing classrooms more than White students, concluding that "the bottom line is that courses using grades and judgements on quality of writing produced more frequent failure in particular racial populations" (p. 338). In a later work, he put it bluntly: "Grading, because it requires a single, dominant standard, is a racist and White supremacist practice. There is no way around it" (Inoue, 2019, p. 5).

Grading, however, is now among the most "sacred and yet deleterious" practices in education (Edwards & Edwards, 1999, p. 260). It is sacred because many teachers cannot

fathom an educational system that exists without grades and deleterious due to the harm they cause learning and well-being of students and teachers, who are burdened with the time-consuming labor of grading. Due to its deleterious impact on well-being, learning, and self-concept, grading, Edwards and Edwards (1999) argued, is "not benign" (p. 260). Concerned with its life-or-death consequences during the controversial Vietnam War, Simon, Kirschenbaum, and Napier (1970) called grading "basically immoral" because they "don't enhance learning" but they do decide "who ships out to Vietnam, who drops out, and who stays on the football team" (p. 478).

Also using religious diction, Feldman (2018) called grading a hallowed practice. He notes that even broaching the subject of grading can be contentious as it is intertwined with teachers' identity, values, and autonomy. In an educational system where teachers must obey content standards from educational boards, grading is one of the few remaining "island[s] of autonomy" (p. 4). Such authority has even been etched into certain states' codes that protect teachers from administrators demanding that they change a grade. Ultimately, grades are "the most concrete symbol of our authority and expertise," and they are "entirely within our control" (p. 4), which is part of the problem. For these reasons, it can be hard to examine grading practices without seeming to challenge a teacher's command over her classroom and eliciting an emotional response, but teachers have forgotten that grades represent evaluation—they are not evaluation itself. A written product represents learning, but it is not learning itself, as Inoue (2019) argued:

> When grades are present, they hijack the students' purposes for their labors and how they understand teachers' labors. Their purpose becomes to get a grade, our to give one (p. 148).

As I update my grade-book each week, it is too easy for me to lose sight of learning and think that my job is to give grades. In fact, over the years, I have received regular reminders to give grades, with no emphasis on what my grades are measuring or what they mean.

The next section outlines the theory and practice of contract grading, before examining its impact on students' task-oriented motivation and perceptions of stress.

2.5.2 Contract Grading

Without eliminating grades entirely, contract grading de-emphasizes grading—that is, it "attempt[s] to make [grades] less present and exert less pressure, by ironically paying attention to how grades are constructed" (Inoue, 2019, p. 142). To Inoue (2019), a grading contract is a "social agreement with the entire class about how final course grades will be determined" (p. 130). While some teachers craft the contract with students (Litterio, 2018), most offer a unilateral (Danielewicz & Elbow, 2009) or "blanket" (Potts, 2010) contract that allows them to maintain full control of the course requirements and the final grade decision while still providing students with a meaningful choice over their learning goals. In "Contracts Help Solve Problems," one of the earliest articles advocating for contract grading, McLaughlin (1961) defines the learner-centered approach:

> The contract method is a system whereby a series of project assignments are devised by the teacher according to an increasing scale of difficulty, with the most difficult tasks rated as an A, intermediates at B, and those of the least difficult at the C level. The students, upon being presented with the prepared list of choices,

> have the option of selecting whichever grade they themselves would like to earn. (p. 418)

In Inoue's (2019) labor-based system, grades are earned exclusively through labor and thus are "free of quality-based grading judgement" (p. 141). Offering only three grades (3.4 for C, 3.7 for B, and 4.0 for A), he does not appear required to use an electronic grading system, like most secondary teachers; however, his goal is to "de-emphasize product" and focus on learning, which happens as a result of labor (p. 132). While conventional grading practices neglect labor and thus are often "unfair to diverse groups of students" (p. 129), students earn higher grades by doing more labor. In the more common hybrid approach examined in this study (also used by Litterio, 2016, 2018), which was first advocated by Danielewicz and Elbow (2009), grades up to B are guaranteed for completing learning tasks, while grades higher than B rest on the teacher's subjective estimate of exceptional writing.

College students have viewed contract grading as a fairer system of assessment than conventional grading practices (Taylor, 1971; Hassencahl, 1979) and have reported fewer negative academic emotions, like stress and anxiety, and preferred it to the conventional approach in course evaluations (Parks & Zurhellen, 1978; Hassencahl, 1979; Farber, 1990; Fairbanks, 1992; Lindemann & Harbke, 2011). Contract grading has been implemented in a variety of courses across the disciplines, including physical education (Smith & Lerch, 1972), social studies (Eastridge, 1975), political science (Bowman, 1977), speech courses (Stelzer, 1975; Wolvin & Wolvin, 1975), introductory psychology (Kirschenbaum & Riechmann, 1975), general education (Taylor, 1980), dental education

(Lewis & Killip, 1971), writing courses (Leahy, 1980; Farber, 1990; Elbow, 1997; Danielewicz & Elbow, 2001), legal education (Aiken, Koplow, Lerman, Ogilvy, & Schrag, 1985); business law (Polczynski & Shirland, 1977); accounting (Zarzeski, 1998); management education (Hiller & Hietapelto, 2001), and technical writing (Litterio, 2016, 2018). Few studies, however, have included students' voices and perception of the grading contract, as Spidell and Thelin (2006) noted. In their study of college students' perspectives of contract grading, they found that "a contract does foster a sense of responsibility for and ownership of learning, albeit with a corresponding rise in anxiety and resistance" (p.44). They surmised that students' resistance was a result of undemocratic educational conditioning.

Although the earliest known work on contract grading advocated for its use in high schools (McLaughlin, 1961), most empirical work with contract grading has been relegated to the 1970s and focused on the development and implementation of contracts (Amsdem, 1970; Barkley, 1975; Bowers & Howard, 1975; Kokus & Mussoff, 1975; James, 1977). Contemporary college students under the contract earned higher grades compared to their peers in conventional grading courses (Lindemann & Harbke, 2011), reported increased involvement in the assessment process (Litterio, 2016), and perceived a stronger sense of control over their grades (Litterio, 2018). The lack of empirical work in secondary classrooms may exist for several reasons.

First, secondary education programs do not appear to include contract grading as a grading system. Of my 12 English colleagues, I was the only one who had learned about contract grading. In college classrooms, instructors often have more freedom to implement alternative approaches, yet its use is still relatively rare, as Lindemann and Harbke (2011) noted. Since the 1970s, two recent studies on contemporary college students (Lindemann and Harbke, 2011; Litterio, 2018) have encouraged the implementation of contract grading for improving academic performance, student responsibility, and perception of control, which can reduce perceptions of stress. Given the role of secondary schools in preparing students for higher education, high school teachers may be hesitant to adopt approaches that are not widely used in colleges and universities.

2.5.3 Contract Grading and Task-Oriented Motivation

Work on contract grading and motivation has focused on students' final grades as the measure of success. Unfortunately, the educational system has become performance-oriented, focused on extrinsic outcomes (Janzow & Eison, 1990; Svinicki, 1998; Page & Alexitch, 2003), not learning. Even educational research has treated grades as an external stimulus to motivate students' behavior, as Hassencahl (1979) observed:

> All grading systems are somewhat Skinnerian in their point of view, equating the grade with the stimulus and the student's effort toward the grade as the response. The traditional assumption has been that all students are motivated in the same way by high grades and that the grades "B," "C," and "D" are given for degrees of falling short of the "A" goal. A grade contract is no less Skinnerian, but allows the students to determine what level of competency and involvement he or she wishes to achieve. For example, a student may take physics in college and find himself involved in the struggle for an "A," when all he really wants to know about physics is on a "C" level. (p. 31)

Grades can also inadvertently motivate students away from learning; as Schinske and Tanner (2017) observed, grades "sometimes promote learning, [yet] they often promote an entirely separate set of behaviors" (p. 219). To Kohn (2006), grades, opposed to written feedback, create students who "think less deeply, avoid taking risks, and lose interest in the learning itself" (p. 12), while Elbow (1997) argued that "conventional grading seduces too many students into thoughtlessness," (p. 82). To Kohn (2006), grades are primarily a means of coercion for teachers who seemingly motivate students with the threat of a bad grade in a sticks-and-carrots approach. Many classrooms foster performance, not learning, goals. As contract grading advocate Elbow (1997) writes, "When we 'motivate' students with grades, we are not building motivation but undermining it: we are gradually *zapping* the ability to work or think or wonder under their own steam" (p. 12-13). If the reward for successful learning is not intrinsic, O'Connor (2007) observed two undesirable results: not only do students focus their attention on the reward or punishment, not the desired learning-focused behavior, but the reward must be increased to sustain motivation for the task.

Performance goals also have potentially dire consequences on well-being and self-belief. Nicholls (1984) observed that for performance-orientated students, "the goal is to develop or demonstrate—to self or to others—high ability, or to avoid demonstrating low ability" (p. 328). To avoid appearing incapable, students focused on ability-linked goals can lead to withdrawal and poor performance, as well as a loss of intrinsic motivation and

self-worth (Dweck & Elliott, 1983; Grant & Dweck, 2003). Martin, Marsh, & Debus (2001) found that performance orientation predicted self-handicapping and defensive expectations. To Kohn (2011), this is a rational response in a world where performance, not learning, is the goal. Fear of failure, then, is often at the root of low engagement; in other words, students who appear to lack motivation for schoolwork are often deeply motivated to preserve their self-esteem in a society that emphasizes and values academic success (Jackson, 2010).

Before this thesis, three studies examined contract grading's impact on student motivation (Polczynski & Shirland, 1977; Parks & Zurhellen, 1978; Callahan, 1979). In one of the largest studies of contract grading, with nine courses and 280 college students enrolled in business law and basic management, Polczynski and Shirland (1977) tested Vroom's motivation theory of expectancy, which asserts that goal-oriented effort increases when the path to the goal is clear. Based on self-report quantitative data, the contract grading group reported significantly higher instrumentality—that is, the link between performance and the goal—leading the researchers to conclude that "it is possible to change levels of students' effort and motivation to perform by clarifying the path to the goal of grade achievement" (p. 241). The next year, Parks and Zurhellen (1978) found that 87.5% of students (n=72) reported 'greater motivation under the contract' and more determination to succeed than their peers in a conventional grading course. Finally, Spidell and Thelin (2006) reported that the majority of students contracted for an A or B, making it appear to succeed at "motivating high student aspirations" (p. 52). This thesis

took a different approach toward motivation, examining the contract's impact on achievement motivation, defensive strategies, and self-protective mechanisms during a high-stakes assessment.

Callahan (1979) also examined contract grading as "a possible stimulus for student motivation and performance" (p. 3). Notably, students were offered contracts for every grade level, even F. While students appeared to prefer the contract system to conventional grading and "enjoyed the relaxed atmosphere afforded the course by the use of the contract grading method" (p. 7), students often contracted for and then earned lower grades than their peers in the conventional grading group, a phenomenon Elbow (1997) also observed. Allowing students to choose their desired effort may result in less striving for excellence (i.e., fewer As), more passing grades, and less academic stress. At GSHS, this was a point of contention: some teachers maintained that we ought to demand all students work toward mastery. Hassencahl (1979) also observed this tension: if a student simply wants a C-level understanding of a subject is it the role of the teacher to challenge all students to work toward their highest potential, even if it is against their choice and brings academic stress?

At GSHS, demanding mastery under the analytic rubric did not yield the desired results, as mentioned earlier: 26% of 9th and 20% of 12th graders earned a D or F in 2018. Contract grading offers students a choice that can bring personal investment while also asking all students to reach their potential. As a result, students' work-life balance may

improve, stress may decrease as a result of grade uncertainty, and one consequence may be fewer As but more passing grades and less academic stress. In 2019, however, my pilot study (Ward, 2021) revealed that students with a history of low grades on the research paper earned significantly higher grades under the contract: six times as many As and 2.5 as many Bs as statistically similar peers in the conventional grading group.

Others corroborate the findings of my pilot study, showing significant grade improvements under the contract grading system, including developmental mathematics (Miller, 1974); behavior contracts for 7th graders (Williams & Anandam, 1973); and introduction psychology (Lindemann & Harbke, 2011). Most recently, Lindemann and Harbke (2011) studied the use of contract grading's effectiveness with 42 contemporary college students, finding that students were three times more likely to earn an A in the course under the contract and one third as likely to withdraw or fail. While Poppen and Thompson (1971) found no statistically significant grades difference, they reflected that they may have constructed a poor contract yet concluded that the contract helped the instructor clarify their objectives; demystified the grading process for students; shifted responsibility to the student to "become educated"; allowed for student choice and voice in the learning and grading process; and fostered an environment of greater cooperation between teacher and student.

2.5.4 Perceptions of Stress Under the Contract

Smith and Lerch (1972) argued that the contract had the ability to "alleviating the threat and stigma of failure in education, a great source of anxiety and guilt would be removed"

(p 82), yet until this thesis, only one empirical study (Fairbanks, 1992) examined student stress under the contract as its main objective, although the only data generated to measure stress was final course grades. To treat 'math anxiety' (defined in the study as 'nervousness' and 'worry') in college students, Fairbanks offered an optional contract to students who feared failing his course with promising results: of the 77% of the students who signed the optional contract, including 69% of the females and only 32% of the males, 86% of the students who were worried, nervous, or afraid of failing fulfilled their contract. Ultimately, the contract group outperformed the conventional grading group, which had more students fail the course. While the study did not report enough data for me to offer a full gendered analysis of the final results, I observed that males were less likely to opt for the grading contract, perhaps because admitting to being nervous and worried would be akin to admitting to weakness, particularly given math-gender stereotypes (Eccles, & Jacobs, 1986; Passolunghi & Tomasetto, 2014). Unfortunately, the conventional grading group, dominated by males, had a higher rate of failure and course withdrawal and may have benefited from the grading contract. Similar stereotypes exist with females and English language arts, the focus of this thesis, and recently, this relationship was confirmed. While no gender achievement differences were found for math, females (grades 3 to 8) outperformed males in English in nearly every district across the country (Reardon, Kalogrides, Podolsky, & Zárate, 2019). Chapter Five examines gender-performance differences under the contract.

Finally, Fairbanks (1992) "sensed" that the contract reduced anxiety because "students stated that knowing they would pass the course relaxed them and that they did far better than they had expected" (p. 430). Other studies have made similar conclusions that contract grading has the ability to reduce stress by promoting a more 'relaxed atmosphere' (Potts, 2010; Callahan, 1979). The only study to ask participants directly, however, was Parks and Zurhellen (1978), who found that 90% percent of the college students (n=82) studied felt less 'anxious' under the contract. This work differed from previous studies by using in-depth qualitative data to thoroughly examine students' experiences and perceptions of stress under the contract.

2.6 Summary

Socially just education requires concern with students' well-being, sensitivity toward the struggles they face, and equitable pedagogical and assessment approaches that foster deep learning. Assessment practices have the power to shape not only teaching and learning but also students' well-being and self-concept; thus, this chapter outlined two key theoretical frameworks—stress theory and self-worth protection theory—which serve as organizing tools and interpretative lenses for this work. This chapter situated this study in the broader field of grading and assessment, examining the issues associated with conventional grading practices and what is known about student stress and achievement motivation under the contract. The next chapter outlines the study's rigorous mixed-methods design.

Chapter 3: Methodology

3.1. Introduction

To 'reduce stress and increase success,' the GSHS English Department met weekly to re-envisioning the annual high-stakes writing assessment in Fall 2018. In January 2019, one year prior to the implementation of contract grading and this thesis research, the English Department reduced the length of the assessment by 600 words (see Table 3.1) but maintained the points-based analytic rubric for all courses, except those in the contract grading pilot study (Ward, 2021), which will be outlined later in this chapter. Kralevec and Buell (2000) assert that homework has an especially harmful impact on disadvantaged students who get penalized when their personal environments challenge their ability to complete assignments at home; thus, reducing the length of the project while maintaining the same number of instructional minutes, the department thought, would promote achievement equity by allowing most students to complete all the work in class. The pilot study found that only change in workload demands did not alter students' perceptions of stress; however, combined with the grading contract, the 12th graders reported a significant decrease in stress from workload demands and earned significantly higher grades. Following the success of the pilot, all English courses adopted the grading contract in January 2020 when this thesis research commenced.

This chapter begins with my ontological and epistemological underpinnings that informed this study. I then outline the pilot study before describing the current study's instrument, participants, data generation, and data analysis. The chapter concludes with

the study's ethical considerations, particularly how I sought to minimize the influence of my positionality as a teacher in the school, show care for participants during COVID-19, and keep the mental health of participants paramount.

3.2 Ontology and Epistemology

All research necessarily starts with a person's view of the world.
—Grix, 2010, p. 66

Undergirding this work is ontological postfoundationalism, a "middle way between the objectivism of foundationalism and the relativism of many forms of nonfoundationalism" ("Postfoundationalism," 2017, para. 1). Consequently, this work recognizes "the possibility of alternative valid accounts of any phenomenon" (Maxwell, 2002, p. 5). Unlike the foundationalists who believe that "reality exists independently of our knowledge of it" and that "true knowledge must rest upon a set of firm, unquestionable [and] indisputable truths" (Grix, 2010, p. 60, 61), I hold that "the ideals of truth, objectivity, and rationality" may exist *out there,* yet human reason and understanding is ultimately "provisional, contextual, and fallible" ("Postfoundationalism," 2017, para. 3). The human lens, bound to a particular context, interferes with a researcher's ability to *know* with objective certainty. In fact, reality may be "socially and discursively 'constructed' by human actors" (Grix, 2018, p. 61), as the post-foundationalists argue, but the pursuit of a truth—or more specifically, the desire for equity and justice for all—serves as a worthy ideal to drive inquiry.

Epistemological critical realism, a reflexive philosophical stance and meta-theoretical position, then follows, which "den[ies] that we can have any 'objective' or certain knowledge of the world, and accept[s] the possibility of alternative valid accounts of any phenomenon" (Maxwell, 2002, p. 5). To the critical realist, one's particular perspective serves as the basis for all theory; an individual cannot attain a single, "correct" understanding of the world, or what some call a "God's eye view," that is independent of any particular viewpoint; as a result, knowledge, held in the human mind, is "partial, incomplete, and fallible" (p. 5). While the constructivists hold that there are "multiple realities," or "the sense of independent and incommensurable worlds that are socially constructed by different individuals or societies," critical realists believe in many perspectives of reality. In this work, each perspective is constructed through the interactions between participants and the researcher. Arising from these ontological and epistemological worldviews, this research sought to understand the impact of grading practices on students' experiences and perspectives, particularly their stress appraisal and self-worth protection behaviors, through a mixed-methods design that values students' voices to understand how and why.

3.3 Pilot Study

In January 2019, at the time the workload decreased, the pilot study (Ward, 2021) compared the impact of conventional grading and contract grading on the academic stress and performance of an identified high-need group: 12th graders enrolled in regular English with a history of low or failing grades. Among seven courses with an identical

prompt and teaching materials, three maintained an analytic rubric while four sections were evaluated with the contracts. Using a mixed-methods approach with a convergent design, a survey entitled "Perceptions of Academic Stress" generated both the quantitative and qualitative datasets, including objective (e.g., time spent doing homework, sleep, relaxing, etc.) and subjective (e.g., students' perceptions of academic stress) measures of stress. The subjective statements were adapted from Bedewy and Gabriel's (2015) Perceptions of Academic Stress (PAS) Scale, which is based on the transactional model of stress (Lazarus & Folkman, 1984).

Given the decrease in workload demands, I expected no significant difference between the groups; however, both the qualitative and quantitative data revealed that the Contract Grading Group was significantly less likely to report stress from workload demands, even though there was no significant difference between the objective measures of well-being or stress between the two groups. While no question mentioned the grading contract, the treatment group overwhelmingly (85%) cited it as a key component in reducing their academic stress during the research unit. Additionally, even with a prior history of low grades, most (84%) fulfilled the contract's requirements. Overall, the contract grading group earned six times as many As and 2.5 times as many Bs as those in the conventional grading group.

While the pilot study was limited to 12th graders with prior experience, the promising results led to the full launch of contract grading in all courses in January 2020.

3.4 Research Design: Mixed Methods

The thesis study built upon the pilot study and maintained a mixed-methods design, as defined by Creswell (2015). To gather a manageable data-set, I employed a mixed-methods approach with an explanatory sequential design, generating data in three phases with the same cohort of participants. First, to understand general trends, I generated matched-pairs quantitative data from a preliminary pre- and post-survey, in which students reflected on their experience in January 2019 (with the reduced workload and analytic rubric) on the pre-survey and then on their experience in January 2020 (with the grading contract) in the post-survey. After comparing the survey responses, I conducted 40 in-depth, semi-structured interviews from March to May with students of all courses and course types to understand how and why the findings occurred. Finally, in May 2020, I collected final grades. Ultimately, integrating two forms of data into a single study provides "breadth and depth of understanding and corroboration" (Johnson, Onwuegbuzie, & Turner, 2007, p. 123) to "simultaneously broaden and strengthen the study" (Yin, 2006, p. 41).

3.4.1 Recruiting Process and Study Withdrawal

Recruitment began in December 2019 when the school's administrative assistant sent parents and students (n=1,236), with the exception of those enrolled in my courses, the respective information sheets. Parents and children had three months (December-March) to withdraw. Most (n=18) withdrew before the study commenced, and no reason was provided, and two were later withdrawn based on parental dissatisfaction with the

inclusive gender question, for a total of 20. Interview volunteers or their parents provided consent and could withdraw their data within one week of the recorded session. No interviews were withdrawn.

3.5 Data Generation

From January to May 2020, the data generation occurred in three phases—pre- and post-surveys; in-depth interviews, and grade collection—with parallel construction, measures, and instruments with the same cohort of participants.

3.5.1 Phase 1

On the first day of the semester in January 2020, 1,216 students received an email with a link to the Qualtrics pre-survey (see Appendices 2 and 3); at the end of February, students received another email with the post-survey. To complete the survey on their iPad, phone, or computer, students were given approximately 20-minutes in their English class. All students at the school are required to bring an iPad to class, and computers were available for students who did not have a charged or working device. While administering the survey, the teachers were directed to monitor the classroom but stand in such a place where they could not see students' devices. This allowed students who did not want to participate to use their devices while preserving their anonymity. To minimize any situational pressure to give socially desirable answers, data generation was confidential and anonymous. In order to pair that data sets, the survey asked students to supply their identification (ID) numbers, but students were informed on the survey that I could not

match an ID number with a student's name. Finally, the surveys were mobile-friendly, allowing students to choose a more discrete phone or tablet.

3.5.2 Phase 2

After analyzing the quantitative data, I conducted 40 semi-structured interviews with students of all course types (i.e., accommodated for learning disabilities, regular, and honors/AP) and courses (e.g., English 1, Honors English 1, etc.). The primary goal for the interviews was to generate a detailed understanding of students' experiences in context and investigate what role, if any, the grading contract had on significant findings related to stress, self-worth protection behaviors, and motivational approach. Each interview lasted, on average, 45 minutes; a few interviews with first-year students without prior experience were shorter (around 35 minutes), while several interviews with returning students lasted over an hour. All participants received a debriefing sheet with mental health resources at the end of the survey and again at the end of the interview via email.

3.5.3 Phase 3

The final phase included collecting students' grades under the analytic rubric in 2019 and under the grading contract in 2020. Unfortunately, participants could not access last year's grade in the electronic grade book, and the majority of interviewees could not recall the exact percentage they earned in 2019 or 2020. With participants' consent, the English Department supplied a list of their students' ID numbers and their grades for

2019 and 2020 in order to compare their academic performance under conventional grading and the grading contract.

3.6 Participants

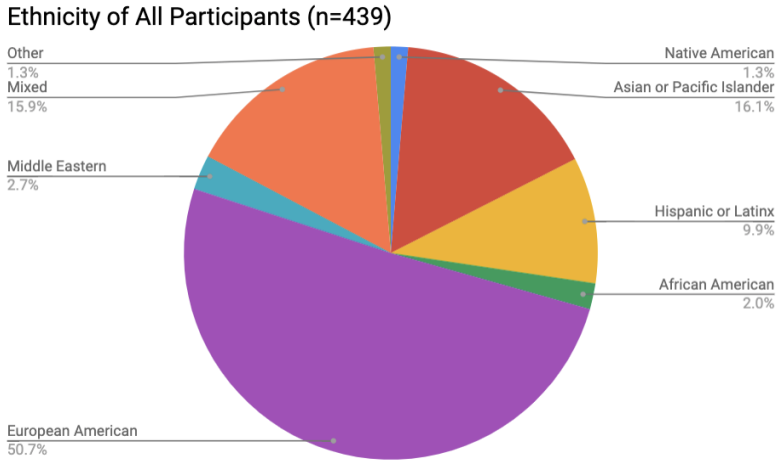
Participants (n=439) were secondary students (grades 9-12, ages 13-19) at a private, religious, college-preparatory school in the Western United States completing the annual five-week research paper unit; 284 had prior experience with the research paper unit, and 155 were first-year students completing the project for the first time at the school (140 entered as 9th graders and 15 transferred from another high school for grades 10-12). Additionally, 241 were in regular English, 160 were enrolled in an honors-level or AP course, and 49 students were diagnosed with learning disabilities and received academic accommodations. Each participant was enrolled in an English course using the grading contract.

3.6.1 Demographics

Of all participants, 226 identified as female, 210 identified as male, one as non-binary, one as transgender male, and two as "other." The study's ethnic sample is somewhat more diverse than the school (see Figure 3.1, top left). Tenth graders were the least represented grade level (see Figure 3.1, top right) due to the fact that I am a 10th grade teacher with courses that were excluded from the study.

Figure 3.1

Participant Demographics



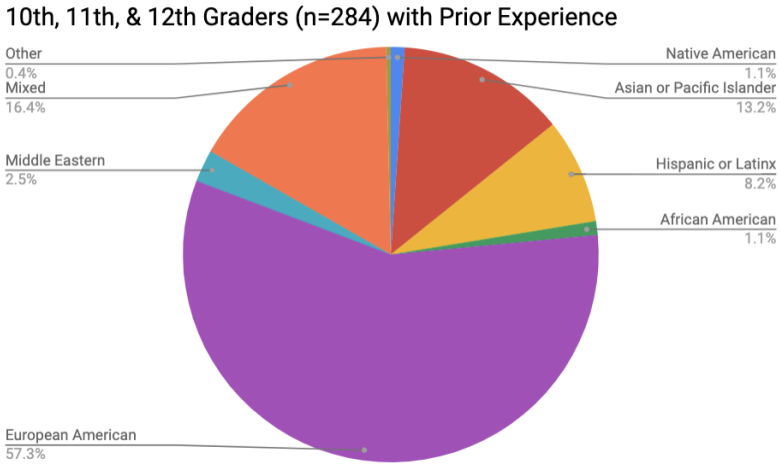
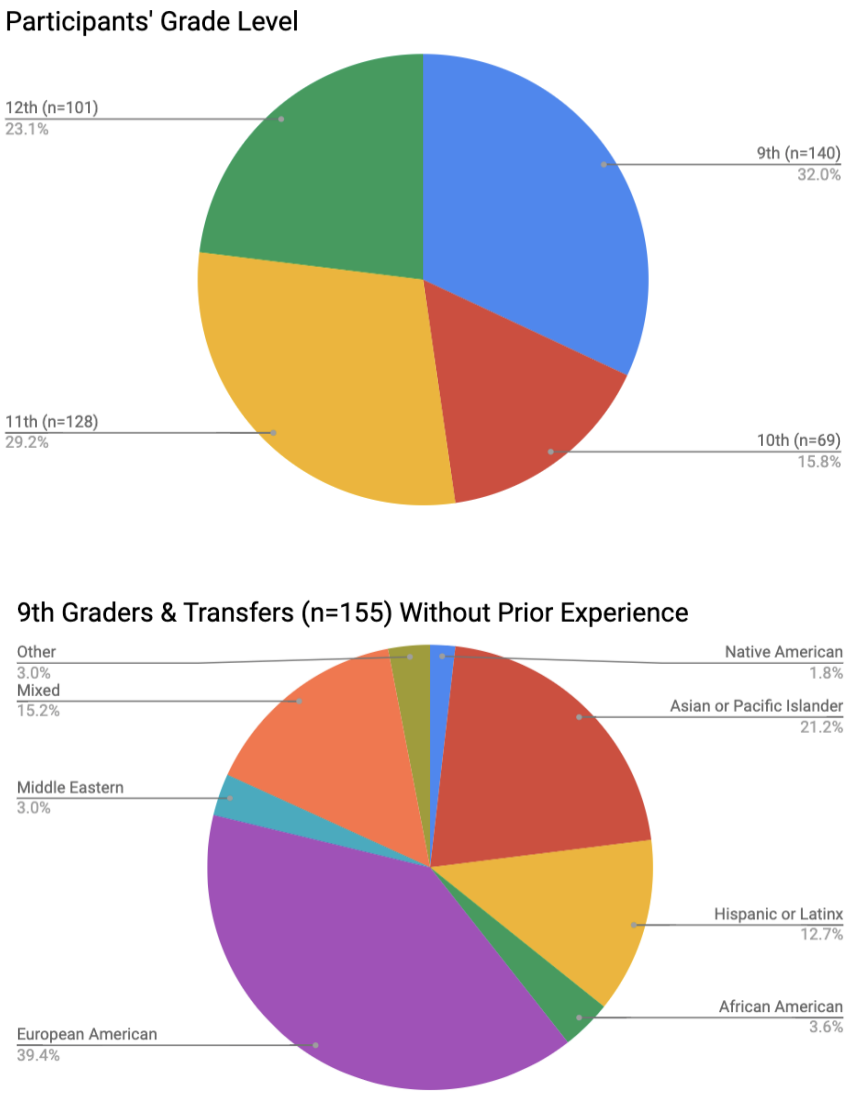


Figure 3.1 Continued



Note. The pie charts reveal participants' ethnic identity and grade level.

