An Integrated Mixed-Methods Study of Contract Grading's Impact on Adolescents' Perceptions of Stress at a Private High School

Emily Ward

Lancaster University, UK

Abstract

This study analyzed the impact of contract grading on adolescents' perceptions of stress amid Good Shepherd High School's annual research paper unit. While college instructors have employed contract grading since the 1970s, the alternative assessment approach appears underused and under-analyzed in contemporary high school classrooms. In spring 2019, participants (n=53) enrolled in one of seven senior-level English courses with identical prompts and teaching materials. While three maintained a traditional grading rubric, four sections were evaluated with mastery-based grading contracts for A or B. The qualitative and quantitative datasets revealed that the Contract Grading Group was significantly more likely to perceive the workload and time constraints as less demanding. Additionally, despite a history of low grades, the majority (84%) fulfilled the contract's requirements, and the Contract Grading Group earned six times as many As and 2.5 times as many Bs as those in the Traditional Grading Group. Within this context, the grading contract reduced the stress of workload demands while significantly improving grades for students with prior experience with each requirement on the contract.

Keywords: academic stress; academic performance; contract grading; high school; adolescents; writing assessment

1. Introduction

Situated in high school English classrooms, this work is concerned with the impact of writing assessment on adolescent well-being. While Schinske & Tanner (2014) characterize now-standardized A-F grading practices as a young, ever-evolving system within the last 100 years of education in the United States, grades—that is, marks on assignments, scores on standardized tests, cumulative GPAs, report cards, and transcripts—are now integral to students' experiences in school. Low grades can increase feelings of defeat (Docksai, 2009) and decrease persistence (Bouffard-Bouchard, 1990). To reduce stress related to grades and shift focus from performance to learning, instructors have used contract grading as an alternative assessment approach across the disciplines since the 1970s.

By clarifying expectations and inviting students to participate in the assessment process, the contract grading contrasts with the traditional grading system. While the latter often increases stress, frustration, and writer's block, contract grading appears to be an accurate assessment approach (Potts, 2010) that reduces college students' stress and anxiety (Fairbanks, 1992), increases task-oriented motivation (Polczynski & Shirland, 1977), encourages responsibility (O'Hanlon & Bock, 1973), and improves academic performance (Hiller & Hietapelto, 2001; Lindemann & Harbke, 2011). To the best of this researcher's knowledge, however, no empirical work has studied the impact of contract grading on contemporary high school students.

To fill this gap, this pilot study occurred at Good Shepherd High School (GSHS) during the annual five-week research paper unit. Each January, the high-stakes assessment occurs in all English courses from freshman to senior year. As evidenced by decreased persistence and an excessive failure rate compared to other assignments, the unit breeds high levels of stress. To ameliorate well-being, the English Department planned to implement contract grading in all

courses during the research unit in 2020. In January 2019, one year before the program's full launch, I conducted a self-initiated, mixed-methods pilot study of contract grading with an identified high-need population: regular senior-level English courses with a history of low and failing grades. The following research question motivated this study: how does contract grading impact high school students' perceptions of academic stress, particularly those with a history of low or failing grades?

This work is significant since contract grading appears under-used and -studied in high school classrooms. Although the earliest known work on contract grading advocated for its use in high schools (McLaughlin, 1961), empirical work with contract grading has been relegated to the 1970s and '90s, and in high schools, work has focused on the philosophy of contract grading (Amsden, 1970; Barkley, 1975; Bowers & Howard, 1975; Kokus & Mussoff, 1975; James, 1977). While recent work has found increased student involvement and control (Litterio, 2016; Litterio, 2018) and increased academic performance (Potts, 2010; Lindemann & Harbke, 2011), no contemporary study has examined the impact of contract grading on adolescents. To fill this gap, this mixed-methods study included survey data and adolescents' voices to examine contract grading's impact on their perceptions of stress.

Thus, the next section explores the study's theoretical framework, stress theory, before examining the English Department's implementation of contract grading for the research unit. The study's methodology precedes the findings and discussion of contract grading's impact on adolescents' perceptions of academic stress under the grading contract.