3.7 Ethical Consideration

This research was conducted with full institutional consent from the administration. In December 2019, I received formal ethical clearance from the Department of Educational Research's Research Ethics Officer and my supervisor before beginning the pilot work,

and I followed all guidelines outlined by BERA (2011). To protect participants' confidentiality and anonymity, the name of the school and the names of all participants have been anonymized; however, on the information sheets to both parents and participants and at the start of each interview, participants were informed that I am a mandated reporter and legally required to file a report if I receive information that a child is a threat to themselves or someone else.

I worked to mitigate the impact of my positionality and the unbalanced power relationships as a teacher who is researching the adolescents at her school by excluding all of my courses as well as my courses in 2018-2019 while I was conducting the pilot study. Care was taken to ensure that no student was pressured to participate. Finally, students were informed that their identities were anonymous unless they volunteered for an interview and supplied their email address. This also minimized pressure to participate or provide socially desirable answers.

Finally, I have sought to provide full transparency about my behavior and practical activity during this process, allowing readers to "understand the researcher's role in the construction of the knowledge that forms the centrepiece of the report" (Scott & Usher, 2010, p. 128). I also recognize that my positionality also brings the potential for my personal bias; thus, I pursued this research with a posture of reflectivity, "a continuing mode of self-analysis (Callaway, 1992, p. 33), and engaged with my supervisor about my claims and assumptions to help negate the influence of biases during the research process.

3.7.1 Data Ethics

Data was secured on a password-protected computer without any identifying information about students or the institution.

3.8 Research Instrument

3.8.1 Validity and Reliability

To best answer the research questions, four psychometrically sound scales, each tested for validity and reliability, were adapted for the research instrument. First, the stress appraisal scales were selected for their suitability for adolescent participants and their alignment to the transactional model of stress. Carpenter (2015) identified only five reliable and valid scales for measuring cognitive appraisal as theoretically described, including the Stress Appraisal Measure (SAM) and Primary Appraisal Secondary Appraisal (PASA). The third scale, Perceptions of Academic Stress Scale (PASS), came out the same year as Carpenter's analysis and thus was not included, yet it is also based on the transactional model of stress. Finally, the Self-Worth Protection Scale (SWPS) was employed for students with prior experience with the research unit (Thompson & Dinnel, 2007).

Each scale appears to be psychometrically sound after item-selection, analysis, and validation procedures. Finally, each scale was found to have an acceptable internal consistency. Determined by Cronbach's alpha, the PASS has an acceptable internal

consistency reliability of 0.7 (Bedewy & Gabriel, 2015) and the PASA has a reasonable to good internal consistency for each dimension: threat (.83), challenge (.63), self-concept (.81), control expectancy (.77), primary appraisal (.80), and secondary appraisal (.74) (Gaab et al., 2005). Using an intraclass correlation coefficient, the Self-Worth Protection Scale is acceptable: Ability Doubts subscale (.91), Importance of Ability subscale (.85), and Avoidance Orientation subspace (.88) (Thompson & Dinnel, 2003). Finally, the SAM has acceptable reliability estimates for each of the six appraisal dimensions: threat (65–.75); challenge (.66–79); centrality (.84–.90); controllable-by-self (.84–.87); controllable-by-others (.84–.85); uncontrollable (.51 –.82); and stressfulness (.75–.81) (Peacock & Wong, 1990). See instrument in Appendix 1.

3.8.2 Scale for First-Time Students

For first-year participants (n=155), the survey focused on anticipatory stress from the assessment. They received a 28-question Stress Appraisal Measure (SAM) created by Peacock and Wong (1989) that was adapted to reference the research paper unit. Peacock and Wong write that "a cognitive-relational view of stress demands the assessment of conceptually important appraisal dimensions" (p. 234); thus, the SAM measures overall perceived stressfulness and taps into six appraisal dimensions mentioned in the previous section. See instrument in Appendix 2.

3.8.3 Scales for Students with Prior Experience

For students with prior experience (n=284), the survey focused on stress appraisal and self-worth protection behaviors. Three scales were adapted to say "research paper" instead of a generic term for schoolwork and then randomized: Perceptions of Academic Stress Scale (PASS) (Bedewy & Gabriel, 2015); Self-Worth Protection Scale (SWPS) (Thompson & Dinnel, 2003); and Primary and Secondary Appraisal Scale (PASA) (Gaab et al., 2005). The table below reveals the survey's scope (see Table 3.1):

Table 3.1

Survey Dimensions for Students with Prior Experience

Scale	Dimension	Number of Items
PASS	Stress from Time Constraints	3
PASS	Stress from Workload Demands	5
PASS	Self-Perceptions	3
SWPS	Avoidance Orientation	12
SWPS	Ability Doubts	11
SWPS	Importance of Ability	4
PASA	Threat Appraisal	4
PASA	Challenge Appraisal	6

Note. The table reveals the dimensions examined in the survey for students with prior experience.

3.8.4 Instrument Pilot

Two instrument pilot sessions were held after school. Pilot participants were recruited by their English classrooms. To volunteer, students simply needed to attend. No other incentives were offered.

During the first session, the focus group of volunteers (n=5) were given printed copies of the surveys and were asked to comment on questions that caused them to hesitate or required clarification. The group debriefed by reviewing the following:

- Are instructions clearly written?
- Is each question and statement easy to understand?
- Are the response choices complete?
- Should all questions be added or deleted?
- Do you have any suggestions for improving the format?

Overall, the volunteers reported that the surveys were clear. Based on the pilot group, the following changes were made to the instrument:

1. A seven-point scale was adopted with 1 as "very untrue for me" and 7 as "very true for me." Previously, two sections of the survey—the Academic Stress and the PASA statements—used a five-point Likert scale while the Self-Worth Protection statements used a seven-point Likert scale. After discussing it, the volunteers felt the seven-point scale helped them choose a more precise response; consequently, the seven-point scale was adopted for all three sections of the survey, as well as the Stress Appraisal Measure (SAM). While the five-point Likert scale is the most popular, Symonds first argued in 1924 that reliability is optimized with a seven-point scale. Churchill and Peter (1984) observed a positive relationship between reliability and the number of scale points. While more options may reduce the response rate and increase frustration levels (Babakus & Mangold, 1992), Malhotra (1993) found that seven may increase response rates. Additionally, in Lewis's (1993) study, seven options brought a stronger correlation in the results of the t-tests.

2. Additionally, for the SAM, the volunteers preferred statements to questions; thus, the SAM was adapted to suit adolescents by turning each item into a statement. The changes were reviewed by my supervisor.

The revised instrument was then field-tested in Qualtrics with a new pilot group on different devices. Volunteers (n=5) received the surveys via email and were asked to provide feedback on the interface, including the design and buttons as well as any questions. All participants finished the survey in 15-20 minutes. To debrief, we discussed the following questions:

- Are instructions for completing the survey clearly written?
- Can you easily navigate the survey and correctly use the buttons?
- Can you easily change (or "correct") your answers?
- Do you have any suggestions regarding the addition or deletion of questions, the clarification of instructions, or improvements in format?

Participants used three devices (e.g., phones, tablets, and computers). All participants found the survey to be clearly written and easy to navigate. Based on the focus group, the following changes were made to the instrument:

1. The "skip logic" was changed so that only students who agreed to an interview were asked to input their email address.

3.8.5 Interview Questions

Two interview guides (see Appendices 4-7) were developed based on the findings of each survey. Each used open-ended questions to generate descriptive, detailed data about students' perspectives and focused on how students "interpret their experiences, how they construct their worlds, and what meaning they attribute to their experiences" (Merriam, 2009, p. 5). In line with semi-structured interviews, questions were used flexibly (Merriam, 2009), although all topics were covered with all participants. Regular

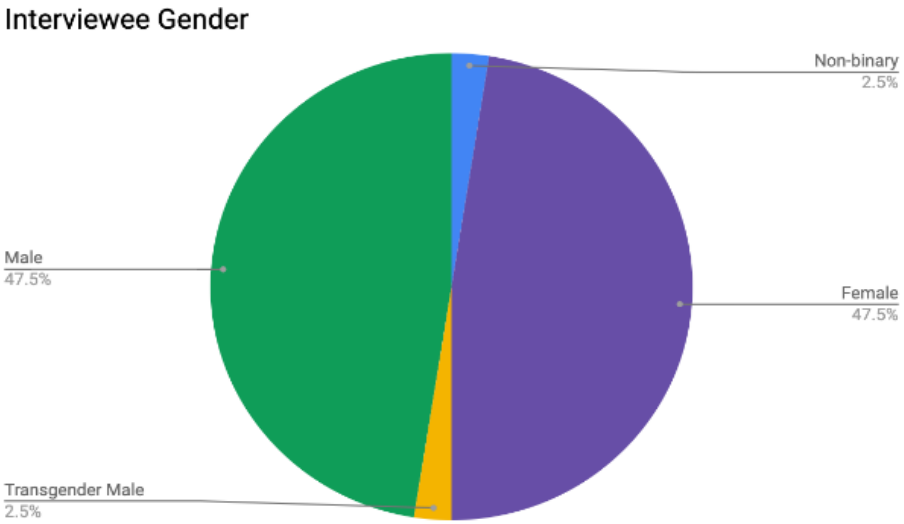
follow-up questions (e.g., can you tell me more about that? why do you feel like that?) were asked regularly to yield rich descriptive data (Merriam, 2009).

3.8.6 Interview Participant Sampling

Stratified sampling was used as a systematic method to choose interview participants from each grade level, English class, and class type. Both the pre- and post-surveys asked for interview volunteers to provide their email addresses. After stratifying the volunteers, interviews took place until English course, course type, gender, and ethnicity were as balanced as possible (see Figure 3.2). Most volunteers were European American (65%, n=26); others identified as mixed-race (n=6), Hispanic (n=4), and Asian or Pacific Islander (n=4). No African American, Middle Eastern, or Native American participants scheduled interviews although I sought an ethnic sample that represented the school's population. Additionally, two interviewees shared they were English Language Learners (ELL) studying abroad.

Figure 3.2

Interviewee Demographics



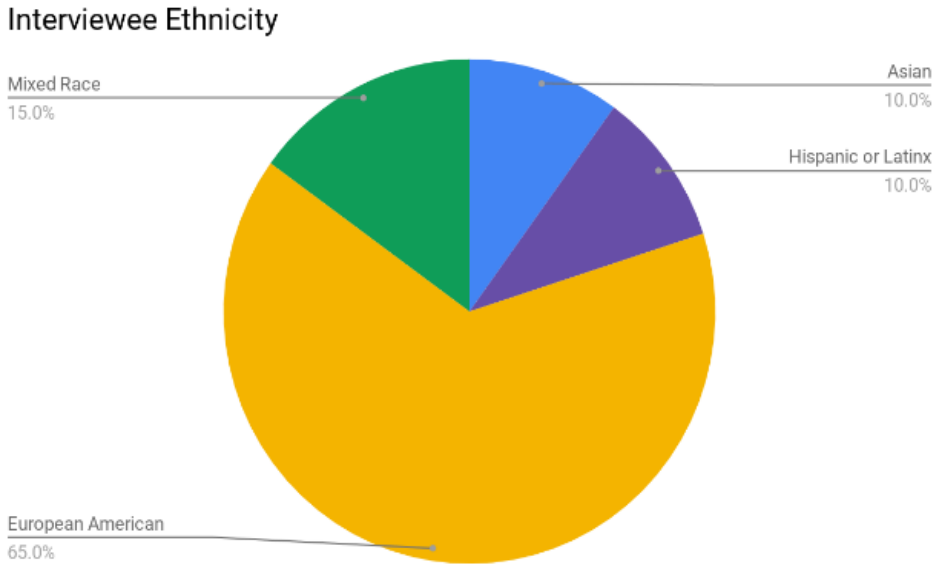
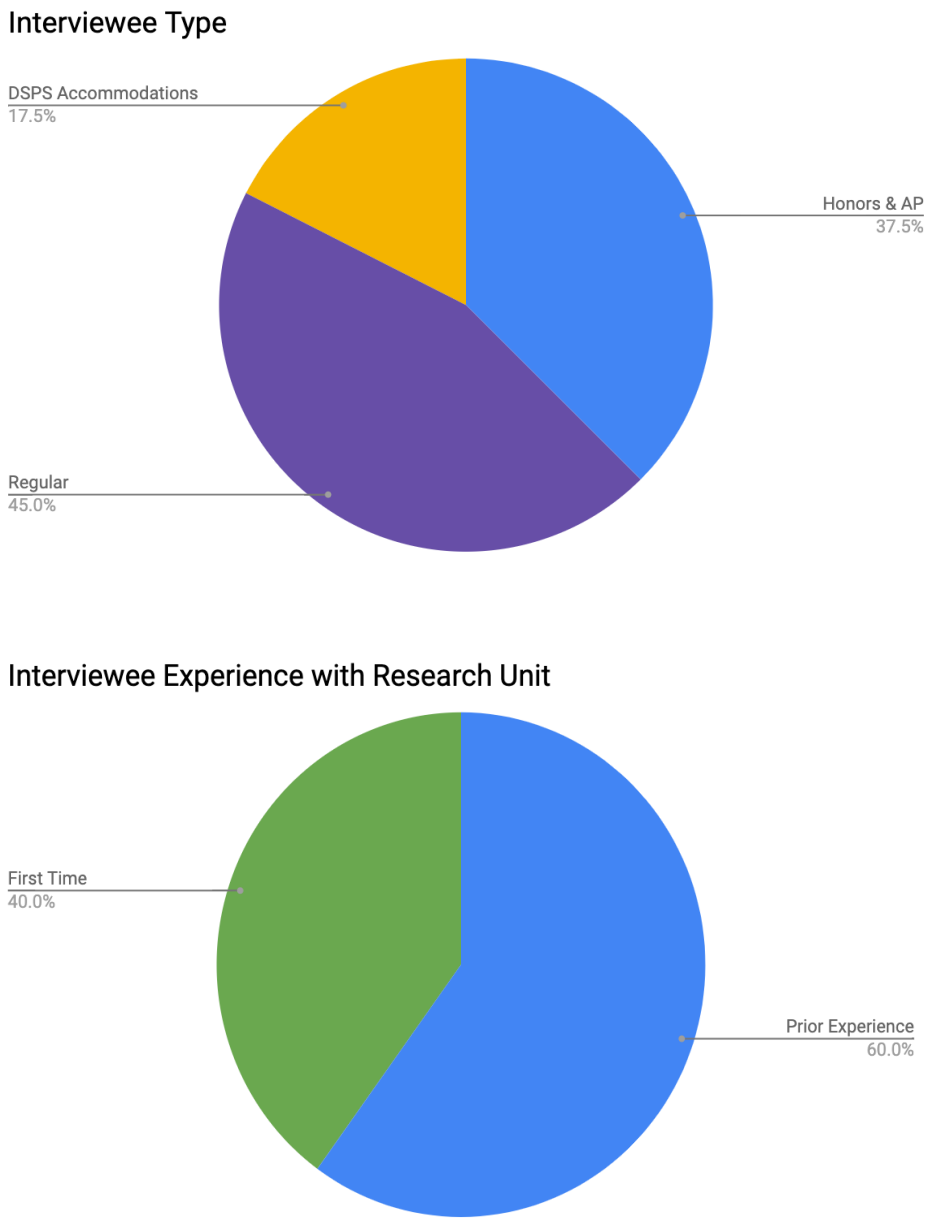


Figure 3.2 Continued



Note. The pie charts show the demographics of the 40 interview participants.

Of all interview participants, 47.5% (n=19) identified as female, 47.5% (n=19) as male, one as non-binary, and one as transgender male; 45% (n=18) were enrolled in regular-level courses, 37.5% (n=15) were in honors or AP courses, and 17.5% (n=7) were enrolled in the Department of Special Services and Programs with a diagnosed learning need, which included Autism Spectrum Disorder; Specific Learning Disability (SLD) in Reading, Writing, and/or Oral Expression; dyslexia, and Attention Deficit Disorder. Sixty-percent (n=24) had prior experience with the research unit, while 40% (n=16) were first-year students completing the project for the first time.

3.8.7 Interview Pilot

Both interview guides were piloted in March 2020. For students with prior experiences, a 45-minute pilot interview was conducted and then several questions were rephrased for clarity, shifted from closed- to open-ended, and then reordered for a more fluid progression. Due to the changes to the interview guide and a recording malfunction, the pilot interview was not used in the study; however, the pilot interviewee, Caleb (pseudonym), was reinterviewed a month after the pilot. The questions were then successfully piloted again with Sara, whose interview was used in the study.

For students without prior experience (grades 9-12), two pilots occurred. After the first 30-minute pilot interview, the guide was expanded to include a section on control and the questions were piloted again. The second pilot interviewee, Talia, was successful and included in the study.

3.8.8 School Closure

This study began on 8 January 2020 at the start of the spring semester, and the interviews commenced the week of 11 March 2020 when the World Health Organization declared the COVID-19 outbreak a pandemic. By 13 March, the local state government issued a stay-at-home order, which eventually shut down schools for the remainder of the school year.

Before the school's closure, six interviews had taken place on campus in my classroom with a window that allows others to see but not hear the details of the conversation. Each student with a previously scheduled interview was contacted to ensure the interview would not be an additional burden; as a result, several students opted to cancel their interview. The remaining interviews moved to a password-protected Zoom meeting to maintain students' safety and confidentiality. In an effort to ensure that participants had a safe and confidential space to undertake the interview, all Zoom interviews were calendared around participants' schedules to ensure that they had an appropriate space to conduct the interview.

To build rapport and ensure the interviewee was mentally and physically healthy, each interview began with a casual off-the-record conversation that included perceptions of well-being while at home, physical health, and experiences with remote learning. One interview was rescheduled based on the off-the-record conversation due to a reported

decline in physical health. In an effort to put interviewees at ease and encourage open and honest responses, I started the interview by sharing a script that emphasized the participants' confidentiality and anonymity. Interviewees appeared open, honest, and vulnerable. I observed that several students refrained from naming their teacher and two asked if it was permissible to use their teacher's name, which I attribute to a misunderstanding about anonymity in the opening script, rather than being guarded.

Additionally, participants were monitored for signs of emotional distress; if any student had shown such signs or turned the conversation about stress related to COVID-19, the interview would have been terminated out of care for the student. This did not happen, but one interview was briefly paused when a student was brought to tears thinking about her prior performance of earning Bs, not As, which emphasized the importance of grades to students in this study, as I will discuss in Chapter Five. Compared to the face-to-face sessions, the Zoom interviews had the addition of slight delays due to internet instability, yet overall, this interference did not impact the quality of the interviews.

3.9 Data Analysis

While the data generation methods have analytical integration as they "can reach into the realm of the other" (Yin, 2006, p. 44), each has a distinct and preferred analytical technique. The quantitative dataset was analyzed using IBM SPSS Statistics version 25. The data was checked for normal distribution, and then paired samples t-tests compared the mean of individual scores on the pre- and post-survey. Individual t-tests were run for

each course type: regular, honors, and courses accommodated for learning disabilities. To control for the familywise error rate, *p*-values were adjusted using the Benjamini-Hochberg procedure. Additionally, to analyze the subscale's dimensions, each scale was checked to ensure it had an acceptable (>.7) internal consistency, as measured by its Cronbach's alpha. If the subscale had an α of <.6, individual questions are reported. Questions for each valid subscales were combined to create a latent variable and then analyzed using a paired samples t-test. Descriptive statistics were used to understand the sample.

I transcribed all interviews in full using Trint, a subscription-based platform for transcribing and editing. Trint was selected for its ISO 27001 certification, which means security risks and threats have been examined and a comprehensive set of information security controls have been implemented to ensure the safety of confidential files. I coded the transcripts in NVivo 12, following Nowell, Norris, White, and Moules's (2017) six-step process: familiarizing oneself with the data; generating initial codes; searching for themes; reviewing themes; defining and naming themes; and then producing the report. To begin the coding process, I first identified relevant passages for the investigation using a broad-brush, or 'bucket' coding, technique. During the second pass, I took stock of the "diversity of opinions in each code, the volume of data and the relative importance participants assign to them while simultaneously coding to more discrete subcodes" (Jackson & Bazeley, 2019, p. 69). I took a deductive stance to generate initial

codes and subcodes that emerged from the data (see Table 3.2), which then led to the generation of broad themes covered in Chapters Four and Five.

Table 3.3
The Interview Codes and Subcodes

Code	Subcodes
Approach to Task	● Achieve Success ● Avoid Failure ○ Playing it Safe ● Grade Orientation ● Learning Orientation
Stress Appraisal	● Challenge Appraisal ○ Definition of Challenge ○ Fueled by Deadlines ○ Eustress Threat Appraisal
Table 3.3 Continued	
Stress Appraisal	● Challenge Appraisal ○ Definition of Challenge ○ Fueled by Deadlines ○ Eustress ● Threat Appraisal ○ Definition of Threat ○ Prior Threat ● Reduced Stress ○ From Contract ○ From Prior Experience ○ From Teacher Support ● Increased Stress ○ From Workload ○ From Final Deadline ○ From Uncertainty about Fulfilling Expectations ○ From Managing Courses ○ From Personal Circumstances
Contract	● Teacher Introduction ○ Efficacy Appeals ○ Student Control ○ Fear Appeals

	○ Teacher Control ● Choice of Effort ● Organization with Coursework and Calendar ● Clear, Direct, or Straightforward Path ○ Easier ○ Freedom ● Self-Assessment ● Relief, Comfort, and Reassurance ● Contrast with Rubric ○ Discomfort with Contract ○ Sense of Doubt and Worry with Rubric ○ Orientation Toward Learning ○ Orientation Toward Performance
Confidence	● Fear of Failure ○ Less ○ More ● Social Comparison ○ Less ○ More
Table 3.3 Continued	
Confidence	● Confidence ○ Decreased ○ Increased ■ From Contract ■ From Prior Experience ● Sense of Responsibility and Independence ○ Increased ○ Decreased ● Sense of Personal Control ○ Increased Decreased
Perception of Workload	● Perception of Workload ○ Reduced ○ Increased ● Perception of Deadlines ○ Accountability ● Perception of Time ○ Less Time ○ More time

	○ Enough Time ○ Not Enough Time
Environment	● Classroom ○ Positive Classroom ■ Collaborative ○ Negative Classroom Environment ● Home ○ Positive Home ○ Negative Home

3.10 Data Storage

Data was encrypted on a password-protected computer and will be stored for 10 years. No paper records were collected.

3.11 Summary

The study employed a mixed-methods approach with an explanatory sequential design to understand the impact of contract grading on adolescents' academic stress appraisal and self-worth protection behaviors. Great care was taken to minimize my positionality. Students without prior experience completed an adapted version of the Stress Appraisal Measure, while students with prior experience completed adapted versions of the Perceptions of Academic Stress Scale (PASS) (Bedewy & Gabriel, 2015); Self-Worth Protection Scale (SWPS) (Thompson & Dinnel, 2007); and Primary and Secondary Appraisal Scale (PASA) (Gaab et al., 2005). Each established scale, tested for reliability and validity, was selected for its alignment with the transactional model of stress and coping put forth by Lazarus and Folkman (1984). Phase one of data generation gathered a

large set of quantitative data, while phase two consisted of 40 semi-structured interviews to understand how and why the significant findings occurred.

Chapter 4: Perceptions of Stress, Self-Worth Protection Behaviors, and Achievement Motivation

4.1 Introduction

This chapter presents the findings on perceptions of stress (Research Question One) and self-worth protection behaviors (Research Question Two) thematically, integrating the quantitative and qualitative data for students of all grade levels (9-12) and course types (regular, honors, and accommodated for learning disabilities). While all high school students (n=439) completed a high-stakes assessment, worth 20% of their final grade, in their required English course, this chapter contrasts the experiences of two groups: returning students (grades 10-12) with prior experience (Prior-Experiencers, noted 'with PE') and first-time students (First-Timers, noted with 'FT') completing the assessment, as the name suggests, for the first time in grades 9-12. Unlike other assessments, the annual research unit lasts approximately five weeks with at least 825 instructional minutes dedicated to teaching ethical research.

Table 4.1

Research Paper Workload Requirements

	9th Grade	10th Grade	11th Grade	12th Grade
Word Count	1,200-1,500	1,500-1,800	1,800-2,100	2,100-2,400
Source Requirement	4	6	6	8

Note. This chart shows the requirements for the paper, which increase with each grade level. Each student in the study completed the longest paper they had written in high school.

In 2019, as mentioned in Chapter Three, the length of the project was reduced by 600 words across all grade levels (see Table 4.1), a change that the English Department surmised may be enough to reduce perceptions of stress. As previously mentioned, however, the pilot study (Ward, 2021) revealed that only 12th graders who received the grading contract reported less stress from workload demands.

The next section contrasts high schoolers' perceptions of stress under the grading contract compared to their prior experiences with conventional grading.

4.2 RQ.1 How does contract grading affect students' perceptions of academic stress?

Both grading contracts (see Appendix 8) were offered to students on the first day of the semester, serving as task-specific learning goals that emphasized each step of the research process. At all grade levels and course types, the majority of adolescents (75%, n=30) interviewed cited the clarity of the grading contract as central to reducing stress this year, alongside teacher support (48%, n=19) and prior experience (25%, n=10) at any grade level, including middle school (grades 6-8). Academic stress during conventional grading, the qualitative analysis revealed, often stems from seeking to ascertain what is expected, or "*figure out what the rubric is saying,*" as Ella (PE honors 11th) explained. Thomas and Rowler (1986) observed that as students prepare for assessments, "clarity of purpose is pronouncedly atypical" as they "rarely know with any precision what they are preparing themselves to do" (p. 30). For Ella, who said her favorite subject is science, the

contract "*adapted to [her] skills and then broke [the paper] down more for a different kind of learner*." She explained how it minimized stress:

> I wasn't as nervous because I'd never really had it laid out before, so this was probably the least nervous I've been out of all the years because everything was just presented, so it's kind of weird: all the requirements that were expected of were kind of given in the beginning and then I saw when everything was due, and then it wasn't a lot to do each time—it was little by little by little—and then every step set up the next step, so it made the next step easier.

While her perception may have been altered by prior experience, she credited the contract as critical to reducing her stress and making the process "*easier*." Like Ella, 56% (n=23) of those I interviewed used the words "easier" or "easy" when comparing this year to their prior experience or expectations for the project, as Natasha (PE regular 10th) stated simply: "*[The contract] made it easier and less stressful*." For Malcolm (FT honors 10th), the rubrics, particularly the qualities matrix, he had received prior to the contract were "*a little overly complex*," continuing, *"It kind of throws big words at you to scare you. But this one was really user friendly.*" Ultimately, the contract's clarity about task requirements and expectations ameliorated stress for students of all course types, as Thomas (PE regular 12th) revealed:

> If you hit the checkpoints that are listed on the A [contract] and you hit them well, you know you're going to get an A or at least very close to it. You'll be somewhat expecting the grade you're gonna get; whereas when I throw everything I have towards one rubric, I just hope that it gets interpreted well or, you know, that I hit as much of it as possible and hopefully get an A, but you don't really know.

In this way, students described experiences in which luck and mystery surround the conventional grading process: expectations are often vague or concealed and the outcome depends upon the teachers' interpretation of the rubric, which varies from instructor to

instructor. For Thomas, putting forth his best effort by "*throw[ing] everything. . . towards one rubric*" led to the insecure position of *hoping* he did it correctly. Mason (PE accommodated 12th) was more direct, revealing that under such conditions, students couldn't "*have as much confidence*." Both reveal that much of the insecurity of assessment is revealed in focusing on performance outcomes. While Chapter Five focuses on outcomes, other research has revealed that performance-orientation predicts self-handicapping and maladaptive coping that can increase stress and fear of failure (Martin, Marsh, & Debus, 2001) and lead to a loss of motivation and self-worth (Dweck & Elliott, 1983; Grant & Dweck, 2003). Under the grading contract, however, high schoolers reported a statistically significant reduction in stress, fear of failure, avoidance orientation, which will be explained in the coming sections, beginning with adolescents' perceptions of stress.

4.2.1 Reduced Perception of Stress

Students with PE were significantly less likely to report that the workload was "very stressful" (M=4.17, SD=1.622) compared to conventional grading (M=4.82, SD=1.445), t(283)= 6.070, p=.000. To describe the grading contract's impact on their well-being during the unit, many students (33%, n=13) used the words "comfort," "relief," or "reassurance, as Lucas (FT honors 9th) observed: "*Everyone's on the same page with this contract, which really kind of brings a sort of comfort to everyone knowing*" that both students and teachers have the same expectations, continuing—

> I find the teachers in the past give you a student rubric and then they have a separate teacher rubric, I would assume, and I feel like this contract is something that the students and the teachers both look at. I mean, it's what the teacher is

> using to grade, so there's nothing different on the contract—it's all there, and you signed off on it, so you just kind of go down the line and if you have it all, then it should be a good thing, so it's just reassuring that you're going to get a good grade and kind of eases the stress.

For Lucas, ambiguous expectations led to the belief that his teachers had separate evaluation standards that were not shared with the class but would be used to assess his work. While Brookhart (2018) noted that rubrics are not always shared with students, the high schoolers in this study received the grading contract on the first day of the semester. Madison (FT regular 9th) said the contract made her "*feel better because it was more of a direct path than normal grading."*

Those with prior experience were significantly less likely to say they "feared failing the research paper this year" (M=4.4, SD=1.931) compared to conventional grading (M=4.85, SD=1.952), t(279)=3.531, p=.000 (see Table 4.6). Stress can impact performance by leading the student's attention to negative thoughts (Angelidis, Solis, Lautenbach, Van der Does, & Putman, 2019). Anna (FT regular 9th) explained:

> It made me feel a lot more at ease just seeing that contract. I really needed something like that. I needed that encouragement from [my teacher], because this is my first high school paper and I came from a smaller middle school, so I've never really done anything like this.

Ultimately, the qualitative data revealed that adolescents' psycho-emotional experiences improved as a direct result of clear, direct expectations, corroborating Fairbanks's (1992) finding that the contract reduced perceptions of stress. While this finding may be self-evident, Thomas and Rohwer (1986) observed that "ambiguity of purpose is more often the rule than the exception" (p. 30). Teachers regularly require students to ascertain

implicit expectations, a taxing task that can exacerbate stress levels, lower confidence, and increase maladaptive coping, according to the students I interviewed. The grading contract, however, presented "just what information they will need, how it must be organized, or the specific form in which it must be exhibited in order to satisfy the instructor's criteria" (Thomas & Rohwer, 1986, p. 30). As Wormeli (2006) wrote, "Some students are better at guessing what's on the teacher's mind than others. No student should be punished for not being a good mind-reader" (p. 23). The impact extended to students' perceptions of workload demands, the focus of the next section.

4.2.2 Reduced Stress from Workload Demands

The size of the workload increased for students with PE since workload demands increase each year, yet the dimension of academic stress from workload demands was statistically significant as students reported less stress under the grading contract (M=3.43, SD=1.22) compared to conventional grading (SD-4.06, SD=1.12), t(278)=7.59, p=.000 (see Table 4.2).

Table 4.2

Workload Demands Subscale for Prior-Experiencers

	Conventional Grading		Contract Grading		*n*	*t*	*p*	Cohen's *d*
	M	*SD*	*M*	*SD*				
Stress from Workload Demands	4.06	1.12	3.43	1.22	278	7.59	.000	.538

Note. The internal consistency for workload demands was good: α=.803 for the pre-survey and α=.849 for the post-survey.

The qualitative data then revealed how and why adolescents' perceptions of stress shifted: the majority of adolescents interviewed (75%, n=30) described as "clear," "concise," "direct," "exact," and/or "straightforward" that accompanied the grading contract, as Caleb (PE honors 11th) explained:

> What decreased my worriedness was that the expectations were clear and that it was listed exactly what grade that I could possibly achieve instead of it somewhat being up in the air as to like—I feel like it removed a lot of the confusion of the English papers can have some times where like I was sometimes you don't really know if what you're turning in is what the teacher is expecting. But the contract removed part of that because it felt like the expectations were clear.

Confusion led to valuable time dedicated to figuring out "*what the teacher is expecting*." The workload demands not only *felt* like less—many actually dedicated less time to doubts, worries, and fear and more to working on the project, which also impacted what Macan (1994) called the subjective experience of time, the focus of the next section.

4.2.3 Reduced Stress from Time Constraints

The quantitative data revealed that students with PE were significantly less likely to experience stress from time constraints under the grading contract (M=3.81, SD=1.355) compared to conventional grading (M=4.07, SD=.823), t(281)=2.84, p=.005, although the effect size was small (see Table 4.3).

Table 4.3

Time Constraints Subscale for Prior-Experiencers

	Conventional Grading		Contract Grading		*n*	*t*	*p*	Cohen's *d*
	M	*SD*	*M*	*SD*				
Stress from Time Constraints	4.07	.823	3.81	1.355	281	2.84	.005	.232

Note. The internal consistency for the time constraints subscale was acceptable: α=.628 for the pre-survey and α=.640 for the post-survey.

The qualitative data revealed that the presentation of the contract led students of all grade levels and course types to reappraise time constraints. For example, Jax (PE accommodated 11th) described how "*the wording of the contract*" and "*the simplicity of the work*" reduced stress from time constraints, continuing—

> It really made me understand what to do and what the goals were. I felt like I've gotten more motivation this year just because the expectations on the paper are easier and more manageable than in the past.

Many students, like Jax, were less likely to engage in avoidance behaviors as a result of evaluative threat or feelings of uncertainty. Similarly, Riley (PE regular 11th) described his experience prior to the grading contract that led to task-avoidance as a result of perceiving the workload as demanding:

> Like I told you previously, I work when there's less work. I'm more inclined to do it earlier, so when I felt like this was less work from previous years, especially freshman year, and I was more inclined to be responsible

Riley noted that he "*works when there's less work."* Although he completed more work this year, the assessment felt easier under the grading contract. Understandably, students

will often perform easier tasks first, as researchers have confirmed (McCown, Petzel, & Rupert, 1987). With the perception of less work through the presentation of the grading contract, students were less inclined to avoid the task. By providing a clear path to the goal, students' task-oriented effort increased, thus corroborating prior work with contract grading and increased motivation (Polczynski & Shirland, 1977; Parks & Zurhellen, 1978; Callahan, 1979).

The qualitative analysis revealed that the majority of students (80%, n=32), including 70% of FT students and 87% of students with PE, perceived having enough time to completed workload demands and meet deadlines this year. While students with PE had a comparable amount of time to complete a longer project both years, many (33%, n=8) perceived having more time under the grading contract, yet struggled to pinpoint why, as revealed through the "I don't know" in their responses:

> *Mia* (PE accommodated 11th): I don't know if they gave us more time this year, but I feel like it was a lot more expanded than last year. I just feel like that. I just feel like last year it was kind of just really fast. And this year felt like a longer period of time that we got to work on it.
>
> *Isabella* (PE regular 10th): I felt I had a lot more time this year than last year. I don't know if it's because of the honors pace or if it's just freshman year they throw it at you, but I definitely think we had a lot more time to do individual parts of it, so it felt more like we were dispersing to work.
>
> *Noah* (PE regular 11th): I feel like I had a lot of time this year. I don't know. Maybe that's just the way [my teacher] had us do it, where she had us do it in chunks, so we did part one and then part two and then we just started editing the rest of the time. It felt like I had a lot of time.

Two students (8%) recognized that they completed more work this year but still noted the time to work on each part felt longer, including Kevin (PE accommodated 10th):

> The workload was obviously more, but it was honestly easier because we had more time and we were able to split everything up to where we'd have like a day or two on each thing, which was more than enough time, and, you know, we were able to perfect each individual thing, so I don't think that workload was a problem. It definitely was more, but we had more time, so it kind of canceled it out. It was even easier.

For many, doubt and worry about meeting expectations led to second-guessing, which added time to last year's paper. This year, however, Natasha (PE regular 10th) said, "*I worked hard on it, but I spent less time on it because of how I did it*." For many, the time felt longer as students spent fewer minutes re-writing or worrying about expectations because, as Thomas (PE regular 12th) said, *"it was very clear*" with "*less guesswork*."

The contract appears to have maximized efficiency by reducing the task's overall cognitive load, as defined by Clark, Nguyen, and Sweller (2006). As previously reported, the workload not only *felt* like less, but many (50%, n=12) reported doing less work this year despite completing a longer paper, like Eric (PE regular 11th) revealed:

> This year probably actually felt shorter because I thought it was a little easier. I didn't have to go back and change as much, so maybe that felt better. Same thing with the workload. I think it was easier this year because I had a contract, so I kind of knew what to expect and stuff, so I don't have to go back and second guess as much. But I don't think it felt much longer. If anything, it felt, like, shorter—I just feel like I didn't do as much work. In the past, I was going back and changing things even when I didn't have to change it. In the past, I think I wasted a lot of time sometimes just second-guessing myself over and over and over again.

Emotions can color one's perception of time (Droit-Volet & Meck, 2007). In other words, worries, doubts, and fear make the time feel longer (Lay & Schouwenburg, 1993), but the grading contract mitigated these issues by providing clarity that enabled students to work confidently and 'do it right' the first time, as Evan (PE AP 11 th) explained:

> Last year, when we did the rubric grading, there's like a hundred different things, and I'm trying to look at each one, going through and being stressed out, like, "Do I have this or do I not?" So, with the [contract], I was just able to look at my paper and be like, "Yeah, I did these things," and then just, like, forget about it and be done.

Additionally, other students with PE (43%, n=10) perceived less time this year; however, the perception of less time was overwhelmingly positive (90%, n=9):

> *William* (PE AP 12th, ELL): I felt like I had less time because, first of all, I had so much to say, but also because I was always enjoying what I was writing and what I was researching so much. Because I wasn't caring about grades as much, time went by so fast.

> *Olivia* (PE regular 12th): I feel like the research paper process went super quick while last year felt like a long time. I kind of liked it because it was just kind of quick.

Droit-Volet and Meck (2007) reported that "our feeling for time is fundamentally inseparable from our subjective experience of the environment" (p. 512). Stressful experiences lengthen the subjective perception of time, as Droit-Volet and Gil (2009) document with several pioneering studies, writing, "The stressful conditions increased the arousal level, which in turn accelerated the clock speed, thus producing an overestimation of the duration" (p. 1945). Most recently, Sarigiannidis, Grillon, Ernst, Roiser, and Robinson (2020) found that increased anxiety increases one's perception of the speed of time. Before the implementation of the grading contract, the students in this study reported significantly higher levels of stress. Unfortunately, decreased well-being is now a nationwide trend in the US (Mental Health America, 2019; Burstein & Greenfield, 2019; Twenge, Joiner, Rogers, & Martin, 2017). One's level of distress depends upon their perception of the potential stressor (Lazarus & Folkman, 1984). The contract's clear

presentation of workload demands positively altered students' perception and significantly decreased distress. This can be explained by Balcetis and Dunning's (2006) finding that motivational states influence perception. In 2010, they concluded that desirable objects appear closer to the perceiver and serve to energize an individual to fulfill the goal. In this study, the clear steps and the promise of a good grade led adolescents to perceive workload demands and time constraints as more manageable led students to reappraise the project as challenging, not threatening to their self-worth or grade, the focus on the next section.

4.2.4 Increased Challenge Appraisal

Under the contract, students with PE reported significantly more challenge appraisal compared to conventional grading (see Table 4.4). They were significantly less likely to feel "worried because the research paper does not represent any threat for me" (from M=3.5, SD=1.471 to M=3.86, SD=1.585) t(274)=-3.599, p=.000; significantly less likely to perceive the paper as "very unpleasant" (from M=5.15, SD=1.5 to M=4.74, SD=1.656) t(277)=3.936, p=.000; and significantly more likely to report that they could think of "lots of solutions to help [them] succeed" (M=4.9 to SD=1.256 to M=5.16, SD=1.178) t(273)= -2.415, p=.036. They were also significantly less likely to report that the research paper "scares" them (from M=3.77, SD=1.456 to M=3.30, SD=1.574) t(274)=3.237, p=.001. Finally, they were significantly more likely to report that they "know what I can do to succeed" (from M=4.86, SD=1.297 to M=5.17, SD=1.223), t(274)=-3.428, p=.021. The effect size, however, was small.

For some (20%, n=5), a sense of personal control fostered the challenge appraisal, as illustrated through Jackson (PE honors 10th):

> *Interviewer*: Did the demands of the workload this year feel more challenging or more threatening?
> *Jackson*: Challenging, just because I had more control over it.
> *Interviewer*: Did you feel like you had control over the workload?
> *Jackson*: Yeah, I felt I did, because, like I said before, I could follow the guidelines and make sure I was meeting all the requirements. It's challenging—it requires more effort from you—but I don't think it's like something to be afraid of.

Table 4.4

Challenge Versus Threat Appraisal Matched-Pairs T-Test Results for Prior-Experiencers

Statement	Conventional Grading		Contract Grading		*n*	*t*	FDR-adjusted *p*-values	Cohen's *d*
	M	*SD*	*M*	*SD*				
I do not feel worried because the research paper does not represent any threat for me.	3.5	1.471	3.86	1.585	274	-3.599	.000	0.235
I find the research paper very unpleasant.	5.15	1.5	4.74	1.656	277	3.936	.000	0.259
I can think of lots of solutions to help me succeed on the research paper.	4.62	1.497	3.40	1.502	273	-3.66	.000	0.815
The research paper scares me.	3.77	1.456	3.30	1.574	274	3.237	.001	0.309
For the research paper, I know what I can do to succeed.	4.86	1.297	5.17	1.223	274	-3.428	.021	0.246

Note. This table displays the t-test results for challenge and threat appraisal for Prior-Experiencers.

Students defined a "threat" at school as a task that feels unachievable, even with effort, and then increases hopelessness and motivation to avoid failure. Conversely, a "challenge" may stretch their resources and abilities but is achievable with hard work; when they can see success and/or growth on the other side, the change leads to increased motivation to achieve success. William (PE AP 12th, ELL) said, "*A challenge pushes my potential and kind of drives me*." Overwhelmingly, students across all grade levels and course types (98%, n=39) appraised the paper as challenging, not threatening, as Olivia (PE regular 10th) and Caleb (PE honors 11th) explained:

> *Olivia* (PE regular 10th): I feel like the research paper is definitely challenging—it's not your average essay, I guess. I don't know. I just feel the research paper has a stigma around it, and it's kind of a big thing, at least for me, but I feel like the research papers are more challenging rather than threatening.
>
> *Caleb* (PE honors 11th): This year felt more challenging, and I didn't really feel this year that I was unable to achieve the grade of A. The larger issue for me was just if I really wanted to put in that much effort into it.

The sense of manageability and control over the workload demands fostered a challenge appraisal that increased self-efficacy and subsequently task-oriented effort. The contract's alignment with the daily deadlines "helped" students stay organized, making it "easier" and "less stressful." Aside from the grading contract, no other change occurred this year, yet students with PE (42%, n=10) also perceived a positive difference in pacing compared to last year:

> *Noah* (PE regular 11th): I think just the way that it was distributed over the course this year was different, because it felt like last year was just, "Go, go, go, go, go." This year, I got to chill out a little bit.

> *Ella* (PE regular 11th): This year, I think, which was different from all my years in the past, was just kind of this step, this step, and this step. . . . and I thought that made it easy.
>
> *Natasha* (PE regular 10th): I felt like I had more time with the way that we had to turn stuff in. We had to turn one thing in every class instead of like a rough draft one day and then the final draft next, like it was just different. The deadline felt farther away.

For students of all levels and course types, the contract positively changed students' cognitive appraisal of the task (Putwain et al., 2016) by explicitly connecting the due dates, classwork, and expectations in students' minds. This suggests that as students felt that their abilities were better matched with the task, evaluative threat and maladaptive coping behaviors decreased, which will be explained later in this chapter.

The next section examines the stress appraisal of FT students, who reflected on their expectations for the assessment in the pre-survey.

4.2.5 First-Timers' Reappraisal

The primary appraisal of the assessment for First-Timers was influenced by what they had heard from their peers and their prior experience with research-based writing. The interviews revealed that their peers, such as upperclassmen on their athletic teams, had warned them about the assessment. Despite what they had heard, however, their appraisal positively (68%, n=11) shifted by the end of the unit, as illustrated through Talia (FT honors 10th):

> *Interviewer*: Did you believe that it would be a scary experience?
> *Talia* (FT honors 10th): At first, yes, because I've never worked on a research paper, but after I was like, "Oh, it's not that scary."

The quantitative analysis revealed that FT adolescents were significantly more likely to report a challenge appraisal after their experience with the grading contract; however, the effect size remained small (see Table 4.5). They were significantly less likely to say the research paper "scares me" (M=3.76, SD=1.742) compared to their expectations (M=4.38, SD=1.772), t(154)= 4.287, p=.000. They were significantly less likely to report that the assessment makes them "feel anxious" (from M=4.97, SD=1.644 to M=4.63, SD=1.67), t(153)= 2.459, p=.019. They also reported feeling significantly less "threatened by the research paper" (from M=3.52, SD=1.546 to M=3.15, SD=1.596), t(154)=2.546, p=.025. Like their peers with prior experience, they were significantly more likely to perceive being "challenged" by the paper under the contract (M=5.12, SD= 1.285) compared to their expectations (M=5.37, SD=1.247), t(154)=2.185, p=.05.

Table 4.5
Challenge versus Threat Appraisal Matched-Pairs T-Test Results for First-Timers

Statement	Perception Before Unit		Perception After Unit		*n*	t	FDR-adjusted *p*-values	Cohen's *d*
	M	*SD*	*M*	*SD*				
The research paper scares me.	4.38	1.772	3.76	1.742	154	4.287	.000	.353
The research paper makes me feel anxious.	4.97	1.644	4.63	1.67	153	2.459	.019	.205
I am threatened by the research paper.	3.52	1.646	3.15	1.596	154	2.546	.025	.228
I can think of lots of solutions to help me succeed on the research paper.	4.9	1.256	5.16	1.178	154	-2.415	.043	.214

The research paper will challenge me.	5.37	1.247	5.12	1.285	154	2.185	.05	.197
I do not feel threatened by the research paper.	3.98	1.694	4.33	1.609	154	-2.139	.034	.212

Note. This table reveals the t-test results for challenge and threat appraisal for First-Timers.

While the effect sizes were small, the quantitative analysis also revealed that First-Timers were significantly more likely to report that they were able to "think of lots of solutions to help me succeed on the research paper" under the contract (M=5.16, SD=1.178) compared to their expectations (M=4.9, SD=1.256), t(154)=-2.415, p=.0425. In addition to the clarity of expectations, FT students also credited their perception of "*an abundance*" of help and resources to assist them with the paper (Malcolm, FT honors 10th). The overwhelming majority of FT students (94%, n=16) felt supported, with sufficient resources to succeed. As Sienna (FT regular 9th) said, "*There was a lot of help everywhere*," which made the process easier, better, or not as "*bad*" as expected for many FT students (65%, n=11), as these students revealed:

> *Bailey* (FT accommodated 9th): I thought it was going to be a little bit harder than it actually was. At first, I was kind of like, "Whoa, I'm going to write all of that?" Then like, once I actually started, I was, "Oh, this isn't that bad."
>
> *Sienna* (FT regular 9th): I've heard before that it was kind of hard. But then like once we actually started that, my teacher explained, and she went through every step, so it was easier than what I've been told.
>
> *Bethany* (FT honors 9th): All of my volleyball teammates had been telling me about it, that it's really stressful, so I was nervous going into it, but it wasn't as bad as everyone said.

Notably, despite warnings from upperclassmen about the difficulty of the assessment (e.g., "I've heard before that it was kind of hard"), the perception of FT students still

shifted upon starting the project (e.g., "It was easier than what I've been told"). While the workload was reduced the previous year, students with PE still shared that it was "really hard"; however, upon introduction of the contract, appraisal shifted for the adolescents in this study, both those with prior experience and those completing it for the first time.

Like their peers with prior experience, First-Timers (50%, n=8) also noted the impact of the project's scaffolding—the alignment of the calendar, deadlines, and contract—on their reduced stress:

> *Daniela* (FT honors 9th): I was, like, so scattered, disorganized, and for me, I like to have a list in front of me in order to follow it. I guess it's just easier for me, seeing the contracts and having it on one paper, both sides, you can flip easily that they're both basically the same 10 points. It was a lot easier so that helped me stay organized, so it definitely helped.
>
> *Anna* (FT regular 9th): [The contract] really helped me stay organized, especially because I'm not usually a super organized person. But it really helped me and helped me just make clear and just make sure I was on the right track.
>
> *Gabriel* (FT regular 11th): One thing that did help me feel a little less stress was the organization of how we were getting it done, so like the calendar of we're doing an intro on this day, we're doing our thesis, this day we're doing topic sentences. It made it a lot easier to not have to do everything at once, like, "Go have fun. Write an entire essay." It's organized, and they're giving you guidance on when to do it.

The organization and schedule for the project has been consistent throughout the years, yet the alignment to the contract's expectations helped First-Timers and students with PE alike perceive the project as easier and more manageable, although some adolescents I interviewed found the new tool uncomfortable, as I will explain in the next section.

4.2.6 Positive and Negative Perception of Flexible Guidelines

> *Micah* (FT accommodated 9th): I thought of the instructions not as rules but as guidelines.
> *Interviewer*: Can you tell me more about what that means to you?
> *Micah*: Yes, rules are things that you have to follow perfectly; guidelines are things to help you on that path, but aren't necessary. The contract has guidelines because it's asking you which one you want to do, while a grading rubric is telling you what you have to do.
> I*nterviewer:* So even the specific items on the contract felt like guidelines?
> *Micah*: Yes, because even if you couldn't do that, you could do something a little lower or a little bit easier.

Some students (23%, n=9) explicitly noted how the contract provided the freedom to explore their own "*creative liberties*" (James, FT honors 9th) as they were "*a little more free. . . to write how [they] wanted*" (Olivia, PE regular 10th).

> *James* (FT honors 9th): A grading contract, I find, like I said, is a little more liberating, a little more open. After I had a better understanding of how it worked, the contract kind of allowed me to write how I write, and it allowed me to be more confident in what I was writing because I knew that taking those liberties wouldn't necessarily be a risk of hurting my grade, so I felt like I was more confident in the grade that I would get because I knew trying to take those risks, if you will, would be detrimental to my grades. Most of the time in previous writing that I've done, every paragraph was very structured according to the rubric, versus with the contract with the expectations of how to write a good paper, it was kind of like, "You know what should be there. Now, write a good paper." Instead of trying to check off all the boxes, I was trying to write a coherent and good paper that would be enjoyable to read.

> *William* (PE AP 12th, ELL): The old rubric had these sections where this is 10 points, this is 20 points, and this is five points, while the new contract only had guidelines. No points. The contract made us have more freedom in writing these research papers, because before we were too busy calculating our grades and we were too busy looking at the rubric: "Did I do this? Did I do this?" But I think for this research paper, I was, like, "I'm just going to write what I want to say on the research paper." We had more freedom, and I thought it fulfilled the purpose of a research paper more, making us write more freely about our thoughts, and our thoughts were more expressive in our research paper.

Both James and William pointed to the grading contract's ability to shift students to learning orientation, allow them to enjoy the learning process, and focus on writing. As Kohn (2006) argued, grades can create students who "think less deeply, avoid taking risks, and lose interest in the learning itself" (p. 12).

On the other hand, a few students (13%, n=5) expressed discomfort with the new approach. Notably, all students who expressed discomfort contracted for a B. While Jax (PE accommodated 12th) said the contract's simplicity "*threw [her] off a little bit*," he felt his expectations were "*easy to fulfill*," which was a "*stress reliever*." A few (8%, n=3) experienced irritation "at first" that appeared to persist through the assessment:

> *Olivia* (PE regular 10th): It just kind of felt pointless at first. I'm like, "Why am I signing this?"
>
> *Tristan* (PE regular 11th): All of a sudden, after so many years of doing it one way, I think it may have been a touch almost irritating, like, "Oh, they just changed it on me."
>
> *Dylan* (PE regular 11th): Originally, I looked at [my teacher] and was like, "Oh, this might be a problem," because the grading rubric normally for someone like me is like a checklist that I get to go down and be like, "OK. Have I done this? Have I done this? Have I done this?" So, I originally thought, "Oh, no grading rubric?" We have a vague contract that made me nervous at first.

These students shared negative appraisals of the contract, as well as negative introductions from their teachers, the focus of the next section. First, while the grading contract described the quality of grammar and syntax needed to earn an A or B, both Dylan and Tristan missed the listing of each grammatical element (e.g., commas, run-on sentences, etc.). Dylan felt he had to "*compensate for the lack of rubric*" by asking more people to read his paper. Change, after all, can be stressful for both students and teachers.

This finding echoes Spidell and Thelin's (2006) finding with college students, who overwhelmingly resisted contract grading due to what the authors call undemocratic educational conditioning.

Tristan also mentioned that his teacher "*attached the old rubric*" to "*make things easier*," telling the class, "*This is a good place if you're one of those people that likes to have the old rubric to start*.'" Like Dylan, Tristan, who described himself as "*one of those people that doesn't like not having some form of a right answer*," called the contract "*vague*" with "*more, I guess, leeway for the teacher to bump grades a little bit based on performance in the class or viewed effort*." To Tristan, the contract was more subjective while the rubric followed a precise formula: "*You get twenty-five points for having good grammar. Every misplaced comma or misspelled word is a half-point off or whatever*." Not only is such a system of writing assessment punitive and thus more potentially threatening for students but it is time-consuming for teachers. Ultimately, while Tristan was the only student to mention receiving both the contract and the old rubric, other teachers undermined the contract while introducing it to students in different ways, which will be explored in the next section.

4.3 RQ.2 How does contract grading affect students' self-worth protection behaviors?

4.3.1 Increased Self-Perception and Decreased Fear of Failure

Some students perceived the option to choose between the A or B contract as trivial, particularly when it was undermined during their teachers' introduction of the contract;

however, to the majority of adolescents I interviewed, the choice was psychologically motivating, increasing confidence and decreasing fear of failure. The quantitative analysis revealed that students with PE report significantly higher self-perceptions under contract grading (see Table 4.6). They were significantly less likely to "fear failing the research paper" (from M=4.85, SD=1.952 to M=4.4 to SD=1.931), t(279)=3.531, p=.000, and significantly less likely to report that it was "difficult to do my best on the research paper due to a persistent fear of failure" (from M=3.68, SD=1.757 to M=3.37, SD=1.7), t(283)=2.857, p=0.028. They were also significantly more likely to report that they were "confident [they] would be successful" (from 4.54, SD=1.578 to M=4.76, SD=1.456), t(284)=-3.403, p=.034, and "confident [they] will write successful research papers in the future" (from M=4.77, SD=1.341 to M=5.01, SD=1.34), t(284)=-3.049, p=.005.

Table 4.6

Self-Perceptions: Matched-Pairs T-Test Results for Prior-Experiencers

Statement	Conventional Grading		Contract Grading			*t*(283)	FDR-adjusted *p*-values	Cohen's *d*
	M	*SD*	*M*	*SD*				
I feared failing the research paper.	4.85	1.952	4.4	1.931	279	3.531	.000	0.232
I am confident I will write successful research papers in the future.	4.77	1.341	5.02	1.34	284	-3.049	.005	0.186
I found it difficult to do my best on the research paper due to a persistent fear of failure.	3.68	1.757	3.37	1.7	283	2.857	0.028	0.179
I was confident that I would be successful during the research paper	4.54	1.578	4.76	1.456	284	-3.403	0.034	0.145

unit.