2. Theoretical Framework: Stress Theory

2.1 Academic Stress in High School

Achievement pressure, particularly on high-stakes assessments, drives adolescents' academic stress, which occurs when a student perceives scholastic demands as exceeding their resources (Wilks, 2008), such as time, energy, or ability. Students at high-achieving high schools, like GSHS, spend more time on homework than the average teen and report chronic levels of sleep deprivation, which are associated with increased stress (Feld & Shusterman, 2015; Galloway, Conner, & Pope, 2013. Higher stress levels are associated with lower satisfaction with life and school (Chambel & Curral, 2005). Unfortunately, the Collaborative for Academic, Social, and Emotional Learning (CASEL, 2018) found that nearly one third (31%) of high school students, 40% of high school seniors, and roughly half (48%) of students with mostly C grades or lower on their last report card report feeling stressed *all* or *most of the time*.

While the APA's (2018) most recent "Stress in America" report found that Generation Z, those born between 1995 and 2010, is the most likely cohort to report poor mental health, including perceived stress and anxiety, they are also the most likely peer group to seek professional help. GSHS counselors (personal correspondence) found that students reported increased anxiety, depressive symptoms, and stress during the English Department's high-stakes research assessment

Annually, all GSHS students move step-by-step through the research process each January and February. To scaffold instruction, regular and honors-level courses dedicate all instructional minutes to teaching ethical academic research. With each grade level, the difficulty of the prompt increases alongside the word count and the required number of sources (see Figure 1). Unfortunately, the research unit became a source of collective suffering for the student body, according to the school's counselors: the unit's daily deadlines and time-intensive nature have

built mounting pressure as students struggle to manage their course loads. The looming threat of course remediation if they fail to submit a paper intensifies students' fear.

Academic stress, then, is intimately connected to fear of failure, which is pervasive in school children, particularly around high-stakes assessments; unfortunately, rather than work to reduce it, teachers often use it as a motivator, relying on its presence by frightening children into working (Jackson, 2010). Though used as a motivational tactic, fear appeals can elicit either a challenge or threat appraisal (Putwain et al., 2016). While the perception of the appeal varies according to the message's severity (Putwain & Roberts, 2012), a challenge appraisal occurs when the student feels capable of responding successfully and can lead to greater engagement as difficulties bring opportunities for mastery and personal growth; however, when failure is anticipated, a threat appraisal serves as a risk to self-worth and leads to lower engagement (Putwain et al., 2016).

Low-performing students appear to resign, putting forth less effort due to their past experiences. Based on department data, students earn failing grades on the paper within the research unit at rates disproportionate to any other assignment. For example, of the approximately 300 students in the freshman class in 2018, 26% earned a D or an F on the paper. In the regular senior-level courses evaluated in the study, 20% of students also earned a D or an F on the previous year's research paper.

2.3 Perception and Academic Stress

The brain is the primary organ for stress (McEwen, 2012), and fortunately, like the brain, perception is plastic. Known as the father of stress research, Selye (1936, 1974) defined stress as the body's non-specific response to a demand for change, whether negative ("distress") or positive ("eustress"). 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, however, led researchers to recognize the importance of one's perception, or appraisal, "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 study operationalizes the seminal work of psychologists Lazarus and Folkman (1984), whose transactional model of stress and coping understands stress as the interplay of two factors: a stimulus and an individual's interpretation of it. For example, 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), yet the first-time skydivers reported significantly higher psychological distress (Hare, Wetherell, & Smith, 2013). Noting that "stress, then, is not a variable but a rubric consisting of many variables and processes" (p. 12), Lazarus and Folkman (1984) called this the process of cognitive appraisal, which like threat or challenge appraisals, can be either positive or negative. While physiological reactivity and doctors' health reports may serve as objective measures of stress, it is ultimately an individual's perception that determines their interpretation of the experience as either positive or negative.

2.4 Contract Grading Review

To minimize academic stress and fear of failure, the GSHS English Department turned to contract grading, which uses the letters and numbers of traditional grading, thus fulfilling institutional requirements to assign grades, while mitigating the presence and pressure of grades "by ironically paying attention to how grades are constructed" (Inoue, 2019, p. 142). Most simply, contract grading outlines the performance criteria—that is, the specific actions and behaviors that

are required to earn each grade (in this study, A and B)—and then invites students to participate in their assessment by choosing the criteria that correspond with their desired goal and energy expenditure. By focusing students' attention on skill acquisition, the contract system values students' labor, which is essential to learning (Inoue, 2019), thereby reducing the threat of failure (Smith & Lerch, 1973) and improving task-oriented motivation (Polczynski & Shirland, 1977).

College students have perceived the system as a fairer assessment approach than traditional grading practices (Taylor, 1971; Hassencahl, 1979), reporting fewer negative academic emotions, like stress and anxiety (Parks & Zurhellen, 1978), and more task-oriented motivation (Polczynski & Shirland, 1977), as previously mentioned. When Fairbanks (1992) used the contract to decrease math anxiety, he found that the contract reduced stress and enabled more students to pass. Callahan (1979) observed that students enjoyed the "relaxed atmosphere afforded by the use of the contract grading method" (p. 7), but they often contracted for and then earned lower grades than students in the traditional grading course, a phenomenon Elbow (1997) also observed. The clear expectations and contract's guarantee (i.e., "you are guaranteed a B if—") may strengthen work-life balance and discourage academic anxiety inspired by grade uncertainty. One consequence may be less A grades and striving for excellence but more passing grades and less academic stress. Thus, I hypothesized that participants would earn fewer Cs and As but more Bs under the contract given their past grades on the project.

More recently, college students under the contract earned higher grades compared to their peers in traditional 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). While both Lindemann and Harbke (2011) and Litterio (2018) concluded their recent work by encouraging teachers to implement contract grading to bolster academic

performance and student responsibility, contract grading has remained unexamined in secondary classrooms where students face considerable threats to their well-being from achievement pressure.

2.5 The Treatment

To strike at the cause of student suffering, rather than its symptoms, two changes occurred to impact the adolescents' psycho-emotional well-being: all study participants received a reduced workload while the treatment group received the grading contract. First, prior to the study, the English Department reduced the workload demands while maintaining the same instructional minutes (see Figure 2) for all courses, including the treatment and control groups. In other words, all students in this study completed a shorter paper than the previous year, a change I hypothesized could alone ameliorate well-being concerns.

The teachers in the study worked together to ensure all participants had identical calendars, course materials, and deadlines, as well as lessons on how to select a research topic, research using academic databases, cite sources correctly using MLA format, and write paraphrased notes to avoid plagiarism. Each teacher checked their students' progress and provided feedback on topic selection, source selection, notes, and drafts. The only instructional difference was the assessment tool. The control group received the analytic grading rubric (see Figure 3), in which each criterion was assessed separately, two weeks before the final paper was due to give them sufficient time to review it and ask questions. The rubric was familiar, as all participants had completed the assessment the previous year.

The treatment group received the grading contracts (see Figure 4) on the first day of the unit when the teacher emphasized the English Department's commitment to students' psycho-emotional well-being and the desire to reduce stress during the unit. The treatment group then

participated in the assessment process by contracting for either an A or B (See Figure 4), which entailed signing their desired contract and then reporting their choice to the teacher. No student was deterred from contracting for an A, although Warner noted that many students who had not earned high grades on previous writing assignments contracted for an A.

Additionally, during each lesson, Warner emphasized the corresponding contract item and encouraged students to self-assess, using it as a tool for learning and assessment. Students received feedback if they fell short of their goal and a stamp when they reached it. Warner observed that this nuance fostered positive affect and self-confidence. In this way, the treatment group participated in the learning and assessment process, while the teacher held the final decision over the grade, a common practice among contract grading practitioners (Danielewicz & Elbow, 2009; Potts, 2010; Litterio, 2018).

Finally, GSHS requires teachers to update their grade books weekly. To accommodate this, all teachers in the study 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, then, was worth 19% of their overall course grade.