Note. This table reveals the significant t-test results for self-perceptions of Prior-Experiencers.

The majority (80%, n=16) received messages from their teachers that inspired confidence, personal control, and responsibility, as illustrated through Vanessa (PE regular 12th):

> [My teacher] gave us a B contract and then the A contract, and she said, "You can choose which one you want to strive for." She gave us options of what was best for us because we know ourselves the best, and so basically, we just kind of chose it. "There is a list of 10 things, so maybe push yourself to do that or like you could do the B," but she put the choice in our hands; it was our responsibility to do what we wanted for the research paper.

Some students explained that actions like signing the contract, along with their parents and teacher, emphasized personal control and responsibility:

> *Lucas* (FT honors 9th): You're both signing that you both agreed that you'd get this certain grade that you signed up for, so I guess it does kind of put more responsibility on you knowing that you actually sign something.

Some students (20%, n=8), however, revealed that the contract was undermined by their teacher who emphasized their ultimate authority and made the students' choice of A or B appear trivial:

> *Sydney* (PE regular 10th): I think [my teacher] basically said something along the lines of, "It's not going to affect your grade either one you pick. You'll still get the grade you're gonna get."
>
> *Kevin* (PE accommodated 10th): I remember her saying that if you circle B, that doesn't mean that you're gonna get a B—you could still get a C or an A.
>
> *Tristan* (PE regular 11th): I was told that even if you contracted out for B—and I did some poking around and found that it was similar across most classes—it wasn't so much a guaranteed B as it was if you do one of the items on it to a substandard degree, as determined by the teacher, it could drop you down to a C or if you did really well, you could end up getting in an A. That especially added to the idea that this is just rubric under a different name.

While a student who did not complete a contract item could earn a C, doing "*one of the items on it to a substandard degree*" would not drop a student off the B contract. (The assessment process will be explained in detail in Chapter Five.) A high percentage of students received messages that their choice was trivial, that "*it's not going to affect your grade either one you pick*," which undermined the contract's participatory element and may have undermined their task-oriented motivation.

Despite a high percentage of students (20%) who had the contract undermined by their teacher, the qualitative analysis revealed that the majority of adolescents were motivated to achieve success (90%, n=36), rather than avoid failure, by setting high goals and monitoring their progress by checking the contract, which increased their confidence while making students less likely to attach their self-worth to doing well on the paper, as revealed by the significant decrease in importance of ability under the contract (M=5.08, SD=1.01) compared to conventional grading (M=5.37, SD=1.06), 4.36(276), p. 000 (see Table 4.7).

Table 4.7

Importance of Ability Subscale Prior-Experiencers

	Conventional Grading		Contract Grading		*n*	*t*	*p*	Cohen's *d*
	M	*SD*	*M*	*SD*				
Importance of Ability	5.37	1.06	5.08	1.01	276	4.36	.000	.280

Note. The importance of ability subscale had an acceptable internal consistency: α=.619 for pre -survey and α=.576 for the post-survey.

Notably, the majority of students (75%) contracted for an A, setting a high goal. Despite inconsistent messaging from teachers, the contract oriented students to achieve success this year:

> *Daniela* (PE honors 10th): I feel like that's where the difference between a contract and a rubric comes into view. Like with a rubric, it's, like, "Okay, if I just follow this, I'll see what I get." But with a contract, it's like you have to follow these points and then it's like, "I can get that," so I guess it's like a mentality sort of thing: I can versus I'll see if I can.
>
> *Mason* (PE accommodated 12th): It impacted my personal confidence by allowing me to strive toward getting an A because I signed on the contract to get an A on it, so I knew what kind of goal I was setting for myself.

This finding corroborates Smith and Lerch's (1972) claim that the contract can alleviate the threat and stigma of failure by promising a good grade upon completion of each item. Additionally, goal-setting theory states that when individuals set specific, measurable goals, task-oriented motivation is more likely to increase to achieve the goal (Friedman & Mandel, 2009). Additionally, with the reduction of evaluative threat, a leading cause of self-worth protection behaviors, students were significantly less likely to fear failure or engage in maladaptive coping behaviors. The next section examines the significant decrease in avoidance orientation under the contract grading system.

4.3.2 Decreased Avoidance Orientation

Prior-Experiencers were significantly less likely to engage in avoidance behaviors and "play it safe" under the contract, although the effect size was small (see Table 4.8). The

quantitative analysis revealed that adolescents with PE were significantly less likely to avoid challenges that could end in failure under the contract (M=3.82, SD=1.468) compared to their prior experience (M=4.22, 1.53) t(283)=3.598, p=.000. They were also significantly less likely to play it safe by "choos[ing] goals that were within my reach" under the contract (M=4.41, SD=1.412) compared to their prior experience (M=4.74, SD=1.262) t(284)=3.495, p=.004. Adolescents were less likely to choose "safe goals," as 75% of students with PE contracted for an A while 25% contracted for a B, and thus they were significantly more likely to report that it was "easy to put my best effort into the research paper because there was little risk of failure" (from M=3.96, SD=1.493 to M=4.22, SD=1.433), t(284)=-2.47, p=.014; and that they were "able to try my hardest because failing the research paper would not reveal low ability" under the contract (M=4.22, SD=1.433; M=4.08, SD=1.33) compared to their prior experience (M=3.96, SD=1.493; M=3.82, SD=1.432), t(284)=-2.452, p=.000.

Additionally, they were significantly more likely to report that "it was easy to try hard because my performance wasn't being judged" (from M=3.54, SD=1.523 to M=3.9, SD=1.415), t(284)=-3.852, p=.000; and significantly less likely to report that "it was hard to try my best" due to feeling of judgment (from M=3.63, SD=1.721 to M-3.2, SD=1.553), t(284)=3.852, p=.000. Chapter Five will examine that impact on students' academic performance, as measured by the final grade on the paper, but these findings reveal positive psycho-emotional shifts under the contract grading system.

Table 4.8

Avoidance Orientation Matched-Pairs T-Test Results for Prior-Experiencers

Statement	Conventional Grading		Contract Grading		*n*	*t*	FDR-adjusted *p*-values	Cohen's *d*
	M	*SD*	*M*	*SD*				
It was easy to try hard because my performance wasn't being judged.	3.54	1.523	3.9	1.415	284	-3.852	.000	0.245
I avoided challenges with the research paper that could end in failure.	4.22	1.53	3.82	1.468	283	3.598	.000	0.266
I felt like my performance was being judged, so it was hard to try my best.	3.63	1.721	3.2	1.553	284	3.852	.000	0.262
I tended to play it safe with my research paper and choose goals that were within my reach.	4.74	1.262	4.41	1.412	284	3.495	0.004	0.237
I was able to try my hardest because failing the research paper would not reveal low ability.	3.82	1.432	4.08	1.33	284	-2.453	0.024	0.188
I found it easy to put my best effort into the research paper because there was little risk of failure.	3.96	1.493	4.22	1.433	284	-2.47	0.014	0.178
I underachieved relative to my level of ability, choosing easy goals in	3.82	1.495	3.35	1.458	279	2.277	0.033	0.183

order to ensure success
on the research paper.

Note. This table reveals the significant t-test results for avoidance orientation among Prior-Experiencers.

One student I interviewed, however, shared that the challenges of the previous year continued to weigh on him:

> *Interviewer*: How worried or stressed were you about the research paper this year?
> *Dylan* (PE regular 11th): I wasn't extremely stressed. I felt OK, and I think it's simply not because there's a new contract or anything, but just because I knew it was another research paper. And so I sort of knew what to expect. But at the same time, I don't know if I would be able to change it, so why should I worry about it? Like, if I feel I am powerless, then why worry? Why make that pain for myself?
> *Interviewer*: Did you feel powerless over the research paper?
> *Dylan*: I mean, I always kind of have. Freshman year especially I felt powerless. Sophomore year, I got a little better. This year, a little better, but I expect not to get an A. I expect to get a high C or low B because I just don't know.

He described confidently turning in the paper freshman year and then the painful experience of earning a C, two grades lower than he expected: "*I signed for a B because I don't have a crazy confidence, so I feel like I did good. I mean, freshman year, I felt like I did great until I saw my grade, then sophomore year, same thing happened*." As a result, he entered his sophomore year's assessment with less confidence and a sense of powerlessness that he has carried with him into 11th grade. Additionally, he had the same teacher for both 9th and 11th grades. The decision to contract for a B was not a healthy decision; a result of his low self-efficacy and -belief as a result of negative academic experiences (Bouffard-Bouchard, 1990), it was a self-protective mechanism to avoid future hurt. Later, he said:

> This year, I didn't feel like I was going to fail. I felt there was no way I was going to get an F or D or something terrible. What I did feel is that my grade was most likely locked out like a C+, so it wasn't a fear of failure, it was just kind of like,

> *Eh, what's the point?* Well, not really what's the point, but *can I improve it?* It was kind of like I thought I was working for a C the whole time despite me signing for a B.

For students like Dylan, who contracted for a B, their mean grade the previous year was 78%, perhaps making the contract's promise feel out of reach or too good to be true. As Dylan explained, the B contract did not feel achievable based on his negative prior experience, which lowered his perception of self-efficacy (Bouffard-Bouchard, 1990). While it did not lead him away from challenges, as Shim and Ryan (2005) reveal can happen, Dylan's attitude and effort were impacted by his low self-efficacy.

Dylan's experience, however, was an anomaly among the interviewees; at all levels, the majority of adolescents (57%, n=23) cited the contract as increasing their confidence:

> *Alexander* (FT regular 12th): It increased my confidence, knowing, again, that it gives you what you need to get the grades you want. It's not like a normal rubric [qualities matrix] where you think you're writing to that standard that's in the A, B, C, or D square, and a lot of times, me and fellow students will think we're writing at A or B level. Turns out, you know, it's graded in parts. I loved how the [contract] didn't have the parts—it just gave you what you needed. In a rubric, we obviously think we're doing great and then we realize we get our grade back that one of our parts is lacking because we didn't focus on it. We don't have to focus on all those parts on this new contract. There's no nitpicking—you just go down the list.

> *Madison* (FT regular 9th): With the contract, it's not like you're guessing to yourself, like, "Did I do this right?" It's like, "Did I do everything I needed to do because there's a list right there?" And it's also kind of like you're getting your self-confidence that you're saying you're gonna get an A. ...I think their mindset changes and they convince themselves they're going to get an A.

Madison observed that her peers' behaviors were the result of their perception that they could reach their goal. For a student to convert their behaviors into performance, they

must first believe they can attain the desired result with the assistance of their actions (Cikrikci & Odaci, 2015). The adolescents reported increased confidence, which resulted in decreased evaluative threat and perceptions of stress, and increased self-efficacy, a predictor of life satisfaction in adolescents (Cikrikci & Odaci, 2015).

Furthermore, the new mindset of self-confidence and sense of assurance fostered by the grading contract led to significantly less social comparison, the focus of the next section.

4.3.3 Reduction in Social Comparison

Table 4.9

Social Comparison T-Test Results for Prior-Experiencers

Statement	Conventional Grading		Contract Grading		*n*	*t*	FDR-adjusted *p*-values	Cohen's *d*
	M	*SD*	*M*	*SD*				
I compared my ability to those around me.	5.21	1.695	4.74	1.777	284	3.976	.000	0.270

Note. This table reveals the significant t-test results for social comparison among Prior-Experiencers.

Under the grading contract, adolescents were significantly less likely to engage in social comparison (see Figure 4.9). More specifically, adolescents with PE were significantly less likely to say that they "compared [their] ability to those around [them]" under the grading contract (M=4.74, SD=1.777) compared to the previous year (M=5.21,

SD=1.695), 3.976(284), p=.000, with an effect size of .270. The interviews revealed that social comparison often arises from doubts about one's work and unclear expectations. In the absence of these fears, the contract reduced social comparison behavior:

> *Sara* (PE honors 11th): I feel like in previous years I've been more eager to ask people like, "Oh, where are you? How many words do you have?" And this year, I was more just focused on myself and getting my own work done.
> *Interviewer*: Why do you think that was?
> *Sara*: I don't know. I just kind of felt like if I did those 10 steps on the contract, I would be fine, and I don't worry about what anyone else is doing.

Students relied on social comparison as a means of self-evaluation. Previously, social comparison served as a way to check their progress against their peers and ensure their work was complete and correct. Social comparison theory, put forth by Festinger (1954) and cited in Taylor, Buunk, and Aspinwall (1990), states that individuals prefer objective, non-social criteria to evaluate themselves, yet in its absence, they will use "social information—namely, other people" (p. 75). Strong emotion, such as fear, can also lead to social comparison: "when people find themselves in challenging, threatening, or unfamiliar circumstances, the need to evaluate the situation, their resources, and their emotional reactions is paramount" (Festinger, 1954, p. 76). With a reduction in their perception of stress and fear of failure, social comparison significantly decreased as they did not need to "*to worry about what anyone else is doing*" (Sara, PE honors 11th).

4 Summary of findings

Quantitative Findings
The adolescents in this study reported— ● Significantly less stress from workload demands and time constraints;

● Significantly less threat appraisal and more challenge appraisal compared to their prior experience or expectations for the unit; ● Significantly higher self-perceptions, including increased confidence and reduced fear of failure; and ● Significantly less avoidance orientation and social comparison

Qualitative Findings
The clarity of the grading contract alleviated many worries, doubts, and fears about assignment expectations that are often associated with conventional grading practices and thus significantly reduced academic stress.
Ultimately, the contract's clarity about task requirements and expectations ameliorated stress for students of all course types.
The contract helped students positively reappraise workload demands and time constraints, even as they completed more work than before.
The contract's clear presentation of workload demands altered students' perception of them and significantly decreased distress.

The findings of this chapter reveal how contract grading altered secondary students' perceptions of workload demands, time constraints, and themselves in contrast to conventional grading practices. The key finding is that compared to conventional grading, contract grading provided an unusual clarity of purpose that reduced perceptions of stress and evaluative threat. Self-worth protection behaviors decreased as students were less likely to report finding their worth in their ability; instead, the majority of students in this study set high goals and used the contract to work efficiently. As a result of these findings, this chapter provides compelling empirical support for the use of contract grading in secondary classrooms to improve the psycho-emotional well-being of high school students.

The next chapter examines students' performance, as measured by their final grade, under improved psycho-emotional conditions.

Chapter Five: Academic Performance

5.1 Introduction

This chapter presents the findings of adolescents' academic performance under the mastery-based contract grading system (Research Question Three), in which all students (grades 9-10) were asked to 'contract' for either an A or B and meet the standards of proficiency. Final grades on the high-stakes assessment were analyzed alongside demographic variables, including grade level, course type, and gender, to ascertain the factors that impacted achievement (Research Question Four). In addition to grade averages for each variable, academic performance was measured by the rate of 'contract fulfillment'—that is, the number of students who fulfilled the requirements for either A or B to meet a high standard of achievement. Two grades for each 10-12th grader with prior experience with the assessment (called Prior-Experiencers) were analyzed in matched-pairs t-tests: their final grade under the conventional grading system (2019) and their final grade under the contract (2020). Since academic growth is expected each year, the prompt and expectations increase in difficulty, making it likely that academic performance will remain relatively stable each year. Academic growth may account for some variance in students' performance, but it is unlikely to be statistically significant given increased task demands.

As mentioned in Chapters One and Two, the impetus for this study was a steady rate of low and non-passing grades on the assessment: in 2018, 26% of freshman and 20% of regular-level seniors at the high school earned a D or an F. While the workload for the

project was decreased in 2019, the findings of the pilot study (Ward, 2021) suggested that a decrease in workload alone did not significantly impact students' academic performance or perceptions of stress. This further supported the argument in this thesis that the number of non-passing grades was related to psycho-emotional issues. Academic performance, then, was this study's secondary focus after psycho-emotional well-being. The findings presented in Chapter Four revealed that the contract grading system significantly altered adolescents' perceptions of the task's demands and their perceived agency over the task, which reduced evaluative threat and oriented students toward success. Building on Chapter Four, this chapter examines academic achievement under significantly improved psycho-emotional conditions of decreased perceptions of stress and fear of failure and increased confidence.

While researchers have examined college students' grades under the contract grading (Poppen & Thompson, 1971; Fairbanks, 1992; Lindemann & Harbke, 2011), this chapter is the first to examine high school students' academic performance under contract grading. In addition to the quantitative data to understand adolescents' academic achievement under the contract, this chapter also includes adolescents' voices, joining one study with college students that included qualitative data (Spidell & Thelin, 2006). While not all interviewees knew their final grade at the time of the interview[1], all were

[1] Students' final grades on the assessment were posted later than previous years due to COVID-19 and the demanding transition to remote learning. Teachers, some of whom have more than 160 students, are normally allotted paid time off to assess papers. School closures, however, changed the working environment as well as the ability to take time off. Many teachers, most of whom are women, had to balance their own workload demands with small children at home.

invited to a follow-up to ensure the record reflected their final experience with the grading contract.

The next section provides clarity and transparency on the process teachers used to assess the papers. The subsequent sections outline academic performance, beginning with overall academic performance across all participants and then examining significant findings based on demographics variables to answer Research Question Four.

5.2 Assessing with the Grading Contract

All students 'contracted' for either an A or B during the first week of the assessment. Some 9th-grade teachers asked students to discuss the contract with their parents, and one teacher asked them to sign it with their parents before acquiring their teacher's signature, as Anna (FT regular 9th) explained: "*Later, we turned in our contracts with [our teacher's] signature and our parents' signature, so it was just like, 'I agree today that I'm willing to do everything this says*.'" Signing the contract was a visible commitment to her goal and her accountability partners. For Anna, this action fostered positive support: "*I think that my mom was really excited about [the contract], like, 'OK. It's a list. This is something that we can do together*.'" While the process could cause pressure from both parents and peers, the majority of First-Timers did not mention the signing of the contract or discussing it with their parents as significant to their experience. Given that achievement pressure is prevalent among wealthy families (Luthar & Latendresse, 2005), parents asked to sign the contract could lead their children to contract for higher goals;

however, combined with the findings of Chapter Four, it appears the teachers' decision to include parents in the decision-making process did not negatively impact well-being.

The majority of students reported signing contract themselves, which was an empowering, real-world experience for some:

> *Alexander* (FT regular 12th): The grading contract—I loved it. It was great because it's like real life, you know; you sign something and it *is* what you sign. It told you what you needed to do and if you didn't do that, then don't expect it.
> *Interviewer*: What do you mean by 'real life'?
> *Alexander*: I mean that in life, you're gonna sign a contract no matter where you are, whether it be a job or, you know, something along those lines, and it gives you expectations. And you just follow that through, and it gives you a chance.

In this way, the signing of the contract appeared to promote a sense of control over the task. Fishman (2014) found that "if students feel as if their academic outcomes are within their control, they are more likely to feel responsible for the completion of those outcomes" (p. 697).

While students participated in the assessment process, each teacher held the final decision over the grade, a practice common among contract grading practitioners (Danielewicz & Elbow, 2009; Potts, 2010; Litterio, 2016, 2018). Before the final papers were collected, all instructors in the study (n=13) gathered in course-alike teams to practice grading holistically. They were taught by the lead teacher of the pilot study to use a check-system for each contract item (see Table 5.1). The average of the marks provided a holistic assessment of the paper as 'high,' 'medium,' 'low,' or 'off-contract,' which were printed

on the top of each student's contract. Each level was associated with one number, reducing the total number of grades from 20 (i.e., 80-100) to six (see Table 5.2).

Table 5.1
Check-Mark System to Assess Papers in the Contract System

Mark	Description
✓+	Exceeds expectations for contract item
✓	Meet expectations for contract item
✓-	Below expectations for contract item
X	Missing

Note. Teachers were encouraged to use a simple check-system for each of the ten items to assess each paper holistically with the grading contract.

Table 5.2
Percentage Grades from Holistic Assessment

Assessment	A Contract	B Contract
High	98+	88
Medium	95	85
Low	92	82

Note. Each contract was associated with three different numerical grades, each corresponding with the teacher's final assessment of the paper as 'high,' 'medium,' or 'low.' Off-the-contract papers earned scores lower than 82.

Each teacher also had a not-for-student-use list of qualities of C and D papers. In the mastery-based approach, all students were asked to strive for proficiency; however, 'off the contract' papers that met an adequate standard of achievement could earn the passing grade of C (i.e., 72, 75, 78). Incomplete papers warranting a non-passing grade could be revised and resubmitted for passing grade. Since D and F are non-passing grades that

require course remediation, the revise-and-resubmit policy encouraged all students to submit a passing paper, yet revised papers received a penalty: at most, they could earn a C. Consequently, revised papers were still 'off the contract' and not included in the contract fulfillment rates in the coming sections.

5.3 Quantitative Findings on Academic Performance

Of all participants (n=439), 90% (n=390) fulfilled the grading contract, earning either A or B on the final assessment, and 97% (n=421) earned a passing grade (i.e., A, B, and C), including 97.5% of Prior-Experiencers (n=269) and 97% of First-Timers (n=152) (see Table 5.3). This reveals that prior experience with the assessment or high-school level writing did not impact the rate of low or failing grades on the assessment. Both groups had a small but notable number (2.5-3%) of students, with similar characteristics that will be explored later, who earned a non-passing grade of D or F.

Table 5.3
Descriptive Statistics of Academic Performance of Students with Prior Experience

Variable	Total *N*	*N* Fulfilled Contract	% Fulfilled Contract	Earned C	Not Passing (>69%)	Not Passing %
All students	278	260	94%	15	7	2.5%
Accommodated English	23	19	83%	2	2	9%
LD in Regular English	10	9	90%	1	0	0%
LD in Honors English	2	2	100%	0	0	0%
Regular English	147	132	90%	12	3	2%

Honors English	112	109	97%	2	1	<1%
10th Grade	56	50	89%	5	1	2%
11th Grade	121	114	94%	6	1	<1%
12th Grade	100	95	95%	4	1	1%
Males	124	112	90%	9	4	3%
Females	154	146	95%	6	2	>1%
In-Season Athletes	126	114	90%	10	2	1%
Non-Athletes	156	146	94%	5	5	3%
Ill During Unit	96	87	91%	6	3	3%
Healthy During Unit	186	173	93%	9	4	2%

Note. This table reveals the rate of contract fulfillment for each demographic variable among students with prior experience with the assessment.

Of the First-Timers, 84% (n=130) fulfilled the grading contract, including 84% (n=111) of 9th graders and 88% (n=14) of students who transferred from another high school for 10th, 11th, or 12th grade (see Table 5.4). Of Prior-Experiencers, 94% (n=260) fulfilled the contract, while only 75% (n=208) of the same students earned a B or higher under the conventional grading system (see Table 5.3). As previously mentioned, since the difficulty and demands of the task increase with each grade level, it is expected that academic performance will remain relatively stable, yet Prior-Experiencers had a 25% increase in the numbers of A and B earned under the contract.

Table 5.4

Descriptive Statistics for First-Year Students' Academic Performance

Variable	Total	Fulfilled Contract *N*	% Fulfilled Contract	Earned C: Passing but Below Contract	Not Passing (>69%)	Not Passing %
All FT Students	154	130	84%	20	4	2.6%
9th Grade	140	117	84%	19	4	3%
Transfers (grades 10-12)	16	14	88%	1	1	6%
Accommodated English	12	9	75%	3	0	0%
LD in Regular English	4	2	50%	1	1	25%
Regular English	94	82	87%	7	5	5%
Honors English	48	39	81%	9	0	0%
Males	84	67	80%	13	4	5%
Females	71	63	89%	7	1	1%
In-Season Athletes	92	78	85%	12	2	2%
Non-Athletes	63	52	83%	8	3	5%
Ill During Unit	46	38	82%	7	1	2%
Healthy During Unit	105	92	88%	9	4	4%

Note. This table reveals the rate of contract fulfillment for each demographic variable among students completing the assessment for the first time.

The findings of the matched pairs t-test also revealed that Prior-Experiencers earned significantly higher grades under the contract (M=88.26, SD= 9.90) compared to the previous year (M=85.49, SD= 9.34), t(283)-4.718, p=.000 (see Table 5.5). Statistically

significant grade improvements were observed in the majority of demographic variables analyzed (see Table 5.5), including non-athletes (M=85.21, SD=9.85 to 88.67, SD=10.91, t(156)=-4.401, p=.000); in-season athletes (M=85.83, SD=8.69 to 87.75, SD=8.5, t(126)=-2.179, p=.04); and students who fell ill during the unit (M=85.31, SD=9.63 to M=87.86, SD=9.66, t(96)=-2.359, p=.028). Notably, the changes were statistically significant yet also marginal in terms of percentage points, which increased, on average, 2-3%, moving students from B to B+. Two grade levels had more significant gains: 10th graders (from M=84.21, SD=12.09 with the rubric to M=89.89, SD=8.97 under the contract, t(56)=-4.166, p=.000) and 12th graders (from M=84.66, SD=7.217 to 89.35, SD=8.67, t(101)=-5.866, p=.000). This amounted to a 5.69% increase under the contract for 10th graders and 4.69% for 12th graders. Also, 10th and 12th graders earned similar grades on the project, revealing that academic performance remains relatively stable each year.

Table 5.5

Results of Academic Achievement T-Test by Demographic Variable (Grades 10-12)

	Grading Rubric		Grading Contract		*n*	*t*	Sig.	Cohen's *d*
	M	*SD*	*M*	*SD*				
1 All Students	85.49	9.34	88.26	9.90	283	-4.718	.000	0.288
2 Accommodated Students	88.17	7.5	84.30	14.11	23	1.610	.144	0.342
3 Regular Students	81.16	10.01	86.96	10.18	147	-6.934	.000	0.574
4 Honors Students	90.62	4.98	90.78	7.77	112	-.220	.895	0.024

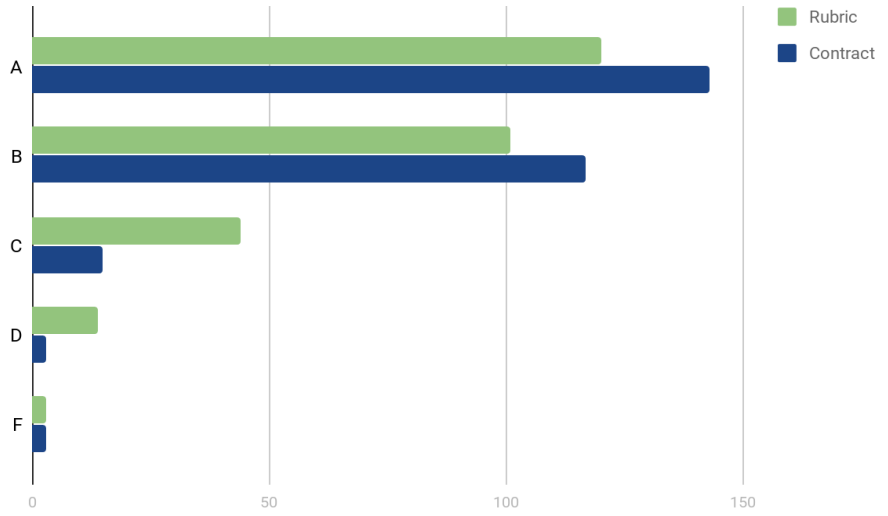
5 Sophomores	84.21	12.09	89.89	8.97	56	-4.166	.000	0.534
6 Juniors	86.66	9.3	86.58	11.36	126	.078	.938	.008
7 Seniors	84.66	7.217	89.35	8.67	101	-5.866	.000	.588
8 In-Season Athletes	85.83	8.69	87.75	8.5	126	-2.179	.040	.223
9 Non-Athletes	85.21	9.85	88.67	10.91	156	-4.401	.000	.332
10 Ill During Unit	85.31	9.63	87.86	9.66	96	-2.359	.028	.264
11 Healthy During Unit	85.59	9.2	88.47	10.05	186	-4.145	.000	.300
12 Males	83.55	9.68	86.74	9.18	125	-3.569	.001	.338
13 Females	87.32	8.36	90.04	7.4	154	-3.858	.000	.345

Note. This table reveals statistically significant and non-significant findings of the matched-pairs t-test, which analyzed changes in returning students' academic performance from 2019 to 2020.

Corroborating the findings of the pilot study (Ward, 2021), this analysis reveals a positive trend of modest grade increases under the contract across multiple grade levels and demographic variables. The contract did not result in dramatic grade increases, yet the analysis reveals a pattern of statistically significant increases (see Table 5.5). The number of students (90%) who fulfilled the grading contract to meet a high standard of academic achievement is most significant. Under the conventional grading system, 19% (n=53) earned a C or lower, compared to 7% (n=20) under the contract. Ultimately, this is a 62%

decrease in grades C or lower and an 11% increase in grade B or higher. As Danielewicz & Elbow (2009) suggested, the goal of "badger[ing] and cajol[ing] every student into a getting a B—that is, into doing everything we specified in the contract" was overwhelmingly successful with the adolescents in this study (p. 254).

Figure 5.1
Academic Performance Outcomes Between Conventional Grading and Contract Grading



Note. The bar chart shows the improved academic outcomes between the conventional grading rubric and the grading contract.

Ultimately, these quantitative findings reveal that the contract grading system, as utilized by 12 different teachers in 13 courses, helped high schoolers of all grade levels, course types, and demographic variables reach high academic standards, as measured by the rate of contract fulfillment and grade improvement. This finding corroborates Johnston and O'Neill's (1973), who reported that students will rise to meet clearly outlined task

demands, regardless of prior performance. In this study, outlining expectations for A and B facilitated high achievement.

The findings of the quantitative analysis also corroborate research with college students, who earned higher grades under the contract compared to their peers in conventional grading courses (Fairbanks, 1992; Lindemann & Harbke, 2011). This research suggests that grading contracts combined with prior experience with high school writing expectations led to higher performance outcomes, as First-Timers who were upperclassmen (88%) had a higher rate of contract fulfillment than 9th graders (84%) with varying prior experiences with academic writing before high school. Prior-Experiencers had the highest rate of contract fulfillment (94%), pointing to the benefit of spaced practice and teacher feedback on durable learning that improves educational outcomes (Kang, 2016). Familiarity with terms and expectations as a result of task repetition, even with increased task difficulty, led to high academic achievement.

The next section examines the qualitative data, which revealed a chain of events that resulted in strong academic performance: clear path to a desirable outcome—> increased control and task manageability —> decreased fear of failure and increased confidence —> increased academic performance.

5.4 Qualitative Findings on Academic Performance

The qualitative analysis revealed that the contract's clear and accessible presentation of the learning and task goals (e.g., "do X, Y, and Z to earn an A") enabled students to

understand and then meet the expectations. To mitigate a formulaic paint-by-numbers approach, the students perceived the contract as offering more flexibility and choice, as William (PE AP 12th, ELL) observed that contract didn't '*trap*' or '*stress*' students; instead, it felt more '*liberating to the students*.' James (FT honors 9th) also called the contract grading "*a little more liberating, a little more open*. The contract allowed for personalization by guiding him to "*what exactly [he] should be doing*" while concurrently having the assurance that he can "*get a good grade and still write the way [he] feels fits the topic*." Additionally, Micah (FT accommodated 9th) revealed he perceived the contract allowed for more personal control and customization than a grading rubric by offering 'guidelines' rather than 'rules':

> *Interviewer*: What parts of the research paper felt within your control?
> *Micah*: Probably everything except for the conclusion. The conclusion was basically a generic conclusion that we wrote in class. It was because I thought of the instructions not as rules but as guidelines. But for the conclusion, there were definitely rules for that one I couldn't just abandon.
> *Interviewer*: You said that the instructions felt more like guidelines than rules. Can you tell me more about what that means to you?
> *Micah*: Yes, the rules are things that you have to follow to be perfect. Guidelines are things to help you on that path, but aren't necessary.
> *Interviewer*: Would you say that the contract has rules or guidelines, and why?
> *Micah*: The contract has guidelines because it's asking you which one you want to do. You, which one do you want to do, while a grading rubric is telling you what you have to do.
> *Interviewer*: Even the specific points on the contract felt like guidelines, items one through 10?
> *Micah*: Yes, 'cause even if you couldn't do that, you could do something a little lower or a little bit easier.

Micah perceived the contract as different from the rubric because it '*ask[ed] you which one you want to do*.' The task's newfound clarity of purpose increased students' sense of control over task demands (e.g., "Step by step, I can do X, Y, and Z"); in turn, with the

doubts and worries about the unknown mostly eliminated, participants reported significantly less fear of failure and significantly more confidence, as detailed in Chapter Four. Consequently, many students contracted for an A, setting higher goals than they had previously earned (e.g., "I didn't know how to earn A before, but now I do").

5.4.1 Clarity of Purpose

The qualitative data revealed the conditions that made the project significantly less stressful and threatening, as detailed in Chapter Four, facilitated high academic achievement. First, among the students I interviewed, the clarity of the grading contract promoted a sense of control over the project's demands that reduced evaluative threat because success was within reach, as Malcolm (FT honors 10th) said:

> I knew exactly what was expected of me to get a certain grade, so that made it very straightforward to what I needed to do.

Since effort ('*what you had to do*') produced an outcome ('*a certain grade*'), the grading contract clarified the performance-reward relationship for the adolescents in this study, as others have observed with college students (Polczynski & Shirland, 1977; Danielewicz & Elbow, 2009). The adolescents described prior writing experiences in which the path to success was opaque, as described in Chapter Four, while the grading contract provided a '*clear*,' '*straightforward*,' and '*direct path*,' a phrase used by 12 interviewees (30%), to a desirable outcome that encouraged task-oriented effort. By spending less time figuring out the expectations, or re-doing work he did incorrectly, Eric (PE regular 11th) described focusing more on the *"quality of the content and the quality of the little parts"* this year compared to previous years. Similarly, Ella (PE 11th honors) observed:

> It was just so clear about what you're supposed to do, so then you could focus your energy on using good diction and improving your argument.

Though an initial concern among teachers in the department, the contract grading system did not lower performance; in fact, it appeared to produce high-order thinking, enabling students to set high goals and improve the quality of their argument and writing. The contract brought each element "to the forefront of student awareness by naming it and insisting on it," as Danielewicz and Elbow (2009) claimed (p. 252), which then freed students to *'focus [their] energy*' on quality and other learning goals, rather than concerns of what they were '*supposed to do*,' as Ella said. More students spent time and energy on the subjective, 'fuzzy qualities' of exceptional writing, which may have also improved academic performance.

By making learning goals clear to students, the contract grading system helped students focus their time and effort on meaningful tasks that promote learning and improve the quality of their writing. Several honors students I spoke to described regularly deducing requirements with success, while regular students shared prior experiences in which they poorly inferred the expectations, completed the work incorrectly, and then performed poorly. Even if they had discovered their error before the due date, it was either too late or in the eleventh hour after significant investment in the task. Insofar as adolescents understood its terms, and the overwhelming majority did, the contract positively impacted academic performance.

The grading contract also strengthened the task-oriented effort and confidence, yet the adolescents in this study were fixated on outcomes. For example, Mason explained (PE regular 12th, SLD):

> [The contract] impacted my personal confidence by allowing me to strive towards getting an A because I signed on the contract to get an A on it, so I knew what kind of goal I was setting for myself. . . This year, my motivation really picked up because the set guidelines made it much easier to understand and made pursuing that goal much more achievable.

He earned a perfect score on the assessment, saying, "*I never expected to get a hundred percent*. *It's incredible for me*."

The contract, which was presented at the start of the assessment, facilitated goal-setting that increased task-oriented effort in the direction of a desirable goal, which the adolescents perceived as attainable. This reveals that increased self-confidence triggered by the grading contract, a statistically significant finding in Chapter Four, preceded strong academic performance. At the time students completed the psycho-emotional scales, they did not know their final grade yet reported significantly more self-confidence while working under the contract grading system, as also reported in Chapter Four. This finding corroborates the work of Polczynski and Shirland (1977) who found that the grading contract promoted self-belief that led to more task-oriented motivation to achieve the desired results. Only when students experience motivation "that produces optimum intellectual development" can they develop their full potential (Nicholls, 1979, p. 10).

The contract's descriptive nature also led students in this study to approach the task as if they would be rewarded for their effort, as Caleb (PE honors 11th) observed:

> My perception of the grading contract was that it made the expectations more clear for the assignment and also that it would perhaps reward students for what they did correctly instead of being marked down for anything that they forgot or did poorly on.

The contract oriented students to achieve success rather than avoid failure, as revealed in Chapter Four. Like Caleb, the English Department recognized that the analytical rubric, which evaluated each criterion separately, was punitive and led some students to task-avoidance and maladaptive coping behaviors. On the other hand, the contract's descriptive and task-specific language, which detailed the performance level required to earn either an A or B, charitably showed students the reward of their effort. In her analysis of rubrics used by writing teachers, Brookhart (2018) found that teachers with task-specific rubrics did not share them with students because they would 'give away' answers, thus making them useful for grading but not for formative assessment. The task-specific contract, however, increased students' sense of certainty that their effort would result in a worthwhile reward, as Mason (PE regular 12th, SLD) observed:

> The rubric was kind of vague. You didn't know when or what level your grade would be if you didn't fall into the set of guidelines. It could possibly be a 92 or an 80-something and you wouldn't know exactly what kind of grade you would have, so you didn't really have as much confidence in yourself; whereas with the contract, you understood beforehand what kind of grade you would get.

Since '*you wouldn't know exactly what kind of grade you would have*,' the rubric increased worries and decreased self-efficacy, as students were uncertain about the outcome of their work. Uncertainty about either expectations or ability status is not conducive to learning (Covington, 1992), and as Mason recognized, teachers who use

analytic rubrics often emphasize subtle differences of degree which lead to low interrater reliability (Starch & Elliott, 1912; Breland, 1983; Rachal, 1984; Brimi, 2011). The A and B contract, with three grades for each letter (e.g., 82, 85, and 88), opposed to 20 (i.e., 80-100), offered more clarity, as Ella (PE honors 11th) explained:

> It decreased my stress and increased my confidence that I'd be able to do it because I would have been I'd say I contacted for an A, and I would have been happy with any of the A grades, so I'd say it decreased my stress.

A clear sense of purpose (i.e., what to do) increased adolescents' confidence in their ability to succeed at their goal, despite students' potentially detrimental emphasis on outcomes. As Ella observed, any of the grades of A would have been good enough, yet the contract provided more precision about what to expect, as Mason noted: "*[the rubric] could be a 92 or an 80-something."* The newfound clarity replaced fears and doubts about performance outcomes with confidence and satisfaction.

Decreased workload alone did not improve academic performance, as the pilot study revealed (Ward, 2021). The contract, in both the pilot study and this thesis study, altered the students' experience with the project as they perceived more agency over the task and felt confident in their ability to handle workload demands. Students, as quoted earlier, felt '*rewarded*' for the effort, rather than treading carefully to avoid errors and mistakes, which are central to the learning process. The fear of making mistakes can lead to task-avoidance, which lowers academic performance. As a result, academic performance improved as a result of first creating healthy psycho-emotional conditions that curbed academic procrastination, the focus of the next section.

5.4.2 Curbing Academic Procrastination

Many adolescents I interviewed shared prior experiences in which lack of clarity or purpose in expectations led to task-avoidance or passive procrastination that results from a student becoming debilitated by indecisions and then failing to complete tasks on time (Lay & Schouwenburg, 1993). Under the contract, however, Evan (PE AP 11th) explained:

> I spent less time procrastinating on it—I was able to just jump in and do what I had to do for that day and just be done, and I wasn't like wasting time being worried, like, "How do I write this?"

In previous years, task delay was often a result of ambiguity of purpose, which led students to dread the task. For example, Riley (PE regular 11th) reported increased evaluative threat during his 9th and 10th grades years from "*not exactly where to go*" with a "*huge paper*" that was "*almost intimidating,"* because—

> Once I got home, like after class—and I pay attention in English, too, and I had a good grade in English 1 and 2—but every time I would just get home and I would look at the paper and look at the rubric, it was just like, "What am I supposed to do?"

Such uncertainty about which direction to take can lead to no action. This year, however, Riley said:

> I was more motivated because sometimes in the past, I had difficulty with the research paper, so it's almost like that dreading feeling, not wanting to get into it because I knew I was struggling, but this year, I did want to get into it because it was just simple, and I wanted to get it done. . . . For me, when it's less work, I'm more likely to just do it immediately. If it's like a three-hour assignment that's due the following day, I'm a push that back to later in the night, so for me, that's horrible.

Solomon and Rothblum (1984) found that procrastination results from having too many things to do, viewing the task as unpleasant, feeling overwhelmed by the task, and feeling afraid to fail. While Rabin, Fogel, and Nutter-Upham (2011) observed that "initiation, planning, and organizational skills were predictive of academic procrastination" (p. 252), the grading contract helped improve these executive abilities.

All students completed a longer task under the contract, yet the contract positively altered many students' perceptions of workload demands, as reported in Chapter Four. While the expectations and word count increased with each grade level, students described prior experiences in which their workload was amplified by first having to ascertain the expectations, a laborious and uncertain task that led to academic procrastination. Defined as task avoidance or delay, academic procrastination is a form of maladaptive coping that can have *'horrible'* consequences, as Riley observed, including negative performance (Moore, 2008) and low self-esteem (Owens & Newbegin, 2000). Tice and Baumeister (1997) found that high procrastination correlated with lower grades and higher levels of stress.

While students may provide explanations for their procrastination to avoid attributing it to low ability (e.g., "it's boring" or "I forgot"), Milgram, Marshevsky, and Sadeh (1995) found that students are more likely to delay tasks they see as unpleasant, boring, or difficult. Covington (1992) argued that excuses allow students to "repackage otherwise questionable actions, like not trying" (p. 83). Riley explained his negative prior experiences in contrast to his experience under the contract:

> In past years, I would forget some parts of the paper. I'd forget some assignments. I forgot, and I had some lates and stuff. And this year, I didn't have any lates. It was just very, very easy.

Missed deadlines are common among procrastinators (Aitken, 1982). As the contract shifted students' appraisal of workload demands, as explained in Chapter Four, students were more likely to '*just do it immediately*.' Under the contract, Riley felt able to complete the task because he had "*clarity and it was easier to comprehend 'cause it's just right there for you what's required*." This research suggests students are less likely to delay tasks over which they perceived a sense of control or manageability. Similarly, Sara (PE honors 11th), a high-achieving student with a history of success on the assessment, described a difference in task-oriented motivation under the grading contract:

> This year, I felt it was better. In previous years, maybe it felt more like a negative stress was put on me, like, "Oh, I don't want to do this. I don't want to do this." But this year, I felt like it was more positive in motivation, like, "Let's just get it done. Let's just get it done" kind of thing. Yeah, this year was that: "Let's just get it done."

As a result of minimizing task delay by "*just get[ting] it done*," she found that she had more "*more time to relax"* because she *"got everything done so early*." The ability to 'check' off the requirements also increased personal satisfaction with her work and move on. As she reflected on her prior experience, she observed much less before-the-due-date stress this year because of her new attitude:

> Last year, I felt like I was procrastinating a little bit. I feel like last year, it was like two days before it was due and I was still editing things, like fixing them and doing like my revisions from the Writing Center. Whereas this year, I went there and edited early, and I got it done, and then I had like a week and a half where I barely touched my paper.

At the time of the interview, Sara did not know her grade, but in the follow-up, she shared that she earned a 95, saying, "*It was comforting to know that this grading was a two-way street: if I completed all the work I needed to, at a high level, I would be rewarded with that grade*." To Sara, contract grading is a reciprocal and fair grading system, in which both student and teacher enter into a mutual agreement. The implication is that conventional grading practices are less rewarding or fair for students as the terms of the agreement are often obscured. With explicit terms, the contract system can increase goal-oriented effort to a desirable reward in adolescents, as Polczynski andd Shirland (1977) explained.

Additionally, many (25%, n=10) participants described enjoying their topic more this year than previously. While potentially coincidental, it may be a positive consequence of experiencing less worry and doubts about one's performance. For example, William (PE AP 12th, ELL) revealed that the grading contract, opposed to a points-based analytic rubric, curbed procrastination by helping him think less about evaluation and judgment and more about writing, saying:

> I would procrastinate or kind of push the work to a later time because I wasn't enjoying myself while I was writing, because I was always being a mathematician in my head, always calculating points. I love being an artist, and I always love being free, and I always love being very expressive. And that's what I could do with this year's research paper, so I got excited about writing, so I would get to it more quickly than in the past. . . . I wasn't thinking that I was writing for a school. It was more about writing this to share to other people and show that I really care about this, because since I'm not caring about grades as much, I was caring more about content.

In addition to lacking clarity of purpose, focusing on grades can also lead to task-avoidance. As James (FT honors 9th) explained:

> I think in a way it helps you to focus less on all the individual points that you should have and allow you to just write and be more open to that. Because in the past I've found people are stressing, making sure they check every single bullet point so they can get the best grade possible. And in a way, I think it could be helpful to have a specific so that you know exactly what you should be doing. But at the same time, it's important to have those liberties of knowing. You know, I can just do these and get a good grade and then still write the way I feel fits the topic.

Conversely, the grading contract allowed more students to get '*excited about writing*,' as Eric said, and they could '*just write*,' as James observed. The grading contract liberated some students to take risks by writing in '*the way [they] feel fits the topic*' while still having the assurance of '*get[ting] a good grade*.' As Danielewicz and Elbow (2009) theorized, some adolescents in this study "experience[d] the value and true 'payoff' of their 'work': the intrinsic rewards and pleasures of writing and learning, their tangible growth and development as they move from draft to draft" (p. 249). The majority, however, remained fixed on outcomes, the focus of the next section.

5.4.3 Grade Orientation

The quantitative analysis revealed that the majority of high schoolers were motivated to achieve success, rather than avoid failure, as revealed in Chapter Four; however, they also remained fixated on performance outcomes, rather than learning or interest. While students like William (PE AP 12th), quoted in the previous section, felt '*free*' of his previous concern about grades and thus *'excited about writing*' as a result of the grading contract, the majority of students I interviewed revealed a grade orientation fueled by

their parents' expectations, their desire to 'succeed' in society, or their need to feel valued. For example, Caleb (PE honors 11th), my pilot interviewee, said in his second interview, "*I contracted for an A because, well, I wanted an A on the paper, which is the main reason*," continuing—

> I think that it's just important to me in general to do well on every assignment for school because I attach my own self-esteem and worth to my grades in school.

Similarly, Alexander (FT regular 12th), who was several weeks away from his high school graduation at the time of the interview, made an observation about his peers and society:

> *Alexander:* We're really motivated by grades nowadays, not by intellectual effort.
> *Interviewer*: Do you feel like you are motivated by grades more than anything?
> *Alexander*: Yeah, it's what makes you succeed. You know, you can be the smartest person in the world, but if your grades don't reflect it, you're not going anywhere.

Alexander's reflection is tinged with anxiety as he considers being the "s*martest person in the world*" who still needs good grades to propel his life forward. While Elbow argues that contract grading, in theory, resists the "capitalism that permeates the classroom," the high schoolers in this study were overwhelmingly oriented toward external outcomes, including their parents' desire to "*see a decent semester grade*" (Tristan, PE regular 11th); "*have good grades*" (Vanessa, PE regular 12th); "*do well in the class*" (Eric, PE regular 11th); and "*succeed and get good grades*" (Jackson, PE honors 10th). Similarly, Riley (PE regular 11th) said:

> It's very important. My parents care. They care more about the GPA and grades than the paper.

Ainsley, the only student I interviewed who did not finish the assignment, described "*huge academic pressure and expectations*" from her family, continuing—

> I know that people are defined by their grades, and I'm not defined by my grades. But I don't know. With the environment at [the school], you can't control that

> parents put extreme pressure on their kids, so then you have a bunch of kids who are constantly talking about their grades and how they're worried, and that feeds into my worry.

After all, private-school parents have paid tuition for a college-preparatory education, as schools are engines of social mobility, as mentioned in Chapter One (McArthur, 2018), and by high school, many students have internalized the pressure to succeed, as Olivia and Isabella (both PE regular 10th) revealed:

> *Olivia*: I really stress getting good grades, so I feel like it's a little important to me to succeed, and I knew as long as I stayed on track and just did all the work on time, I knew I could get the grade that I wanted.
>
> *Isabella*: It's pretty important to me just because I try to keep my grades up as much as possible. I feel like if I don't do as much as I can, then I feel like I'm cheating myself out of being better.

Within the context of capitalist logic, '*cheating [herself] out of being better*' did not mean failing to learn; it meant not earning the highest grade she could. This finding is unsurprising given that students' beliefs and motivations are influenced by their context: the discourse in US society emphasizes final grades, GPAs, and rankings, rather than an interest in the subject or learning knowledge and skills.

Some (13%, n=5) adolescents cited their intrinsic motivation toward the task, yet this appeared to be a pre-existing belief that was unaffected by the grading contract. For example, Hannah (PE AP 12th, ELL), a hard-working student I taught as a 10th grader during her first year in America who then moved into the honors track, surprised me with her answer:

> *Interviewer*: How important is it for you or for you and your family that you do well or succeed on the research paper?

> *Hannah*: It's not very important if we're being honest. It's just—I do have good grades, but I don't really care about my grades that much as other people do. I'm not defined by my grades. Also, my family, they don't even care what college I go to. So, as I said, everything I do to pass. It all depends on my self-motivation.
> *Interviewer*: What do you think motivates you to take a class like AP Language or to do as well as you do in school?
> *Hannah*: What motivates me? It's mainly because I've always believed that what I'm doing right now is always going to benefit me in the future. So, I'm the only one who can get the good things out of this, so why not do better to help yourself, right?

She recognized that the skills she has learned will make college easier but it will also aid her future career: "*I just personally like reading the sources because it's a way of learning, and I believe this knowledge will help me because I want to be a filmmaker*." Additionally, Jackson (PE honors 10th) said, "*It's important to me because writing is something I really like to do*." This was not a result of contract grading but the intrinsic value of learning.

5.4.4 Goal Setting

The majority of participants (76%, n=334) set a highly desirable goal by contracting for an A, including 85% (n=213) of Prior-Experiencers. The previous year, under the grading rubric, only 43% of Prior-Experiencers earned an A on the final assessment. Covington (1992) observed students set unattainable goals as a self-handicapping strategy, but only if the goal includes a degree of irrationality, making their inevitable failure to achieve their goal would be the result of an exceedingly difficult goal, not their ability status; however, for each student I interviewed who contracted for an A, the choice was strategic, following a logical syllogism: if they contracted for A and fell short, they would earn a B, but if they contracted for a B and fell short, they would earn a C; thus, it was

most logical to at least strive for an A, even if it was likely out of reach, as Tristan (PE regular 11th) revealed:

> Truth be told, as I'm sitting here now, I think it might have been a better idea to go ahead and get the A contract and then just let stuff slip because when you contract out for the A, you're starting bar's up here instead of down here and you've got that little bit of give if you don't do well.

With clear expectations offered to them and a perceived manageable workload that appeared less than previous years, they *might as well try* for the highest grade. After the assessment, Micah (FT accommodated 9th) observed the benefit of contracting for a higher grade and regretted his decision to contract for a B:

> The grading contract made me feel confident enough in my own abilities that I could do the A one, and if I did choose the A one, then I might have gotten a better grade. Now, I know what it's like, and I can try to go for the A contract next time.

Choosing the A meant working for the destination of A with the contract as their map. The results of the quantitative analysis reveal the majority of Prior-Experiencers aimed for and then reached high standards of academic achievement. This corroborates Fishman's (2014) conclusion that "those who perceived the capability to achieve academic outcomes were more likely to feel internally obligated to produce such outcomes" (p. 685). As William (PE AP 12th, ELL) observed, "*I'm promising myself to get an A, and I will get an A. That kind of promise is very binding morally and, in the brain, too, and freeing—I got to be more passionate about writing.*" Worry and doubts about performance can stifle the creative process, yet for William, the contract paradoxically bound him goal while also freeing him to care less about grades and more

about learning. Additionally, Jackson (PE honors 10th) explained how the contract impacted his self-efficacy:

> I always, with any assignment, try to get at least a B, but on the grading contract, I said I was shooting for an A, and I always believed I could, but there was part of me that was like, if I got like a B, I'd be fine with it but I wouldn't have lived up to that grading contract, so part of me would have felt a little disappointed but like not too much.

The contract buoyed students to '*liv[e] up to the grading contract*,' or reach their highest potential. Setting the goal led to an increased desire to fulfill the goal, which involved less task delay, as Jackson said, "*I knew the way you succeed is by actually working on it*." While Jackson had earned an A the previous year as well, Vanessa (PE regular 12th) had a history of B grades on the assessment, which she credited to her busy competition schedule as an in-season athlete. Like Jackson and many others, she contracted for and then earned her first A on the assessment, saying, "*I just kind of wanted to push myself and see if I could do it even during my season. I still did it*."

Only 67% (n=143) earned an A under the contract system, meaning some did not reach their goal, but this was expected, following the syllogism. Alexander, for example, contracted for an A but earned a B, saying, "*I felt like it was well deserved. I mean, you know, I've never really done a research paper before. It was something new, and I did my best.*"

Ainsley (PE regular 11th), the only interviewee who earned an F, never submitted her paper. After two previous poor experiences with the assessment, she described herself as

"extremely" likely to avoid tasks at which she might not succeed and said, "*I have previous bad experiences with research papers that interfere with my ability to tell myself that I'm going to do good on this one*." Despite this reflection during the interview, she described "*start[ing] off pretty well with gathering sources*," continuing—

> But that's really kind of where it stopped, and after I missed sort of one or two assignments, that's where something in my brain just clicks to where I didn't think that. . . [pause] I didn't think that I could catch up, so I just kind of stopped.

Like her peers, Ainsley said the contract made the task *seem* smaller, but that perception shifted as she worked and found herself unable to meet the first progress check, which led her teacher to enter a zero since teachers are required to update their grade books each week; thus, the progress checks were entered into the grade-book as completion grades worth just 1% of the assessment's 20%, but as the first assignment of the semester, the 1% was the entire weight of the grade:

> It was very scary because it brought down my grade a lot, and it just continued. It felt hard to get up and start again.

Representing the 2.5% of students in this study who earned failing grades, Ainsley's experience further breaks down the myth of the motivational F and speaks to the unintended detrimental impact of grade applications that allow students to check their grades in real-time, as mentioned in Chapter One. Low grades can shut down learning by motivating students away from the task.

Fortunately, 97.5% of Prior-Experiencers earned a passing grade, and of Prior-Experiencers who contracted for an A, 99% (n=211) earned a passing grade on the final assessment. Under the grading contract, the number of adolescents who earned As

(n=143) increased by 19% and Bs (n=117) by 16%. This finding points to the potential positive impact of setting challenging goals to improve academic performance. This work corroborated a finding that researchers (Phan, 2009; Koenig, Eckert, & Hier, 2016; Schippers, Morisano, Locke, Scheepers, Latham, & De Jong, 2020) have observed in elementary school-aged children to college students: "explicitly setting goals can markedly improve performance at any given task" (Morisano, Hirsh, Peterson, Pihl, & Shore, 2010, p. 256). The mastery-based contract grading has the added benefit of folding goal-setting into the assessment to aid performance, thus aligning student-teacher expectations and enabling teachers to better support their students' goals. Since the teacher has written the contract, it ensures that each goal is well-articulated, specific, and measurable.

This finding reveals that under the contract grading system, high schoolers perceived reduced task demands, even while completing more work, and were less likely to procrastinate; as a result, they achieved significantly higher goals than the previous year.

5.4.5 Clarifying Subjective Terms

The interviews revealed a difference between Prior-Experiencers' and First-Timers' view of the subjective qualities of exceptional writing on the A contract. For the former, the A contract brought relief by clearly outlining the qualities of exceptional work and eliminating unnecessary areas in which they had previously invested significant time and energy. First-Timers were the only students I interviewed who reported challenges, at least initially, with the A contract's subjective expectations, pointing to the learning curve

of new experiences. The interviews revealed that not only was prior experience with academic writing beneficial but the contrast between the rubric and the helped Prior-Experiencers perceive the grading contract, including the subjective qualities, as beneficial.

First-Timers, like Daniela (FT honors 9th), described more challenges with subjective language. For her, the A contract was clear but demanding and even irritating: "*I knew I wasn't going to hit everything on the A contract because most of them were like, 'no mistakes at all.' And I'm like, 'OK, well, I'm obviously not perfect and I don't know why that's on the contract*.'" In a form of defensive pessimism (Norem & Cantor, 1986), she told her teacher that she contracted for a B, yet she described flipping back and forth between the contracts, quietly trying her best to fulfill the expectations for A. In her follow-up interview, her attitude shifted as she shared that she did, in fact, earn an A, which she credited to the contract's clarity and organization, a finding that ran throughout all grade levels and course types:

> *Interviewer*: What impact, if any, do you think the contract had on your academic performance?
> *Daniela*: [The contract] totally helped me because before we had to look at the rubric and multiple different things, like notes, and now, it's like, 'Here's a contract.' Before, I was, like, so scattered, disorganized, and for me, I like to have a list in front of me in order to follow it. I guess it's just easier for me, seeing the contracts and having it on one paper, both sides, you can flip easily that [the A and B contracts are] both basically the same 10 points. It was a lot easier so that helped me stay organized, so it definitely helped.