2.6 Creating the Contracts

To develop the grading contracts, the English Department adopted Danielewicz & Elbow's (2009) hybrid model, creating two contracts for each course—one for mastery (A) and another proficiency (B)—using a three-step process we developed: identify the need(s) for change; identify the learning goals for the assignment; and turn the learning goals into concrete, specific items for the contract. Identifying the need for change was crucial to unifying the team and helping them work through the challenges that would follow. Next, articulating the learning goals, what students

should know and be able to do, was the lengthiest step: 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. The list of learning goals then turned into concrete, explicit, and objective actions that lead to proficiency. As a result, all items on the B contract were "concrete, do-able, yes-no tasks" (Elbow, 1997, p. 12), while the A contract included the subjective, "fuzzy" qualities of exceptional writing (See Figure 4). Unlike Inoue's (2019) labor-based model, the A and B contracts asked for the same quantity of work.

Additionally, the English Department only outlined the criteria for A and B, not C or lower, as Johnston & O'Neill (1973) found that the teacher's criteria controlled their performance to a high degree, regardless of the participants' prior performance. They 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. The goal was that all students, even those who had previously earned low or failing grades on the project, would at least achieve proficiency. The grading process will be explained in the next section.

3 Methodology

This study utilized a mixed-methods approach with a convergent design to explore the impact of contract grading on students' perceptions of stress. The instrument generated both closed-ended quantitative and open-ended qualitative data to analyze participants' subjective and objective well-being. 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).

3.1 Institutional Setting

This work is situated at Good Shepherd High School (anonymized), a private, religious college-preparatory institution for grades 9-12 (ages 13-19). With approximately 1,350 students, the comprehensive, co-educational school resides in an affluent, suburban county in the Western United States. The student body is predominantly upper-middle-class and European American (72%). After graduation, the majority of students (96%) attend a two- or four-year college or university.

3.2 The Instrument

A survey entitled "Perceptions of Academic Stress" (see Appendix A) gathered 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 & Gabriel's (2015) Perceptions of Academic Stress (PAS) Scale, which is based on the transactional model of stress (Lazarus & Folkman, 1984) and uses a five-point Likert scale with one as strongly disagree to five as strongly agree. When applicable, "research paper" replaced a generic term for schoolwork in the survey statements. The survey also included free-response questions, which asked students to write about their perceptions of stress, what they liked, and what they did not like about this year's unit. No question asked about or referenced contract grading.

The mobile-friendly survey was distributed to students' emails after the students submitted the final paper and before they knew their final grade.

3.3 Participants

Participants were high school seniors (n=50) and juniors (n=3) enrolled in a regular, senior-level English course. Females (n=32) outnumbered males (n=21) in the study but not in the

classrooms. Participants were randomly assigned to one of seven courses completing the research paper unit with identical assignments. To minimize the influence of teacher variance, care was taken to ensure the three teachers worked closely together, maintaining identical prompts, teaching materials, and expectations for the final paper, which will be covered in the next section.

The three control group courses, taught by Timothy Johnson (one section) and Karly Von Kleist (two sections), were under the supervision of lead teacher Bethany Warner, who volunteered four sections for the pilot contract grading. Following school-wide trends, 20% of Warner's students also earned a D or F on the paper last year, as previously mentioned, and more than half (57%) of her seniors who previously earned a D or F submitted the paper late, which prompted her willingness to pilot the new assessment. With these factors considered, teacher variance can account for some but not all of the results of the study.

Randomization resulted in seven courses with statistically comparable demographics, including age, gender, GPA, final course grade in English last semester, and final grade on last year's research paper. The four courses in the treatment group (n=33) were evaluated with the contract grading system while the three control groups (n=20) retained the traditional grading system.

Of the 119 students who signed consent forms, 53 students completed the survey, making the study participation rate 44%. Of the 53 participants, no one was removed as all met the study criteria of completing the research paper unit the previous year. While 71% (n=38) of participants had a GPA average of 3.0-3.9, 53% (n=28) of participants earned a C or lower on last year's research paper.

3.4 Assessing with the Contract

To grade with the contract, Warner used a check-system for each contract item (see Table 1). The average of the marks provided a holistic assessment of the paper as 'high,' 'medium,' 'low,' or 'off the contract,' which were printed on each student's contract. Each level was associated with one number, thus reducing the total number of grades from 20 (i.e., 80-100) to six (see Table 2).