Similarly, in the pursuit of an A, Malcolm (FT honors 10th) thought it looked "*a little subjective*," but when he "*looked at it a little more*," the contract made the expectations "*real clear*." For Amelia (FT regular 10th), the subjectivity was an issue until the end:

> A teacher's definition of 'well-crafted' versus a student's definition of it could be different. And so when a student writes a paper and thinks that they did a really good job on it, the teacher might look at that and say that it's horrible, but there's no way to connect that gap, to *bridge the gap,* and make sure that the teacher and the student have the same viewpoint and the same ideas on what's 'well crafted.' (emphasis added)

As Danielewicz and Elbow (2009) explained, "the pedagogical principle behind fuzzy criteria is to highlight what we value most about good writing no matter how indefinable" (p. 252). By outlining the teacher's expectations, including the 'fuzzy criteria,' the contract grading makes the subjective nature of good writing more explicit, yet issues related to subjectivity and ultimately bias persist under any grading tool. As Inoue (2019) argues, subjective standards reproduce White language supremacy, punishing diverse students who "simply do not have enough time or fluency in the dominate White discourse" (p. 132).

While Prior-Experiencers found the expectations for A-quality work clearer than other assessment tools, First-Timers faced a natural learning curve while adjusting to the expectations of high school and academic writing. This finding suggests that the contract serves as a tool for learning but is best supplemented by feedback from the teaching during the writing process and examples of student writing, perhaps both well-written models to emulate or poor examples to learn from, to help '*bridge the gap*' between student and teacher.

To understand the impact of other factors and variables (Research Question Four) on student performance, as measured by their final grade, the next section analyzes students' grades alongside demographic variables.

5.4 Variables Impacting Performance

5.4.1 Regular-Level English

Of all course types (i.e., regular, honors/AP, and accommodated for SLDs), the academic performance of Prior-Experiencers enrolled in regular English courses most benefited most from the grading contract. As previously mentioned, regular students had a history of low or failing grades on the assessment at each grade level prior to the intervention and thus were recognized as a high-need group. Under the grading contract, 87% of regular 9th graders earned a B or higher, along with 91% (n=50) of regular 12th graders, only 64% (n=35) of whom earned a B or higher the previous year under the conventional grading system.

Additionally, First-Timers enrolled in regular English courses had the highest percentage of students (87%, n=82) who fulfilled the grading contract, compared to 81% (n=39) of honors courses and 75% (n=9) of accommodated courses (see Figure 5.1). This surprising finding again points to the impact of high goal setting: regular students, who are typically middle-ability, were the most likely to contract for an A. Compared to just 58% of honors students and 66% of students in accommodated English, 90% of students in regular

English courses contracted for an A. This finding suggests that their ambitions may have increased their effort in the direction of success. While fear of failure or disappointment held some students back, such as Micah in the previous section, many regular students were motivated to increase effort toward success, corroborating Polczynski and Shirland's (1977) findings with college students.

This finding continued for regular Prior-Experiencers, who had the most significant grade increase under the contract (M=86.96, SD=10.18) compared to last year (81.16, SD= 10.01), t(147)-6.934, p=.000 (see Table 5.5). Compared to their performance under conventional grading, the majority of regular Prior Experiencers (75%, n=100) earned a higher grade under the contract system, while 13% (n=19) earned a lower percentage of the same grade under both systems. Notably, compared to their performance under conventional grading, 17 regular students jumped two letter grades under the contract system: 10 moved from a D to a B and seven from C to A, thus showing significant improvement under the contract system.

Like FT students enrolled in regular courses, the majority (63%) of Prior Experiencers in regular courses also contracted for A, although just 23% earned an A the previous year with the rubric. As mentioned previously in this chapter, honors students described prior situations in which they had to ascertain their teacher's implicit expectations and thus were able to meet high standards without explicit directions, yet regular students may

have worked hard in the past but misplaced their effort, as Eric (PE regular 11th) observed:

> If you don't have a set requirement or what you have to do to get to a certain grade, then, who knows? Like, maybe put in all your effort, but you still get a D or something like that, but now I know exactly what I have to do.

Teachers may assume that the academic performance of middle-ability students is low due to a lack of effort; however, to '*put in all your effort*' and then '*still get a D'* reveals misplaced effort, often as a result of unclear expectations, and thus a surprising low grade. Conversely, the contract enabled more regular students to focus on beneficial tasks and find newfound success. The contract not only made him '*less worried to fail*,' but he observed newfound independence both in class, when his teacher did not check on him ("*I can see who they're checking on—it's people who are having more trouble*," he said), and at home as a result of his increased effort in the right direction, which subsequently increased his self-worth and self-confidence:

> When [my parents] saw me putting in that much time this year, they really just let me sit back. And I told them that instance of the teacher applauding me and of this contract. Normally, they're involved with me like, 'Hey, how you doing?' I think they checked in with me once, like, 'Are you on time?' or something like that. It was just all up to me. I was able to just do it on my own, but in a good way, and I was actually happy about that because it left just I didn't feel like a presence was over my head like, 'Hey, are you doing your work?' It was more like, 'You can do it. You're capable of doing your work by yourself.'

Increased task-oriented effort brought rewards beyond academic success, such as independence and self-confidence. As Covington (1992) observed, "Everything goes better with success" (p. 78). The analysis revealed that while all students of all course types experienced less stress under the grading contract, reduced stress was most

important for middle-ability students' academic success. Other research reveals that high-stress circumstances can lower the grades of middle-ability students, yet high-ability students are often better able to mobilize their resources during higher levels of perceived stress and anxiety to maintain high grades (Phillips, 1962, cited in Kirschenbaum, Napier, & Simon, 1971). Ultimately, for regular students, who are generally middle-ability, decreased perceptions of academic stress under the contract may have significantly increased academic performance.

5.4.2 Honors and AP Students with Prior Experience

Prior-Experiencers enrolled in honors and AP courses earned equivalent mean grades under both the grading contract and the rubric: M=90.62 with the rubric and M=90.78 under the contract (see Figure 5.5). Further analysis, however, revealed that 59% (n=55) earned higher grades under the contract system, while 41% (n=38) earned lower percentages of the same letter grade both years. Enrollment in honors-level courses at GSHS means fulfilling the high prerequisite. Upperclassmen in honors courses have a history of academic success, which reveals why honors students earned an A-average under both the conventional and contract system. For example, Joshua (PE AP 11th), who earned similar grades all three years, said:

> I liked that [the contract] laid out very clearly what was expected, but it wasn't that much different than the rubric for me, but, overall, I feel like my response to it was maybe slightly more like I knew what was happening, but there wasn't much of a difference for me.

As measured by outcomes, "*there wasn't much of a difference"*; he had earned an A on the assessment each year. High-ability students are often better able to mobilize their

resources during higher levels of perceived stress and anxiety to maintain high grades (Phillips, 1962, cited in Kirschenbaum, Napier, & Simon, 1971). For Jackson (PE honors 10th), who earned a perfect score on the paper under the grading contract, mobilizing his resources meant the ability to ascertain implicit expectations. He described prior experiences, in both English and other courses, in which he made '*a mental list of guidelines*' based on his notes in order to '*make sure [he was] doing it right*.' Like Joshua, his process was similar under the contract: "*I pretty much had the same approach with the last year's research paper, like, 'Okay, I gotta get this done and get it done correctly*.'" Where students enrolled in regular courses often reported being unsure how to do it '*correctly*' and then misplacing their effort, high-ability students, like Joshua and Jackson, had a strong approach to the paper and the ability to ascertain implicit expectations before contract and, thus, were more likely to find academic success under any grading tool.

Notably, the majority of honors students, like their regular peers, reported perceiving significantly less stress under the contract. Combined with the findings of Chapter Four, this reveals that the grading contract significantly reduced their perceptions of stress while keeping their grades high. For example, Evan (PE AP 11th) described high-stress prior experiences in which he struggled to ensure he had done all that was expected of him, which led to increased work, worry, and fear. Under the contract, however, he reported a significant decrease in what he called 'fretting':

Interviewer: Why is that?

> *Evan*: By, like, simplifying what was required of me, I was able to check off that I met the requirements more easily and, like, let go of the stress of doing badly on the paper.
> *Interviewer*: You mentioned 'letting go' earlier. Would you say that you were able to let go of anything else?
> *Evan*: Yeah, I was able to let go of stressing about the paper earlier. Instead of spending as much time being stressed about like, "Oh, did I get everything?" I was able to be like, "Oh, I got everything," and I was able to let go of being worried about that.

Under the contract, Evan was able to find satisfaction with his work after checking the contract. As Ella (PE honors 11th) explained, "*you can look back through your essay and make sure that all those points were covered,*" continuing—

> The contract just eased my mind knowing just when everything was due and then just seeing how the contract—I think I touched on this earlier—but how it laid everything out so that it led to the next step and supported the next step, and that made it easy.

This finding suggests that high-performance outcomes do not take into account honors students' psycho-emotional experiences, which should concern educators concerned with declining adolescent well-being (Twenge et al., 2017; APA, 2014; CASEL, 2018; APA, 2018). Of all course types, honors students can earn high grades while also experiencing tension and worry. The grading contract, however, 'eased' the majority of adolescents' worries and enabled them to have an easier and less stressful experience under the grading contract. While the physical and psychological correlates of stress are common among students at high-achieving high schools (Feld & Shusterman, 2015), this research studied the psychological impact of the grading contract, finding that it improved honors students' psycho-emotional experiences while they achieved high grades.

5.4.3 Students with Diagnosed Learning Disabilities

Of Prior Experiencers in courses accommodated for SLDs, 57% (n=23) earned higher or equivalent grades under both the contract and conventional grading systems, while 43% (n=10) earned lower grades. Accommodated courses had the lowest contract fulfillment rate (83%) and the highest rate of low or failing grades (9%). The previous year, 48% (n=11) earned an A, yet all (n=23) contracted for an A. Job and Klassen (2012) found that adolescents with LDs showed an overestimation of ability and overly optimistic efficacy, which may reflect a lack of metacognitive awareness. In this study, however, 30% (n=7) of students with LD reported that they were unsure if they had fulfilled the contract.

This finding warrants further investigation, yet I surmise that the contract may have aligned the expectations for accommodated classrooms, where students are expected to meet regular course expectations with added support via accommodations. The standardization of expectations may have lowered their grades under the contract, yet further study is needed to confirm this explanation.

5.4.4 Gender

Of the Prior-Experiencers, both males and females had significant grade increases from the rubric year to the contract year. Females earned a mean score of 90.04 (SD=7.4) under the contract compared to M=87.32 (SD=8.36) the previous year, t(154)=-3.858, p=.000, while males earned a mean score of 86.74 (SD=9.18) under the contract compared to M=83.32 (SD=9.68) with the previous year, t(125)=-3.569, p=.001. Further

analysis revealed that 61% (n=94) of females earned higher grades under the contract; 24% (n=37) earned lower percentages but equivalent letter grades both years; and 14% (n=22) earned a lower grade compared to conventional grading. Of males, 65% (n=80) earned higher grades under the contract; 18% (n=22) earned lower percentages but equivalent letter grades both years; and 18% (n=22) earned lower grades compared to conventional grading.

First-time females (89%, n=63) fulfilled the contract grading at a higher rate than FT males (80%, n=67). Honors-level FT students tested into the course, and of the 30 females and 18 males enrolled in honors English, 87% (n=26) of females and 72% (n=13) fulfilled the contract. This finding continued in regular English courses, where 46 of 55 (84%) males and 36 of 39 (92%) females fulfilled the contract.

While Reardon, Kalogrides, Podolsky, and Zárate (2019) found that females (grades 3-8) outperformed males in English classrooms across every school district in the country, it is notable that in this study, the majority of both males and females, both FT and with PE, reached high standards of achievement; however, of the students who earned non-passing grades (n=7), 71% (n=5) were males enrolled in regular English. Only one volunteered for an interview, but he did not respond to email requests to schedule the interview, leaving this finding unexplained. Future work should examine the experience of students with low or failing grades, particularly those enrolled in regular courses.

5.5 Summary of Findings

Quantitative Findings
Adolescents set and then reached high academic goals, as measured by the number of those who contracted for an A.
The majority of adolescents (90%) in the study fulfilled the contract to earn an A or B on the high-stakes assessment, thus meeting high standards of achievement.
97% of both first-time and returning students earned a passing grade (e.g., A, B, and C).
Returning students had the highest rate of contract fulfillment (94%), followed by transfer students (grades 10-12) with high school experience (88%). Ninth graders completing the assessment for the first time had the lowest rate of contract fulfillment: 84%.

Qualitative Findings
Interview participants found the grading contract's language overwhelmingly clear; however, several first-time students struggled with the subjective terms, revealing a learning curve for those adjusting to high school writing expectations.
Clarity of purpose and desirable grade outcomes facilitated productive task-oriented effort.
The academic performance of typically middle-ability adolescents enrolled in regular English courses most benefited from clarity of purpose.
Procrastination and task avoidance diminished as the result of a clear path to a desirable goal.
While reaching higher or comparable levels of academic performance, the overwhelming majority of students described more positive affect under the contract grading system.

The use of a mastery-based grading system, in which all students were asked to fulfill high academic standards, was overwhelmingly successful with the study participants, as

90% earned either an A or B, while 97% of participants earned a passing grade of C or higher. While this thesis study was the first to examine the impact of contract grading on high school students' academic performance the findings presented in this chapter corroborate the work of other researchers who found increased academic performance (Fairbanks, 1992; Lindemann & Harbke, 2011) and goal-oriented effort (Polczynski and Shirland, 1977; Parks & Zurhellen, 1978) under the contract with college students.

The qualitative analysis revealed that the academic performance of the adolescents in this study was a result of two key factors: clarity of purpose and productive, task-oriented effort. First, the grading contract offered clarity of purpose and expectations and task-specific goals. While the subjective language on the A contract was a challenge for some first-time students, the contract's language was overwhelmingly clear to the adolescents I interviewed. As a result, the grading contract facilitated productive, task-oriented effort in the direction of a desirable goal. While this improved the psycho-emotional experience for the majority of adolescents in this study, it most impacted the academic performance of typically middle-ability students enrolled in regular English courses. Under the conventional grading system, where they were often left to ascertain implicit expectations, regular students described experiences in which they either lacked direction or worked on the wrong task. As a result of clear purpose and expectations, regular students showed the strongest improvement under the grading contract, as the contract specified the necessary tasks for success and thus encouraged productive, task-oriented effort and goal achievement.

Chapter Six: Conclusion and Recommendations

6.1 Introduction

Based on the quantitative and qualitative analysis, the contract grading system, when compared to conventional grading practices, improved adolescents' psychological, emotional, and social conditions, which then had a direct and significant impact on academic achievement. The findings of this thesis also reveal that lowering perceptions of stress and meeting academic standards are not mutually exclusive; in fact, improved psycho-emotional conditions appeared to facilitate high academic performance. This study provides the first empirical support for implementing mastery-based contract grading in secondary classrooms to facilitate goal-setting and productive task-oriented effort, limit task avoidance and procrastination, and improve academic performance, particularly for middle-ability students, by clarifying expectations and helping them to place their effort on the right tasks. In this way, contract grading creates a sound pedagogy and positive learning environment that motivates task-oriented effort and promotes a success-oriented approach by providing a path to high standards. Additionally, mastery-based contract grading provides a grade that more accurately reveals students' achievement of proficiency or mastery.

As I have argued throughout this thesis, in a competitive, neoliberal education system, focusing on grades has caused significant achievement pressures that can not only exacerbate academic stress but hijack learning. For example, Schneider and Hutt (2014) observed the behavior of college students during the controversial Vietnam War, which

highlighted the dire implications of grades: since students needed to maintain good academic standing to avoid the war, they enrolled in 'easy' classes: "If learning had to occur along the way, so be it, but otherwise, students would do the least amount of work possible in order to attain the token of highest value" (p. 216). While the US has not had a draft since the 1960s, many students have adopted the approach of eligible draftees: make safe academic choices by avoiding risk-taking. When academic outcomes are perceived as more important than durable learning, the consequence is a different kind of death, one of the students' full potential.

Grades, as Kohn (2011) argued, "don't prepare children for the 'real world' — unless one has in mind a world where interest in learning and quality of thinking are unimportant" (n.p.). Edwards and Edwards (1999) also urge teachers to move away from conventional grading:

> If grades in public schools were abolished, colleges and universities would be forced to look for more meaningful criteria for admission. They would then have to address the problem of low correspondence between college achievement and success in the world of work. SAT and ACT scores would no longer be appropriate college admission data, and getting acceptable SAT and ACT scores would not drive the high school curriculum. (p. 262)

In May 2020, while I was conducting this research, the world-renowned University of California system voted to eliminate the admission requirement for the SAT and ACT (University, 2020) due to their unfairness toward diverse student groups (Coleman, 2011; Sackett et al., 2012). While a surprising decision, given the non-profit College Board's multi-million dollar 'assessment' programs, this was a promising move that may change

the course of secondary education, where decisions in higher education have a great influence.

Unfortunately, however, contract grading is not a panacea for the gross economic injustices in the US that permeate classrooms nor does it get rid of grades entirely, which may be best for students' learning, development, and well-being. Contract grading does, however, minimize negative impact of grades and allow students to work without being fueled by stress and anxiety. In this final chapter, I will argue that contract grading serves as a viable option for democratizing secondary classrooms by promoting student agency. While using the letters and numbers of the conventional grading system, the contract grading system stands in contrast to traditional practices by making an unusual arrangement with students that removes the looming "shadow of a grade" (Danielewicz & Elbow, 2009, p. 249). However, as Litterio (2018) discovered with her college students, the adolescents in this study were still performance-oriented, but importantly, many of the negative outcomes associated with such an orientation diminished as students reported many positive psycho-emotional outcomes that will be reviewed in this chapter. This work corroborates Litterio's (2018) and Inoue's (2019) argument that contract grading can break through barriers by making classrooms that require grades less threatening and therefore more student- and learning-centered.

While many educators may see this as a critical issue in contemporary classrooms, there has been a scarcity of research exploring techniques that combat this pervasive problem.

This thesis, however, makes a significant contribution by focusing on a more socially just assessment tool called contract grading that significantly improves psycho-emotional well-being in secondary school during high-stakes assessment. This work provides empirical evidence from an entire English Department, with 13 instructors who each used the contract grading system for the first time with a diverse group of adolescents. The original findings of this thesis provide empirical support for teachers, department chairs, and administrators to implement contract grading in secondary schools to create psycho-emotionally healthy learning environments and ultimately improve students' feelings about themselves and about academic writing.

This chapter includes a summary of the quantitative and qualitative results, a discussion of the study's strengths and limitations, and recommendations from the findings for teachers and administrators when developing and implementing contract grading.

6.2 Summary of Quantitative and Qualitative Results

The quantitative results of this thesis reveal that the contract grading system had a statistically significant impact on lowering adolescents' perception of evaluative threat, fear of failure, perceptions of stress, and social comparison in comparison to conventional grading practices. My analysis of the qualitative data examined what factors promoted such outcomes, finding that clear expectations for a desirable goal increased task manageability and fostered a sense of control over the assessment, which led them to appraise workload demands as challenging, increase task-oriented effort, and experience more confidence. The adolescents in this study reported a statistically significant

reduction in their perceptions of stress from workload demands and time constraints, compared to their prior experience under the conventional grading system or their expectations for the assessment, and then met high academic standards: the quantitative analysis found that 90% of adolescents in the study fulfilled the contract by reaching mastery or proficiency, including 94% of students with prior experience (grades 10-12) and 84% of first-time students (grades 9-12). Typically middle-ability regular students experienced the most significant grade improvement as a result of more focused effort.

6.3 Strengths and Limitations of the Study

6.3.1 Strengths

The first strength of this study is the rigorous mixed-methods design to examine adolescents' perceptions of stress (Research Question 1), self-worth protection behaviors (Research Questions 2), and academic performance (Research Question 3) under the grading contract during high-stakes writing assessment. In the first phase of collection, this study generated matched pairs quantitative data from 439 adolescents (grades 9-12, ages 13-19) from five established scales: Stress Appraisal Measure (SAM) (Peacock & Wong, 1989); Perceptions of Academic Stress Scale (PASS) (Bedewy & Gabriel, 2015); Primary and Secondary Appraisal Scale (PASA) (Gaab et al., 2005); and Self-Worth Protection Scale (SWPS) (Thompson & Dinnel, 2007). In the exploratory follow-up, 40 interviews with a diverse group of adolescents investigated six outcomes related to students' experiences with and perceptions of the grading contract: fear of failure; the

importance of ability; avoidance orientation; perception of workload demands; perceptions of academic stress; and ability doubts.

Another strength of this study is its direct examination of students' perceptions of stress and the use of qualitative data to understand students' experiences in context. Few studies have used interviews to examine students' self-worth protection behaviors, and only one other study (Spidell & Thelin, 2006) used qualitative data, in the form of reflection letters and video interviews, to understand the impact of contract grading on college students' experiences. This thesis contributes 40 semi-structured interviews with adolescents to understand their experiences under the grading contract.

Another strength is the large sample (n=439) of diverse adolescents of varying grade levels, abilities, and experiences, which can strengthen the trustworthiness of the study. The study included all high school grade levels; high-, middle-, and lower-ability high schoolers; those enrolled in honors and AP courses; international students and English Language Learners (ELL); and students with diagnosed learning needs, a growing group of students at GSHS and in the US (Riser-Kositsky, 2019). The population of this study was also more diverse than the predominately White school.

Additionally, the English Department's collaboration also strengthened the study, as explained in Chapter One. This is not to say, however, that all 12 teachers enthusiastically embraced the new approach, as the interviews revealed and I will further explain in the

next section, but the findings revealed that even when teachers expressed concern or anxiety about the approach to students, their psycho-emotional experience improved, as explained in Chapter Four.

Finally, given the survey's focus on psycho-emotional well-being, another strength is training in socio-emotional learning for grades 10-12, particularly identifying and managing emotions, which may have helped students answer the survey questions on self-worth protection behaviors, academic stress, and stress appraisal. Unfortunately, however, no such training occurred in 2020 for first-year students. The school's emphasis on socio-emotional learning may have impacted the findings, as I will discuss in the next section.

6.3.2 Limitations

This study had several limitations, the first of which is a potential Pygmalion effect, or other-imposed prophecy, as the emphasis on adolescents' well-being may have impacted, consciously or unconsciously, their perceptions and behavior. This research also took place within an institution that has proclaimed its commitment to academic rigor alongside socio-emotional well-being. While this pilot study analyzed the first pedagogical change to address the evaluative stress caused by high-stakes assessment, all participants in this study have received direct and palliative changes to address the decline in mental health. The impact of the grading contract may have been augmented by the school's culture.

One design limitation is the absence of a control group in the pre-post test design, which can be subject to internal validity threats. First, 10-12th graders were asked to reflect on their academic stress during last year's research paper, and in a year's time, they could have matured; however, with each grade level, the project demands also increase, making it more likely that academic performance will remain steady. Furthermore, this does not explain the positive change in the perspective of first-year students. To mitigate the threat of instrumentation, equivalent measures were used in both the pre- and post-survey with identical populations. To strengthen the study's design, four reliable and valid scales were adapted for this study with test-retest validity, and all *p*-values were corrected for Type 1 Error as a result of multiple testing using the Benjamini-Hochberg procedure. Finally, since students were given time to complete each survey in their English course, the time of day was not a factor since the survey was taken at the same time. The final threat is historical events: within the five-week period of the study, a global pandemic was on the rise; this, however, could explain an increase in perceived stress but not a decrease.

Another design limitation in the study is the focus on self-report data and subjective well-being. Perception, however, is key to the transactional model of stress appraisal, the theory that informs this research; thus, examining students' feelings is inherently subjective but valid and important. Additionally, in the pilot study, there was no difference in the objective measures of well-being, yet the subjective measures were statistically significant. Connected to this is the limitation of self-report bias as students

may have offered socially desirable answers. To encourage authentic responses, students were informed that their responses were anonymous unless they volunteered for an interview.

Additionally, a data limitation is the participation rate of 45%, a similar percent as the pilot study (Ward, 2021). While 1,216 students were emailed the survey, only 978 were eligible, after removing my students from 2019, during the pilot study, and those enrolled in online English, which is taught and run by a separate online department. To mitigate the influence of my positionality as a teacher at the school at which I was researching, I took great care to ensure that no student was pressured, either formally or informally, to participate. Additionally, no incentives were offered. As a result, the participation rate for the survey was lower than desired but still includes an acceptable participation rate.

An impact limitation is the small effect sizes of many of the significant quantitative results. This may be due to the fact that students completed the post-survey within a week of turning in the paper but before receiving the final grade, which marks the end of the process (Murray, 1972). Fortunately, the in-depth interviews I conducted to understand the significant results helped to not only understand "what is" through the distribution of variables across multiple dimensions through the quantitative strand; it also understood and explained "why" by uncovering participants' meanings behind each phenomenon in the qualitative strand (Merriam, 2009; Creswell, 2015). By the time I interviewed students, many had received the outcome, which brought relief when they reached their

goal and altered their perceptions. This means that their grade marked the conclusion of the process, not just the act of submitting it. Given that it was participants' first experience with contract grading, and many of whom had difficult prior experiences with the research paper, there was a lingering sense of uncertainty that may diminish over time to increase the effect of the grading contract. From the qualitative strand, it is clear that the grading contract improved students' psycho-emotional well-being.

The final limitation is the school's upper-middle class and predominantly White (67%) study body. While participants in the study were more diverse than the school's population (see demographic pie charts in Chapter 3), their socio-economic status and private school education have given them access to privileged discourses. Given the school's cost and location in an affluent neighborhood, many students, with white-collar professionals as parents, speak the English of schools and have attended well-rated, tuition-based institutions since birth. To Shor (2009), in this setting, the contract inadvertently functions in the neo-liberal logic of education, providing a seller's warranty to customers. Since others do not have the privilege of shopping around for a better school or grading plan, Shor writes that this marketization "privileges the strongest consumers, the already-privileged, further empowering the already powerful" (p. 16). Even with their advantages, however, many adolescents in this study have suffered. Coming from a working-class background, I attended public schools and personally have grave misgivings about private education, believing that it further privileges the affluent

and exacerbates injustice.[2] I hope this research benefits public institutions and provides confidence for implementing contract grading in secondary schools beyond GSHS.

6.4 Future Research

The findings of this thesis reveal many areas for further investigation. First, as previously mentioned, follow-up qualitative research is needed for students who earned low and failing grades to understand the factors that influence achievement, as well as students with diagnosed learning needs. Unfortunately, only one participant who earned a low or failing grade volunteered for an interview, but he did not respond to emails to schedule the interview. Additional research is also needed to understand the impact of the grading contract on students with diagnosed learning disabilities, the only identified group to earn lower grades under the contract. Further work can include interviews with teachers to test the hypothesis that the grading contract equalized the standards of performance for each grade level.

Future work can examine how the assessment approach impacts adolescents' feelings toward academic writing and research over time as they move through high school. Additional work should also follow GSHS students into college-level writing to examine their adjustment to college-level expectations, which may not be clearly communicated (Thomas & Rohwer, 1986).

[2] In May 2021, I accepted a position as a Coordinator for Equity-Focused Professional Learning within the newly developed Atlanta Public Schools Center for Equity and Social Justice

Finally, as McArthur (2018) argued, socially just assessment is fair to students as well as academic and administrative staff, who are burdened by the time-consuming labor of grading. Future work should examine how grading impacts English teachers' perceptions of workload stress, which causes physical symptoms, including eye strain, shoulder strain, and neck pain. This is particularly important for secondary educators with large class sizes.

6.5 Conclusion and Recommendations

> Only extraordinary education is concerned with learning; most is concerned with *achieving*; and for young minds, these two are very nearly opposites.
> —Marilyn French (quoted in Kohn, 2006)

I conclude this thesis with a call to action for secondary teachers concerned with equity in assessment and well-being: the findings of this thesis provide empirical support for using contract grading in secondary classrooms to improve psycho-emotional and academic performance. While GSHS is indebted to Danielewicz and Elbow's (2009) recommendation for implementing hybrid contract grading, we had to adapt their approach for secondary education; thus, the next section provides procedural details on creating and implementing contract grading in an English Department through a three-step process the GSHS English Department developed: identify the need(s) and theory for change; identify the learning goals for the assignment; and turn the learning goals into concrete, specific items for the contract.

6.5.1 Recommendations for Secondary Schools

Like many secondary teachers who must work closely with course teams, the GSHS English teachers met weekly as a department for an entire semester to redesign the project and implement contract grading. The process started with identifying the need for change, which was crucial for building a shared vision and helping us work through future challenges. Together, we brainstormed and discussed the following: how our grading practices impacted student learning; how they have impacted students' well-being as well as ours; and whether or not students were meeting their full potential under the current system. Through these in-depth conversations, most of the department agreed that our analytical rubric focused on areas to deduct points for missteps; punitive and harsh, it motivated students away from learning and the grades it produced did not accurately reveal what students could do.

After identifying the need for change, articulating the learning goals followed. This was the lengthiest step as it led to conversations about *what* teachers have done in their classrooms for years, not *why* they do what they do or if it produced durable learning. We interrogated each aspect of the process and eventually created the list of learning goals that we believed all students should and could reach, regardless of prior skill or training.

Next, to create the B contract, we turned each learning goal into an actionable task by asking, *What activity do they need to do to learn X?* Unlike conventional grading, in which students can cut corners in pursuit of the highest grade (further proof that grades

are separate from learning), the contract system values labor, which is often taken for granted but essential to learning (Inoue, 2019). Each concrete, explicit, and objective action then became the baseline proficiency contract, following Danielewicz and Elbow (2009) point that "almost any feature of writing or studying that a teacher wants to emphasize as crucial for learning can be handled as a yes/no binary requirement" (p. 252). Those I interviewed shared that the contract helped them see the importance of each step of the research process. Each step was equally weighted and important to growth.

The A contract then built upon the B contract; it did not demand a longer paper for the highest grade but a higher quality of work, which results from more labor. Each action on the B contract was infused with the subjective qualities of exceptional writing, such as '*compelling'* analysis and a '*well-crafted*' thesis. As previously mentioned, students can benefit from student-writing examples to animate each subjective writing quality, but as Inoue (2019) argues, "Asking students to meet a teacher's standards for an 'A' on a paper is not so clear, even when rubrics and examples are given—nor are such standards always attainable by anyone in the classroom" (p. 140). As mentioned in Chapter Five, subjective qualities reproduce White language supremacy and thus removing such quality-based judgements about student writing "builds equity among diverse students with diverse linguistic competencies" (p. 132).

Additionally, during each lesson, teachers emphasized the corresponding contract item and encouraged students to self-assess, using it as a holistic tool for not only assessment

but learning. In this way, contract grading draws on backward curriculum design, which stems from constructivist learning theory and "may improve instruction" by first identifying the desired results and identifying "what will constitute acceptable evidence of learning" (Mazur, 2018, p. 2). Only after creating the contract, which requires significant time investments, does the teacher plan the individual lessons. The result, however, is transformative and meaningful learning experiences that improve the quality of instruction and develop the key socio-emotional competency of self-regulation. While GSHS has taught this research unit for decades, and thus no lesson was created from scratch, the grading contract likely focused and refined their existing lessons, even without teachers' conscious knowledge, which helped students to see the reason and purpose behind each task, as the interviews revealed.

Additionally, to "affirm positive steps toward personal accountability," students received feedback if they fell short of their goal and a stamp when they reached it (Wormeli, 2006, p. 25). One teacher who used a stamp told me that this nuance fostered positive affect and self-confidence, although no student I interviewed brought up the stamp, perhaps because it was overshadowed by the grade-book. Many secondary teachers are required to update their grade books weekly and thus cannot follow Inoue (2019) or Danielewicz and Elbow (2009) who only make note of students who fall below the contract. To accommodate this, GSHS teachers made 1% of the project's 20% total weight reflect the initial progress checks, or mini-deadlines, which satisfied the school's requirement and provided transparency to students and parents about their progress. The final paper and contract

grade, then, was worth 19% of their overall course grade. After the success of this study, an English Department chair and I discussed seeking administrative approval to remove the mini-deadline grade to further eliminate grade orientation and help the persistence of those who fell behind during the process. In March 2021, I began data generation for follow-up study to analyze the impact of electronic grade-books on goal orientation and academic self-efficacy, and self-handicapping. Grades send powerful messages, but this research also emphasized the importance of teacher messaging.

Mastery-based contract grading begins with confidence in students' abilities and the belief that almost all students can reach high standards, whether A or B. While contract grading does not mean lowering expectations, even if subjectivity is partially removed, it does make learning goals and expectations explicit to students to improve their well-being, experience, and outcomes. Since contract grading is more democratic and participatory than conventional grading, students' decisions *do* matter because they impact their subsequent actions and behaviors and task-oriented motivation.

Unfortunately, not all adolescents in this study received positive messages that emphasized their agency in the learning and assessment process. A quarter of the (25%, n=10) students I interviewed reported that the contract was undermined by their teacher who told them that their decision did not matter or that they held total control, as I explained in Chapter Four. These few teachers, it seems, were concerned that students would take advantage of the contract by not taking the project seriously because of the

new assessment. This approach views students as lazy, sneaky, or untrustworthy, yet ironically, it is conventional grading practices that promote such behavior, leading some to or cut-corners or cheat in pursuit of a good grade, rather than work in pursuit of durable learning. Contract grading can facilitate an environment that supports students' growth but it must begin with students receiving efficacy appeals that emphasize their agency and meaningful participation. While some teachers strive to motivate students through grades and rely on fear appeals, contract grading "capitalize[s] on the extrinsic motivation by sending a message that many adolescents need to hear: 'Don't panic or be anxious; just do the work!" (Danielewicz & Elbow, 2009, p. 257).

6.5.2 Conclusion

Riley (PE regular 11th): This year was straightforward. I had a general path in my mind of where to go, and that gave me comfort, if you will, and how I should do my paper and less risk of failure.
Interviewer: You just used the word "comfort." What gave you comfort?
Riley: The contract's just, like, a piece of paper, but in a way, what it gives me as a student—the guidelines for the paper—gives me comfort.

This study reveals that contract grading can reduce evaluative threat for high schoolers, thereby reducing perceptions of stress (Research Question One) and self-worth protection behaviors (Research Question Two), while also improving academic performance (Research Question Three), particularly for typically middle-ability students enrolled in regular English courses (Research Question Four). Throughout this thesis, I have argued for the positive benefits of contract grading during high-stakes assessment. Overwhelmingly, the adolescents I interviewed revealed that the contract grading system

contrasts with conventional grading practices by offering an unusual clarity of purpose that reduced evaluative threat, thereby minimizing perceptions of stress and self-worth protection behaviors, that facilitated productive task-oriented effort that led more students to reach high levels of excellence.

The findings of this thesis can have rebounding effects on writing assessment in secondary schools, where contract grading is underused but beneficial. Contract grading is a holistic tool for teaching and learning that can reduce evaluative threat during high-stakes writing assessment to improve psycho-emotional well-being and academic performance. As Riley (PE 11th regular) said, the contract was "*just a piece of paper*" but it offered "*comfort*" that reduced students' perceptions of stress, reduced their self-worth protection behaviors, and enabled them to reach high goals. This is not to say that five weeks of contract grading eradicated students' concern with grades—high schoolers, after all, exist in complex social environments with profound influence on their goals and behavior—but it does appear to allow them to work without being fueled by stress.

Educators and school administrators must work to combat the mental health crisis facing contemporary adolescents, who are increasingly worried about grades as a direct result of the marketization of knowledge, as I outlined in Chapter One. Contract grading resists the neoliberal logic of education, valuing cooperation over competition; rather than sorting students into winners and losers, contract grading aims to help all students meet their fullest potential. The significant and original findings of this thesis revealed the

psycho-emotional benefits of contract grading on adolescents. I conclude, then, by arguing that contract grading, if used more widely in secondary schools, has the potential to shift the culture of secondary education from obtaining grades to setting and achieving high learning goals.

References

Aborn, M. (1953). The influence of experimentally induced failure on the retention of material acquired through set and incidental learning. *Journal of Experimental Psychology, 45*(4), 225-231.

Akgun, S., & Ciarrochi, J. (2003). Learned resourcefulness moderates the relationship between academic stress and academic performance. Educational Psychology, 23, 287–294. http://dx.doi.org/10.1080/ 0144341032000060129

Albdour, M., Lewin, L., Kavanaugh, K., Hong, J., & Wilson, F. (2017). Arab American Adolescents' Perceived Stress and Bullying Experiences: A Qualitative Study. *Western Journal of Nursing Research, 39*(12), 1567-1588.

Allison, A.L, Peres, J.C, Boettger, C., Leonbacher, U., Hastings, P.D, & Shirtcliff, E.A. (2012). Fight, flight, or fall: Autonomic nervous system reactivity during skydiving. *Personality and Individual Differences, 53*(3), 218-223.

Aiken, J.H., Koplow, D.A., Lerman, L.G., Ogilvy, J. P. & Schrag, P.G. (1985). The Learning Contract in Legal Education. *Maryland Law Review, 44*, 1047-1111.

Aitken, M. E. (1982). A personality profile of the college student procrastinator. *Dissertation Abstracts International Section A: Humanities and Social Sciences, 43*(3-A), 722–723.

Ainslie, R. C., Shafer, A., and Reynolds, J. (1996). Mediators of adolescents' stress in a college preparatory environment. *Adolescence, 31*(124), 913-24.

American Institute of Stress. (2018). What is Stress? Retrieved from https://www.stress.org/daily-life

American Psychological Association (APA) (2018). Stress in America: Stress and Generation Z.

Amsden, S. (1970). Have You Ever Tried Contracting for Grades? *English Journal, 59*(9), 1279-1282.

Arsenio, W., & Loria, S. (2014). Coping with Negative Emotions: Connections with Adolescents' Academic Performance and Stress. *The Journal of Genetic Psychology, 175*(1), 76-90.

Babakus, E, & Mangold, W G. (1992). Adapting the SERVQUAL scale to hospital services: An empirical investigation. *Health Services Research, 26*(6), 767-786.

Backward Design. (2018). *The SAGE Encyclopedia of Educational Research, Measurement, and Evaluation*, 2018.

Balcetis, E., & Dunning, D. (2006). See What You Want to See. *Journal of Personality and Social Psychology, 91*(4), 612-625.

Balcetis, E., & Dunning, D. (2010). Wishful Seeing: More Desired Objects Are Seen as Closer. *Psychological Science, 21*(1), 147-152.

Barkley, M.V. (1975). To grade or not to grade. *Illinois Teacher*, 38(4),201-203.

Barnett, M.A. (1989). "Writing as a Process." *The French Review* (63.1),31-44

Becker, D. (2013). *One Nation Under Stress*. New York: Oxford University Press.

Bedewy, D., & Gabriel, A. (2015). Examining perceptions of academic stress and its sources among university students: The Perception of Academic Stress Scale. *Health Psychology Open*. https://doi.org/10.1177/2055102915596714

BERA (2011). *Ethical Guidelines for Educational Research*. London: BERA.

Borman, D.L. (2018). De-Grading Assessment: Rejecting Rubrics in Favor of Authentic Analysis. *Seattle University Law Review*, Vol. 41:1.

Bouffard-Bouchard, T. (1990). Influence of Self-Efficacy on Performance in a Cognitive Task. *The Journal of Social Psychology, 130*(3), 353-363.

Bowers, G. & Howard, M. (1975). Sameness is not a virtue. *English Journal*. 64(8),49-52.

Bowman, J. S. (1977). Winning the Grading Game: A "Modified" Contract Approach. *Teaching Political Science, 4*(2), 225-234.

Bracho-Sanchez, E. (2019). Suicide rates in girls are rising, study finds, especially in those age 10 to 14. Retrieved May 1, 2019, from https://www.cnn.com/2019/05/17/health/suicide-rates-young-girls-study/index.html

Breland, H. (1983). The Direct Assessment of Writing Skills: A Measurement Review. *ETS Research Report Series, 1983*(2), I-23.

Brimi, H.M. (2011). Reliability of Grading High School Work in English. *Practical Assessment, Research & Evaluation, 16*(17), 12.

Brookhart, S.M. (2018). Appropriate Criteria: Key to Effective Rubrics. *Frontiers in Education (Lausanne), 3*, F2018-04-10, Vol.3.

Brookhart, S.M, Guskey, T.R, Bowers, A.J, Mcmillan, J.H, Smith, J.K, Smith,

L.F, Welsh, M.E. (2016). A Century of Grading Research: Meaning and Value in the Most Common Educational Measure. *Review of Educational Research, 86*(4), 803-848.

Burstein, B,. Agostino, H., and Greenfield, B. (2019). Suicidal Attempts and Ideation Among Children and Adolescents in US Emergency Departments, 2007-2015. *JAMA Pediatrics.* Published online April 08, 2019.

doi:10.1001/jamapediatrics.2019.0464

Callahan, R. (1979). Contract Grading and Student Performance. *Journal of Instructional Psychology, 6*(3), 3.

Callaway, H. (1992). Ethnography and experience: Gender implications in fieldwork and texts. In J. Okely & H. Callways (Eds.), Anthropology and autobiography (pp. 29-49). New York, NY: Routledge.

Carpenter, R. (2016). A Review of Instruments on Cognitive Appraisal of Stress. *Archives of Psychiatric Nursing, 30*(2), 271-279.

CASEL (2018). "Respected: Perspectives of Youth on High School & Social Emotional Learning."

Chambel, M., & Curral, L. (2005). Stress in Academic Life: Work Characteristics as Predictors of Student Well‐being and Performance. *Applied Psychology, 54*(1), 135-147.

Churchill, Gilbert A, & Peter, J. Paul. (1984). Research Design Effects on the Reliability of Rating Scales: A Meta-Analysis. *Journal of Marketing Research, 21*(4), 360.

Cikrikci, Özkan, & Odaci, Hatice. (2016). The Determinants of Life Satisfaction Among

Adolescents: The Role of Metacognitive Awareness and Self-Efficacy. *Social Indicators Research, 125*(3), 977-990.

Clark, R., Nguyen, F, & Sweller, J. (2006). Efficiency in learning evidence-based guidelines to manage cognitive load (1st ed.). San Francisco: Jossey-Bass.

Coleman, R. (2011). Stratification, Inequality, and the SAT: Toward an SAT-Optional Movement. *Georgetown Journal on Poverty Law & Policy., 18*(3), 507-531.

Covington, M.V. (1992). *Making the grade: A self-worth perspective on motivation and school reform*. Cambridge ; New York: Cambridge University Press.

Covington, M.V. (1984). The Self-Worth Theory of Achievement Motivation: Findings and Implications. *The Elementary School Journal, 85*(1), 5-20.

Covington, M.V., & Beery, R.G. (1976). *Self-worth and school learning*. New York: Holt, Rinehart and Winston.

Creswell, J.W. (2015). *A Concise Introduction to Mixed Methods Research.* Los Angeles: Sage.

Creswell, J.W. & Plano Clark, V.L. (2017). Designing and conducting mixed methods research (3rd ed.). Thousand Oaks, CA: Sage.

Creswell, J., & Guetterman, T.C. (2018). *Educational research. Planning, conducting and evaluating quantitative and qualitative research (*Sixth ed.). New York: Pearson.

Danielewicz, J., & Elbow, P. (2009). A Unilateral Grading Contract to Improve Learning and Teaching. *College Composition and Communication,61*(2), 244-268.

de Anda, D., Baroni, S., Boskin, L., Buchwald, L., Morgan, J., Ow, J., Gold, J.S., Weiss,

R. (2000). Stress, stressors and coping among high school students. *Children and Youth Services Review, 22*(6), 441-463.

De Castella, Krista, Byrne, Don, & Covington, Martin. (2013). Unmotivated or motivated to fail? A cross-cultural study of achievement motivation, fear of failure, and student disengagement. *Journal of Educational Psychology, 105*(3), 861-880.

Droit-Volet, S., & Meck, W. H. (2007). How emotions colour our perception of time. *Trends in Cognitive Sciences, 11*(12), 504-513.

Droit-Volet, S. & Gil, S. (2009). The time–emotion paradox. *Philosophical Transactions. Biological Sciences, 364*(1525), 1943-1953.

Droit-Volet, S., Fayolle, S., Lamotte, M., & Gil, S. (2013). Time, Emotion and the Embodiment of Timing. *Timing & Time Perception, 1*(1), 99-126.

Dweck, C. S. (1986). Motivational processes affecting learning. *The American Psychologist, 41*(10), 1040-1048.

Dweck, C. S., & Elliott, E. S. (1983). Achievement motivation. *Handbook of child psychology*. New York: Wiley.

Dweck, C.S, & Leggett, E.L. (1988). A Social-Cognitive Approach to Motivation and Personality. *Psychological Review, 95*(2), 256-273.

Eastridge, H. E. (1975). The Contract Method of Instruction. *American Secondary Education, 5*(4), 30-33.

Eccles, J., & Jacobs, J. (1986). Social Forces Shape Math Attitudes and Performance. *Signs, 11*(2), 367-380.

Edwards, C.H, & Edwards, L. (1999). Let's End the Grading Game. *The Clearing House,*

72(5), 260-263.

Elbow, P. (1983). Embracing Contraries in the Teaching Process. *College English, 45*(4), 327-339.

Elbow, P. (1997). Taking time out from grading and evaluating while working in a Conventional system. *Assessing Writing, 4*(1), 5-27.

Eiland, L, & Romeo, R.D. (2013). Stress and the developing adolescent brain. *Neuroscience, 249*, 162-171.

Fairbanks, P. J. (1992). Treating Mathematics Anxiety: The Optional Contract. *Mathematics Teacher, 85*(6), 428 -30.

Farber, J. (1990). Learning How to Teach: A Progress Report. *College English, 52*(2), 135-141.

Feld, L.D., & Shusterman, A. (2015). Into the pressure cooker: Student stress in college preparatory high schools. Journal of Adolescence, 41(C), 31-42.

Feldman, J. (2019). *Grading for equity: What it is, why it matters, and how it can transform schools and classrooms*.

Festinger, L. (1954) A Theory of Social Comparison Processes. Human Relations, 7, 117-140.

Fleischmann, R., Harrer, M., Zarski, A., Baumeister, H., Lehr, D., & Ebert, D. (2018). Patients' experiences in a guided Internet- and App-based stress intervention for college students: A qualitative study. *Internet Interventions, 12*, 130-140.

Fishman, E. J. (2014). With great control comes great responsibility: The relationship

between perceived academic control, student responsibility, and self-regulation. *British Journal of Educational Psychology, 84*(4), 685-702.

Friedman, B. A., & Mandel, R. G. (2009). The Prediction of College Student Academic Performance and Retention: Application of Expectancy and Goal Setting Theories. *Journal of College Student Retention : Research, Theory & Practice, 11*(2), 227-246.

Gaab, J. (2009). PASA – Primary Appraisal Secondary Appraisal. *Verhaltenstherapie, 19*(2), 114-115.

Galloway, M., Conner, J., & Pope, D. (2013). Nonacademic Effects of Homework in Privileged, High-Performing High Schools. The Journal of Experimental Education, 81(4), 490-510.

Guskey, T. R. (2013). The Case Against Percentage Grades. *Educational Leadership, 71*(1), 68.

Graff, N. (2011). "An Effective and Agonizing Way to Learn": Backwards Design and New Teachers' Preparation for Planning Curriculum. *Teacher Education Quarterly (Claremont, Calif.), 38*(3), 151-168.

Grant, H., & Dweck, C.S. (2003). Clarifying Achievement Goals and Their Impact. *Journal of Personality and Social Psychology, 85*(3), 541-553.

Grau, H.J. (1999). “Streamlined” contract grading - a performance-measuring alternative to traditional evaluation methods. *Journal of College Science Teaching,* 254.

Grix, J. (2010). *The Foundations of Research*. Basingstoke: Palgrave Macmillan, 51-61.

Hanesworth, P., Bracken, S., & Elkington, S. (2019). A typology for a social justice

approach to assessment: Learning from universal design and culturally sustaining pedagogy. *Teaching in Higher Education, 24*(1), 98-114. doi:10.1080/13562517.2018.1465405

Hancock, D. R. (2001). Effects of Test Anxiety and Evaluative Threat on Students' Achievement and Motivation. *The Journal of Educational Research (Washington, D.C.), 94*(5), 284-290.

Hare, O.A, Wetherell, M.A, & Smith, M.A. (2013). State anxiety and cortisol reactivity to skydiving in novice versus experienced skydivers. *Physiology & Behavior, 118*, 40-44.

Hassencahl, F. (1979). Contract Grading in the Classroom. *Improving College and University Teaching, 27*(1), 30-33.

Hjeltnes, A., Binder, P., Moltu, C., & Dundas, I. (2015). Facing the fear of failure: An explorative qualitative study of client experiences in a mindfulness-based stress reduction program for university students with academic evaluation anxiety. *International Journal of Qualitative Studies on Health and Well-being, 10*(1), 27990.

Hiller, T.B., & Hietapelto, A.B. (2001). Contract Grading: Encouraging Commitment to the Learning Process through Voice in the Evaluation Process. *Journal of Management Education, 25*(6), 660-84.

Inoue, A.B. (2019). *Labor-Based Grading Contracts: Building Equity and Inclusion in the Compassionate Writing Classroom* Perspectives on

Writing. The WAC Clearinghouse; University Press of Colorado. https://doi.org/10.37514/ PER-B.2019.0216

Inoue, A. B. (2019). How Do We Language So People Stop Killing Each Other, or What Do We Do about White Language Supremacy? *College Composition and Communication, 71*(2), 352-369.

Inoue, A. B. (2014). Theorizing Failure in US Writing Assessments. *Research in the Teaching of English, 48*(3), 330-352.

Jackson, C. (2010). "Fear in Education." *Educational Review*, vol. 62, no. 1, 2010, pp. 39–52.

Jackson, K., & Bazeley, P. (2019). *Qualitative data analysis with NVivo*. Los Angeles: Sage.

James, S. M. (1977). The evolution of a language arts program for pre- and early adolescent students. *English Journal*, 197,6(4),47.

Jamieson, J. P., Crum, A. J., Goyer, J. P., Marotta, M. E., & Akinola, M. (2018). *Optimizing stress responses with reappraisal and mindset interventions: An integrated model* Routledge. doi:10.1080/10615806.2018.1442615

Jamieson, J. P., Peters, B. J., Greenwood, E. J., & Altose, A. J. (2016). Reappraising stress arousal improves performance and reduces evaluation anxiety in classroom exam situations. *Social Psychological and Personality Science, 7*(6), 579-587. doi:10.1177/1948550616644656

Jaschik, S. (2018). "Suicide Note Calls Out Pressure on Students." *Inside Higher Ed*. Retrieved July 19, 2019, from

https://www.insidehighered.com/admissions/article/2018
/02/12/suicide-note-16-year-old-renews-debate-about-pressure-top-highschools

Janzow, F., and J. Eison. (1990). Grades: Their influence on students and faculty. *New Directions for Teaching and Learning* 42: 93-102.

Jayanthi, P., Thirunavukarasu, M., & Rajkumar, R. (2015). Academic stress and depression among adolescents: A cross-sectional study. *Indian Pediatrics, 52*(3), 217-9.

Job, J., & Klassen, R. (2012). Predicting performance on academic and non-academic tasks: A comparison of adolescents with and without learning disabilities. *Contemporary Educational Psychology, 37*(2), 162-169.

Johnston, J.M., & O'Neill, G. (1973). The analysis of performance criteria defining course
grades as a determinant of college student academic performance. *Journal of Applied Behavior Analysis, 6*(2), 261-268.

Johnson, R.B., Onwuegbuzie, A.J. & Turner, L.A. (2007), "Toward a definition of mixed methods research", Journal of Mixed Methods Research, Vol. 1 No. 2, pp. 112-33.

Kessler, R., Berglund, P., Demler, O., Jin, R., Merikangas, K., & Walters, E. (2005). Lifetime Prevalence and Age-of-Onset Distributions of DSM-IV Disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry, 62*(6), 593-602.

Kirschenbaum, H., Napier, R., & Simon, S. B. (1971). *Wad-Ja-Get?: The Grading Game in American Education*. New York City, NY: Hart Publishing Company.

Kirschenbaum, D., & Riechmann, S. (1975). Learning with Gusto in Introductory Psychology. *Teaching of Psychology, 2*(2), 72-76.

Koenig, E., Eckert, T., & Hier, B. (2016). Using Performance Feedback and Goal Setting to Improve Elementary Students' Writing Fluency: A Randomized Controlled Trial. *School Psychology Review, 45*(3), 275-295.

Kohn, A. (2006). The Trouble with Rubrics. *English Journal, 95*(4), 12.

Kohn, A. (2011). Poor Teaching for Poor Children... In the Name of School Reform. *Education Week, 30*(29), 32.

Kokus, N. & Musof, L. (1975). Assessing English: An Alternative. *English Journal*.64(3),73-76.

Kyriacou, C. (2003). *Helping troubled pupils.* Cheltenham, England: Nelson Thornes.

Lay, C.H., & Schouwenburg, H.C. (1993). Trait Procrastination, Time Management, and Academic Behavior. *Journal of Social Behavior and Personality, 8*(4), 647-662.

Lazarus, R., Folkman, S. (1984). *Stress, appraisal, and coping*. New York: Springer Pub.

Leahy, R. (1980). Writing Made Possible: A Contract Approach. *Improving College and University Teaching*, 28 (4), 155-157.

Leaper, C., Farkas, T., & Brown, C. (2012). Adolescent Girls' Experiences and Gender-Related Beliefs in Relation to Their Motivation in Math/Science and English. *Journal of Youth and Adolescence, 41*(3), 268-282.

Lewis, J. (1993). Multipoint Scales: Mean and Median Differences and Observed Significance Levels. *International Journal Human Computer Interaction*. 5, 383-392. 10.1080/10447319309526075.

Lewis, L.A, & Killip, D.E. (1971). Contract grading: An alliance for learning. *Journal of Dental Education, 35*(10), 616-620.

Lindemann, D., & Harbke, C. (2011). Use of Contract Grading to Improve Grades Among College Freshmen in Introductory Psychology. *SAGE Open, 1*(3), 1-7.

Litterio, L. M. (2016). Contract Grading in a Technical Writing Classroom: A Case Study.

Journal of Writing Assessment, 9(2).

Litterio, L. (2018). Contract grading in the technical writing classroom: Blending community-based assessment and self-assessment. *Assessing Writing, 38*, 1-9. *Community Colleges*, *15*(1), 29–42.

Luthar, S.S, & Barkin, S.H. (2012). Are affluent youth truly "at risk"? Vulnerability and resilience across three diverse samples. *Development and Psychopathology, 24*(2), 429-449.

Luthar, S., Barkin, S., & Crossman, E. (2013). "I can, therefore I must": Fragility in the upper-middle classes. *Development and Psychopathology, 25*(4pt2), 1529-1549. doi:10.1017/S0954579413000758

Luthar, S. S., & Latendresse, S. J. (2005). Children of the affluent: Challenges to well-being. *Current Directions in Psychological Science, 14*(1), 49-53. doi:10.1111/j.0963-7214.2005.00333.x

Martin, A. J., & Marsh, H. W. (2008). Academic buoyancy: Towards an understanding of students' everyday academic resilience. *Journal of School Psychology, 46*(1), 53-83. doi:10.1016/j.jsp.2007.01.002

Martin, A.J, Marsh, H. W, & Debus, R.L. (2001). Self-Handicapping and Defensive Pessimism. *Journal of Educational Psychology, 93*(1), 87-102

Maxwell, J. A. (2005). Qualitative research design: An interactive approach. Thousand Oaks, Calif: Sage Publications.

McArthur, J. (2016). Assessment for social justice: The role of assessment in achieving social justice. *Assessment & Evaluation in Higher Education, 41*(7), 967-981. doi:10.1080/02602938.2015.1053429

McArthur, J. (2018). *Assessment for social justice perspectives and practices within higher education* (1st ed.). London: Bloomsbury Academic.

McCown, W., Petzel, T., & Rupert, P. (1987). An experimental study of some hypothesized behaviors and personality variables of college student procrastinators. *Personality and Individual Differences, 8*(6), 781-786.

McEwen, B. S. (2012). Brain on stress: How the social environment gets under the skin. *Proceedings of the National Academy of Sciences of the United States of America, 109* (Supplement 2), 17180-17185. doi:10.1073/pnas.1121254109

McLaughlin, J. G. (1961). Contracts Help Solve Project Problems. *The Clearing House, 35*(7), 418.

McShane, S. L., & Von Glinow, M. A. (2005). Organizational behavior: Emerging realities for the workplace revolution. McGraw-Hill Irwin, Boston, MA.

Mental Health America (MHA). (2019). The State of Mental Health in America. Retrieved September 19, 2020, from https://mhanational.org/sites/default/files/2019-09/2019%

20MH%20in%20America%20Final.pdf

Merriam, S. B. (2009). Qualitative research: A guide to design and implementation. San Francisco, CA: Jossey-Bass.

Middleton, Michael J, & Midgley, Carol. (1997). Avoiding the Demonstration of Lack of Ability. *Journal of Educational Psychology, 89*(4), 710-718.

Milgrim, N., Marshevsky, S., & Sadeh, C. (1995). Correlates of academic procrastination: Discomfort, task aversiveness, and task capability. *The Journal of Psychology, 129*(2), 145-55.

Moore, R. (2008). Academic Procrastination and Course Performance among Developmental Education Students. *Research and Teaching in Developmental Education, 24*(2), 56-67.

Morazes, J. L. (2016). Educational background, high school stress, and academic success. *Children and Youth Services Review, 69*, 201-209.

Morisano, D., Hirsh, J., Peterson, J., Pihl, R., & Shore, B. (2010). Setting, Elaborating, and Reflecting on Personal Goals Improves Academic Performance. *Journal of Applied Psychology, 95*(2), 255-264.

Murray, D. (1972). "Teach Writing as a Process Not Product." Ed. Villanueva, V. *Cross-Talk in Comp Theory*. 2nd. Urbana: National Council of Teachers of English, 1972. 3-6.

Ng, V., Koh, D., & Chia, S. (2003). Examination Stress, Salivary Cortisol, and Academic Performance. *Psychological Reports, 93*(3_suppl), 1133-1134.

Nicholls, J. (1979). Quality and equality in intellectual development: The role of

motivation in education. *American Psychologist, 34*(11), 1071-1084.

Nieto, S., & Bode, P. (2007). *Affirming diversity: The sociopolitical context of multicultural education* (5th ed.). Boston: Allyn & Bacon.

Norem, J., & Cantor, K. (1986). Anticipatory and post hoc cushioning strategies: Optimism and defensive pessimism in "risky" situations. *Cognitive Therapy and Research, 10*(3), 347-362.

Norton, L. (2004). Using assessment criteria as learning criteria: A case study in psychology. *Assessment & Evaluation in Higher Education, 29*(6), 687-702. doi:10.1080/0260293042000227236

Nowell, L. S., Norris, J. M., White, D. E., & Moules, N. J. (2017). Thematic Analysis: Striving to Meet the Trustworthiness Criteria. *International Journal of Qualitative Methods*. https://doi.org/10.1177/1609406917733847

O'Connor, K. (2012). *Fifteen fixes for broken grades: A repair kit*. Toronto, ON: Pearson Canada.

Östberg, V., Plenty, S., Låftman, S., Modin, B., & Lindfors, P. (2018). School Demands and Coping Resources−Associations with Multiple Measures of Stress in Mid-Adolescent Girls and Boys. *International Journal of Environmental Research and Public Health, 15*(10), 2143.

Owens, A., & Newbegin, I. (2000). Academic Procrastination of Adolescents in English and Mathematics: Gender and Personality Variations. *Journal of Social Behavior and Personality, 15*(5), 111-124.

Page, S., & Alexitch, L.R. (2003). Learning- and grade-orientation, sex, and prediction of

self-reported academic performance. *Psychological Reports, 92*(1), 320.

Page, C.E. & Coutellier, L. (2018). Adolescent Stress Disrupts the Maturation of Anxiety-related Behaviors and Alters the Developmental Trajectory of the Prefrontal Cortex in a Sex- and Age-specific Manner. *Neuroscience, 390*, 265-277.

Parks, A.G, & Zurhellen, H.S. (1978). Student Attitudes toward the Grade Contract. *Improving College and University Teaching, 26*(4), 239-242.

Passolunghi, M., Ferreira, T., & Tomasetto, C. (2014). Math-gender stereotypes and math-related beliefs in childhood and early adolescence. *Learning and Individual Differences, 34*, 70.

Patterson, S. (2016). The effect of emotional freedom technique on stress and anxiety in nursing students: A pilot study. *Nurse Education Today, 40*, 104-110.

Peacock, E.J, & Wong, P.T. (1990). The stress appraisal measure (SAM): A multidimensional approach to cognitive appraisal. *Stress Medicine, 6*(3), 227-236.

Phan, H. (2009). Relations between goals, self-efficacy, critical thinking and deep processing strategies: A path analysis. *Educational Psychology, 29*(7), 777-799.

Phye, G.D, Schutz, P., & Pekrun, R. (2007). *Emotion in education* (Educational psychology series). Burlington u.a: Elsevier, Academic Pr.

Poe, M. & Inoue, A.B. (2016). Toward Writing as Social Justice: An Idea Whose Time Has Come. *College English, 79*(2), 119-126.

Polczynski, J., & Shirland, L. (1977). Expectancy Theory and Contract Grading Combined as an Effective Motivational Force for College Students. *Journal of Educational Research, 70*(5), 238-241.

Poppen, W.A, & Thompson, C.L. (1971). The Effect of Grade Contracts on Student Performance. *The Journal of Educational Research (Washington, D.C.), 64*(9), 420-423.

Postfoundationalism. (2017). *Encyclopedia of Science and Religion. Encyclopedia.com.* <http://www.encyclopedia.com>.

Potts, G. (2010). A simple alternative to grading. *Inquiry: The Journal of the Virginia Proceedings of the National Academy of Sciences of the United States of America, 109 Suppl 2*, 17180-5.

Purser, R. (2019). *McMindfulness: How mindfulness became the new capitalist spirituality*. London: Repeater Books.

Putwain, D. (2007). Researching academic stress and anxiety in students: Some methodological considerations. *British Educational Research Journal, 33*(2), 207-219.

Putwain, D., & Remedios, R. (2014). Fear appeals used prior to a high-stakes examination: What makes them threatening? *Learning and Individual Differences, 36*, 145.

Putwain, D. W., & Roberts, C. M. (2012). Fear and efficacy appeals in the classroom: The secondary teachers' perspective. *Educational Psychology, 32*(3), 355-372. doi:10.1080/01443410.2012.659845

Putwain, D. W., Nicholson, L. J., Nakhla, G., Reece, M., Porter, B., & Liversidge, A. (2016). Fear appeals prior to a high-stakes examination can have a positive or negative impact on engagement depending on how the message is appraised. *Contemporary Educational Psychology, 44-45*, 21-31. doi:10.1016/j.cedpsych.2015.12.001

University of California Drops SAT and ACT. (2020). *The Wall Street Journal. Eastern Edition*, pp. The Wall Street journal. Eastern edition, 2020-05-29.

Rabin, Laura A, Fogel, Joshua, & Nutter-Upham, Katherine E. (2011). Academic procrastination in college students: The role of self-reported executive function. *Journal of Clinical and Experimental Neuropsychology, 33*(3), 344-357.

Rachal, J. R. (1984). Community College and University Instructor Consistency in the Evaluation of Freshman English Themes. *Community Junior College Research*

Romano, J. (1992). Psychoeducational Interventions for Stress Management and Well‐Being. *Journal of Counseling & Development, 71*(2), 199-202.

Reardon, S., Fahle, E., Kalogrides, D., Podolsky, A., & Zárate, R. (2019). Gender Achievement Gaps in U.S. School Districts. *American Educational Research Journal, 56*(6), 2474-2508.

Sackett, P.R, Kuncel, N.R, Beatty, A.S, Rigdon, J. L, Shen, W., & Kiger, T.B. (2012). The Role of Socioeconomic Status in SAT-Grade Relationships and in College Admissions Decisions. *Psychological Science, 23*(9), 1000-1007.

Sarigiannidis, Ioannis, Grillon, Christian, Ernst, Monique, Roiser, Jonathan P, & Robinson,

Oliver J. (2020). Anxiety makes time pass quicker while fear has no effect. *Cognition, 197*, 104116.

Sanders, A., & Lushington, K. (2002). Effect of perceived stress on student performance in dental school. *Journal of Dental Education, 66*(1), 75-81.

Sarkar, S., Menon, V., & Kumar, S. (2020). Reducing stress among medical students: A qualitative study of students' perspectives. *Indian Journal of Psychiatry, 62*(2), 198-201.

Sarigiannidis, I,, Grillon, C., Ernst, M., Roiser, J.P, & Robinson, O.J. (2020). Anxiety makes time pass quicker while fear has no effect. *Cognition, 197*, 104116.

Schinske, J., & Tanner, K. (2014). Teaching More by Grading Less (or Differently). *CBE Life Sciences Education, 13*(2), 159-166.

Schneider, J., & Hutt, E. (2014). Making the grade: A history of the A-F marking scheme. *Journal of Curriculum Studies, 46*(2), 201-224.

Schippers, M., Morisano, D., Locke, E., Scheepers, A., Latham, G., & De Jong, E. (2020). Writing about personal goals and plans regardless of goal type boosts academic performance. *Contemporary Educational Psychology, 60*, Contemporary Educational Psychology, January 2020, Vol.60.

Schraml, K. (2010). Stress symptoms among adolescents: The role of subjective psychosocial conditions, lifestyle, and self-esteem. *Journal of Adolescence., 34*(5), 987-96.

Scott, D. & Usher, R. (2010). *Researching Education: Data Methods and Theory in Educational Inquiry, 2nd edition.* London: Continuum.

Selye, H. (1974). *Stress without distress*. London: Hodder and Stoughton.

Seijts, G.H., Latham, G.P., Tasa, K., & Latham, B.W. (2004). Goal Setting and Goal Orientation: An Integration of Two Different Yet Related Literatures. *The Academy of Management Journal, 47*(2), 227-239.

Schutz, P., & Pekrun, R.. (2007). *Emotion in education* (Educational Psychology). Amsterdam; Boston: Academic Press.

Shankar, N.L, & Park, C.L. (2016). Effects of stress on students' physical and mental health and academic success. *International Journal of School & Educational Psychology, 4*(1), 5-9.

Shor, I. (2009). Critical Pedagogy Is Too Big to Fail. *Journal of Basic Writing, 28*(2), 6-27.

Shim, S., & Ryan, A. (2005). Changes in Self-Efficacy, Challenge Avoidance, and Intrinsic Value in Response to Grades: The Role of Achievement Goals. *Journal of Experimental Education, 73*(4), 333-349.

Smith, B., & Lerch, H. (1972). Contract Grading. *Physical Educator, 29*(2), 80.

Spidell, C. & Thelin, W.H.. (2006). Not Ready To Let Go: A Study of Resistance to Grading Contracts. *Composition Studies, 34*(1), 35-68.

Spivey, C., Havrda, D., Stallworth, S., Renfro, C., & Chisholm-Burns, M. (2020).

Longitudinal examination of perceived stress and academic performance of first-year student pharmacists. *Currents in Pharmacy Teaching & Learning, 12*(9), 1116-1122.

Simon, S., Kirschenbaum, H., & Napier, R. (1970). The Day the Consultant Looked at Our Grading System. *The Phi Delta Kappan, 51*(9), 476-480.

Solomon, L. J., & Rothblum, E. D. (1984). Academic Procrastination: Frequency and Cognitive-Behavioral Correlates. Journal of Counseling Psychology, 31, 503-509.

Starch, D. (1913). Reliability and Distribution of Grades. *Science (New York, N.Y.), 38*(983), 630-636.

Starch, D., & Elliott, E. (1912). Reliability of the Grading of High-School Work in English. *The School Review, 20*(7), 442-457.

Stelzner, S. L. (1975). A case for contract grading. *The Speech Teacher, 24*(2), 127-132.

Svinicki, M. D. (2005). Student Goal Orientation, Motivation, and Learning. *Accounting Education News, 33*(3), 7.

Symonds, P. M. (1924). On the Loss of Reliability in Ratings Due to Coarseness of the Scale. *Journal of Experimental Psychology, 7*(6), 456-461.

Taylor, H. (1980). Contract Grading. *ERIC Clearinghouse on Tests, Measurement, and Evaluation*. *75(1),* 5-75.

Taylor, H. (1971). Student reaction to the grade contract. *The Journal of Educational Research, 64*(7), 311-314. doi:10.1080/00220671.1971.1088417

Taylor, S. E., Buunk, B. P., & Aspinwall, L. G. (1990). Social comparison, stress, and coping. *Personality and Social Psychology Bulletin*, 16(1), 74–89.

Terp, U., Bisholt, B., & Hjärthag, F. (2019). Not Just Tools to Handle It: A Qualitative Study of Nursing Students' Experiences From Participating in a Cognitive Behavioral Stress Management Intervention. *Health Education & Behavior, 46*(6), 922-929.

Thompson, T. (1999). Underachieving to protect self-worth: Theory, research and interventions. Ashgate, UK: Avebury.

Thompson, T., & Dinnel, D. (2007). Is Self-worth Protection Best Regarded as Intentional Self-handicapping Behaviour or as an Outcome of Choking Under Pressure? *Educational Psychology (Dorchester-on-Thames), 27*(4), 1.

Thompson, T., & Parker, C. (2007). Diagnosing the Poor Performance of Self-worth Protective Students: A product of future outcome uncertainty, evaluative threat, or both? *Educational Psychology (Dorchester-on-Thames), 27*(1), 111-134.

Thompson, T., & Richardson, A. (2001). Self-handicapping status, claimed self-handicaps and reduced practice effort following success and failure feedback. *British Journal of Educational Psychology, 71*(1), 151-170.

Thomas, J., & Rohwer, W. (1986). Academic Studying: The Role of Learning Strategies. *Educational Psychologist, 21*(1-2), 19-41.

Tice, D. M., & Baumeister, R. F. (1997). Longitudinal study of procrastination, performance, stress, and health: The costs and benefits of dawdling. *Psychological Science, 8*(6), 454–458. https://doi.org/10.1111/j.1467-9280.1997.tb00460.x

Travis, J., Kaszycki, A., Geden, M., & Bunde, J. (2020). Some Stress Is Good Stress: The Challenge-Hindrance Framework, Academic Self-Efficacy, and Academic Outcomes. *Journal of Educational Psychology, 112*(8), 1632-1643.

Twenge, J. M. (2017). Have Smartphones Destroyed a Generation? Retrieved July 19, 2019, from https://www.theatlantic.com/magazine/archive/2017/09/Has-the-smartphone-destroyed-a-generation/534198/

Twenge, J. M., Joiner, T. E., Rogers, M. L., & Martin, G. N. (2017). Increases in Depressive Symptoms, Suicide-Related Outcomes, and Suicide Rates Among U.S. Adolescents After 2010 and Links to Increased New Media Screen Time. *Clinical Psychological Science, 6*(1), 3-17. doi:10.1177/2167702617723376

Ward, E. (2021). An integrated mixed-methods study of contract grading's impact on adolescents' perceptions of stress in high school English: A pilot study. *Assessing Writing, 48*, 100508.

Warner, J. (2016). I Have Seen the Glories of the Grading Contract...: Inside Higher Ed. Retrieved April 22, 2019, from https://www.insidehighered.com/blogs/just-visiting/i-have-seen-glories-grading-contract

Warner, J. (2017). Grading Contract Success, At Last, At Last: Inside Higher Ed. Retrieved April 22, 2019, from https://www.insidehighered.com/blogs/just-visiting/grading-contract-success-last-last

Wolfe, J. L., & Fodor. (1996). The Poverty of Privilege: *Therapy with Women of the*

"Upper" Classes, Women & Therapy*, 18*(1), 3-4, 73-89, DOI: 10.1300/J015v18n03_08

Wolvin, A.D, & Wolvin, D.R. (1975). Contract Grading in a Technical Speech Communication Course. *The Speech Teacher, 24*(2), 139.

Wormeli, R. (2006). Accountability: Teaching Through Assessment and Feedback, not Grading. *American Secondary Education, 34*(3), 14-27.

Wilks, S.E. (2008). Resilience amid Academic Stress: The Moderating Impact of Social Support among Social Work Students. *Advances in Social Work, 9*(2), 106-125.

Williams, R. L., & Anandam, K. (1973). The effect of behavior contracting on grades. *The Journal of Educational Research, 66*(5), 230–236. https://doi.org/10.1080/00220671.1973.10884461

Wilson, M. (2006). *Rethinking Rubrics in Writing Assessment.* Portsmouth, NH: Heinemann Educational Books.

Wilson, M. (2007). Why I Won't Be Using Rubrics to Respond to Students' Writing. *The English Journal, 96*(4), 62-66.

Yarber, W. (1974). Retention of Knowledge: Grade Contract Method Compared to the Traditional Grading Method. *The Journal of Experimental Education, 43*(1), 92-96.

Yin, R. K. (2006) "Mixed Methods Research: Are the Methods Genuinely Integrated or Merely Parallel," *Research in the Schools*, vol. 13, no.1, pp. 41-47.

Zarzeski, M.T. (1998). The use and benefit of flexible student contracts. *Issues in Accounting Education, 13*(3), 585-594.

Appendices

Appendix 1: Instrument for Students with Prior Experience

Workload Demands
● The size of the workload was excessive.
● I believe that the amount of work the assignment required was too much.
● I was unable to catch up if I got behind on my work.
● The research paper was unusually difficult.
● The research paper's workload was very stressful for me.
Self-Perceptions
● I was confident that I would be successful during the research paper unit
● I feared failing the research paper.
● I am confident I will write successful research papers in the future.
Time Restraints
● The deadlines stressed me out.
● The timeframe was too short to complete the project.
● I had enough time to relax after working on the paper.
Importance of Ability
● I found it necessary to do well on the research paper in order to preserve a sense of self-worth.
● I strove to ensure success on my research paper through persistence and hard work.
● I closely followed the assignment requirements in order to avoid being marked down.
● I strove to ensure success on the research paper by a thorough approach.
Ability Doubts
● I compared my ability to those around me.
● I was afraid that I would fail the research paper even though I often do pretty well.
● I lacked confidence in my ability to do well.
● I found it difficult to do my best on the research paper due to a persistent fear of failure.
● I was afraid of putting too much effort into the research paper in case I failed.
● I worried about not succeeding on the research paper, even though others around me had considerable confidence that I would do well.

● If I received praise or recognition for my research paper, I tended to discount the importance of my own achievement.
● When I faced new challenges, I doubted my ability to do well.
● I'm not particularly threatened by the research paper, even if it is likely to reveal low ability.
● I did not doubt my ability to do well on the research paper.
● I do not rate my ability to write a good research paper very highly.
Avoidance Orientation
● I avoided working on the research paper because I couldn't do as well as I would like.
● I felt like my performance was being judged, so it was hard to try my best.
● I tended to play it safe with my research paper and choose goals that were within my reach.
● I withdrew from challenges with the research paper due to fear of failure.
● I avoided challenges with the research paper that could end in failure.
● I underachieved relative to my level of ability, choosing easy goals in order to ensure success on the research paper.
● Setting goals for the paper that were within my ability allowed me to preserve a sense of self-worth.
● I was able to try my hardest because failing the research paper would not reveal low ability.
● I performed at my best because there was little risk of failure.
● It was easy to try hard because my performance wasn't being judged.
● I tried to avoid situations during the research paper in which failure would reveal low effort.
PASA: Threat Appraisal
● I do not feel threatened by the research paper.
● I find the research paper very unpleasant.
● The research paper scares me.
● I do not feel worried because the research paper does not represent any threat for me.
PASA: Challenge Appraisal
● For the research paper, I know what I can do to succeed.
● It mainly depends on me whether my teacher evaluates my work positively.
● The research paper challenges me.

● I can think of lots of solutions to help me succeed on the research paper.
● If my teacher evaluates my research paper positively, it will be a consequence of my effort and personal commitment.
● I can best protect myself against failure through my behavior.

Appendix 2: Instrument for First-Time Students

● The research paper has important consequences for me.
● The research paper has long-term consequences.
● I will be greatly affected by the outcome of the research paper.
● The research paper has serious implications for me.
● I am eager to tackle the research paper.
● The research paper is going to have a positive impact on me.
● I will become a better student because of the research paper.
● I am excited about the outcome of my research paper.
● I have someone I can turn to if I need help with the research paper.
● There are enough resources available to help me deal with the research paper.
● There is help available to me while working on the research paper.
● There is someone who can help me manage the research paper.
● I will be able to overcome the research paper.
● I have the ability to do well on the research paper.
● I have the necessary skills needed to achieve a successful outcome on the research paper.
● I perceive the research paper as stressful.
● The research paper will tax or exceed my coping resources.
● The research paper creates tension in me.
● I am threatened by the research paper.
● The research paper is going to have a negative impact on me.
● The outcome of the research paper will be negative.
● The research paper makes me feel anxious.
● The outcome of the research paper is uncontrollable.
● This is a totally hopeless project.
● It is beyond anyone's power to do anything about the research paper.

Appendix 3: Contract Questions for All Interviews

1. Can you tell me about your experience with and perception of the grading contract?
2. How did your teacher introduce the grading contract?
3. Did anything about your teacher's introduction of the grading contract increase your personal level of stress this year? Decrease?
4. To you, what is the difference between the old rubric and the new contract?
5. To you, what was the difference between the two contracts?
6. Which one did you choose, and why?
7. What did it mean to you that you were contracting for a grade?
8. Which of the contract's requirements did you find difficult to understand or fulfill? Easy to understand or fulfill?
9. Do you believe you had it within your ability to fulfill the contract you chose?
10. Would you say that the grading contract increased your nervousness or worry about the research paper? Decreased?

Appendix 4: Interview Guide for Participants with Prior Experience

Fear of Failure

1. In the past, how worried were you about not succeeding on the research paper?
2. This year, would you say that there was any less worry or fear than in the past?
3. In the past, do you recall feeling afraid of getting marked down?
4. This year, were you less afraid of getting marked down, and why?
5. Last year, did you feel like your teacher was judging your performance?
6. This year, did you feel like your teacher was judging your performance?
7. Compared to last year, do you feel like you had an easier time doing your best this year?

Importance of Ability

1. How important is it to you or your family that you do well on the research paper?
2. In general, how likely are you to compare your ability and/or performance to peers around you?
3. While writing the paper this year, how likely were you to compare your ability to those around you, and why?

Avoidance Orientation

1. Would you feel better if you worked hard at something and got a low grade or if you didn't work hard and you got a low grade?
2. Are you likely to avoid working on something if you think you might get a low grade?

    a. Follow-up: In the past, can you tell me about a time you avoided working on the research paper? This year?

3. Have you ever pretended that you had not worked hard when you really had?
    a. Follow-up: Last year, do you recall pretending you had not worked hard on the paper when you really had? This year?
4. In the past, have you played it safe with your research paper goals and effort?
    a. Follow-up: Do you think you were less likely to play it safe this year?

Perception of Workload Demands

1. What was your perception of the workload and its size this year compared to last year?
2. What about the time frame to complete the project this year compared to last year?
3. When it comes to school work, what is the difference between a challenge and a threat?
4. Did the demands of the workload feel more challenging or more threatening?
5. Compared to last year, did the deadlines cause you less stress this year?
6. Compared to last year, would you say you had more or less time to relax after working on the paper?

Perceptions of Stress

1. Do you think that a certain level of stress ever helps you perform well in school?
2. Compared to last year, did you feel any helpful stress this year during the research paper?

3. Can you tell me about anything this year that increased your personal stress level?
4. Can you tell me about anything this year that decreased your personal stress level?
5. What impact, if any, do you think that the grading contract had on increasing or decreasing your personal stress?
6. Do you believe you fulfilled your contract, and why or why not?
7. After this year, would you say that the research paper is any less unpleasant or scary?
8. After this year, are less likely to be nervous or worried about future research papers?

Ability Doubts

1. Based on your experience this year, how confident are you that you can write a successful paper in the future?
2. Do you think anything about the grading contract increased your confidence in your ability to do well?
3. Finally, what solutions or resources do you think helped you with the research paper?

Appendix 5: Interview Guide for First-Time Students

Primary Appraisal

1. Can you tell me about any prior experience with research papers?
2. What had you heard about the research paper?
3. Did you believe the research paper unit would be a scary experience?

Secondary Appraisal

1. Tell me more about your actual experience with the research paper this year.
2. While working on the paper, what was easy or went smoothly for you?
3. While working on the paper, what was difficult or challenging for you?
4. Can you tell me about anything this year that caused you to feel nervous, worried or stressed this year?
5. Can you tell me about anything this year that decreased your feelings of nervousness, worry, or stress this year?
6. What was your perception of the workload and its size this year?
7. How much stress did the workload cause you?
8. Did you feel like you had enough time frame to complete the project this year?

Challenge/Threat

1. Do you think that a certain level of stress ever helps you perform well in school?
2. Did you feel any helpful stress during the research paper this year?
3. At school, what is the difference between something being challenging and something being threatening for you?

4. Did the demands of the workload this year feel more challenging or more threatening?
5. After your experience this year, would you say that the research paper is less scary than you thought it would be?

Lacked Resources & Solutions

1. Did you ever feel unsure of what to do to succeed?
2. Did you feel like you lacked people to turn to for help with the research paper?
3. Did you feel like you lacked sufficient resources to help you succeed?
4. What resources did you need to help you with the research paper?

Control

1. Did your own fear of not doing as well as you wanted ever stop you from working on the paper this year?
2. This year, who did you feel like was in control of your success, and why?
3. What about the research paper felt within your control? Uncontrollable? Controlled by others?

Learning

1. Based on your experience this year, how confident are you that you can write a successful paper in the future, and why?
2. Finally, what solutions or resources do you think helped you with the research paper?

Appendix 6: Sample Grading Contracts

HFE2/H-Eng. 2/Eng. 2 Contract for a B

The following contract is intended to detail the major elements of a sophomore-level research paper.
High (88%) **Medium** (85%) **Low** (82%)

You are guaranteed a B if you:

1. Format the paper and Works Cited in **MLA format** with no major errors.
2. Write a **thesis statement** that includes a **supportable title** for the decade and corresponding **topic sentences**.
3. Find and **use six sources** (three per topic) from the library databases. *Use no more than eight sources.*
4. Write a minimum of 60 paraphrased **notes** that **support the thesis** and then use those only to write your paper. Cite each note with an **MLA parenthetical citation**.
5. Write a minimum of **eight academic paragraphs** (introduction, six body paragraphs, and conclusion).
6. Cite at least **two different sources** in each body paragraph and balance the use of each source throughout the paper.
7. **Analysis** is consistently present and reasonably explained in each paragraph.
8. Meet the **minimum word** count of 1,500 without exceeding 1,800 words (-1% per 1% over or under).
9. Writing must be edited and pruned for **proper grammar and mechanics**.
10. Submit printed essay and submit to turnitin.com **on time** (on PSL) and receive a **clean report.**

Total: _____/100

Mandatory Rewrite Threshold — *rewrite for up to 70 percent*

- Well-below word count (25% or more of the paper is missing)
- Minimum sources used are half (or fewer) of the requirement
- No controlling idea and no attempt made at answering the prompt

HFE2/H-Eng. 2/Eng. 2 Contract for an A

The following contract is intended to detail the major elements of a sophomore-level research paper.

High (98%) **Medium** (95%) **Low** (92%)

You may earn an A if you:

1. Format the paper and Works Cited in **MLA format** with no errors.
2. Write a **well-crafted thesis statement** that includes a **strong, supportable title** for the decade and corresponding **well-crafted topic sentences**.
3. Find and **use six sources** (three per topic) from the library databases. *Use no more than eight sources.*
4. Write a minimum of 60 paraphrased, **interest-grabbing notes** that **prove the thesis** and then use those only to write your paper. Cite each note with an **MLA parenthetical citation**.
5. Write a minimum of **eight academic paragraphs** (introduction, six body paragraphs, and conclusion).
6. Cite at least **two different sources** in each body paragraph and balance the use of each source throughout the paper. All sources should work together **seamlessly.**
7. **Insightful analysis** is consistently present within each paragraph.
8. Meet the **minimum word** count of 1,500 without exceeding 1,800 words (-1% per 1% over or under).
9. Writing must be edited and pruned for **sophisticated use of grammar, mechanics, diction, and sentence structure**.
10. Submit printed essay and submit to turnitin.com **on time** (on PSL) and receive a **clean report**.

Total: _____/100