Table 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. Warner used a simple check-system to assess each paper holistically with the grading contract.

Table 2
Percentage Grades Based on the 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.'

Warner also had a not-for-student-use list of qualities of C, D, and F papers. In the mastery-based approach, all students were asked to strive for proficiency (that is, A or B), as previously explained; however, 'off the contract' papers that met an adequate standard of achievement could earn the passing grade of C (i.e., 72, 75, 78).

Teachers in the study met as a course team to establish a benchmark by first grading together and then grading their students' papers individually. All papers received a percentage score (0-100%) worth 20% of the students' final grade. With their students' consent, teachers supplied final grades for analysis.

3.5 Data Analysis

In line with the convergent mixed-methods approach, each data set was analyzed separately before merging the data to compare the results (Creswell & Guetterman, 2018). Qualitative survey responses were analyzed and hand-coded using a six-step thematic coding process outlined by Nowell, Norris, White, and Moules (2017): familiarizing oneself with the data; generating initial codes; searching for themes; reviewing themes; defining and naming themes; and producing the report. The initial codes were "stress level," "workload," and "contract grading"; however, based on its frequency in the responses, "duration of the project" was added. The qualitative data was also transposed into quantitative data with the purpose of comparing both datasets. No theme was considered final until all data had been read through and coded (King, 2004; cited in Nowell et al., 2017).

Quantitative data was analyzed by conducting t-tests that compared the results of three stress-related outcomes (i.e., ability doubts, workload demands, and time constraints) between the treatment and control groups using version 24 of IBM's SPSS Statistics. The alpha level of .05 was used for all statistical tests, and to correct for family-wise error as a result of multiple testing, Benjamini-Hochberg adjusted *p*-values are reported. The significant quantitative results were also critically examined beside the qualitative data's themes.

3.6 Ethical issues

This study was conducted with full institutional consent. To further mitigate ethical issues, parents were notified via email of the research plans via an information sheet that allowed them six weeks to opt their child out of the study. Participants also received information sheets and consent forms. Care was taken to ensure students did not feel pressured to participate, either formally or informally. Participation was voluntary and confidential to provide each student assurance that their teacher's attitude toward them would not change based on their willingness to take part in the study. Furthermore, all names, including the name of the institution, have been anonymized.

Additionally, I followed the Ethical Guides for Education Research (2011) laid out by the British Educational Research Association (BERA) and received Lancaster University's ethical approval procedure, receiving formal ethical clearance from the module convenors.

Furthermore, I recognize my positionality as an insider within the institution's English Department and thus the potential for my personal bias as a teacher at the school. Throughout the data collection process, I pursued a posture of reflectivity, "a continuing mode of self-analysis" (Callaway, 1992, p. 33), to be vigilant toward my own practices, claims, and assumptions, about which I have sought to be transparent in this study. Additionally, in an effort to negate the influence of biases during the research process, I engaged with colleagues outside of the English Department and heeded feedback from my module convenors and reviewers.

4 Findings

4.1 Academic Stress

By self-reported objective measures, the quantitative data revealed no statistically significant difference between the hours sleeping, doing homework, or relaxing between the

Contract Grading Group and Traditional Grading Group (see Figure 5). On average, participants in both groups spent somewhat *less* time sleeping and relaxing and somewhat *more* time completing homework during the research paper process, yet the difference between an average week and the research paper was not statistically significant. From an average weeknight of M=2.85 (SD=.82), sleep slightly decreased to M=2.42 (SD=.95) during the unit. Time spent relaxing slightly decreased, too, from M=2.36 (SD=.71) on an average weeknight to M=1.85 (SD=.89) during the research paper process. As sleep and relaxation slightly decreased, the number of hours spent completing homework increased by about a half an hour from M=2.25 (SD=.78) on an average weeknight to M=2.7 (SD=.95) during the research paper process.

While the two groups reported comparable objective measures of well-being, the analysis of the qualitative and quantitative data revealed two themes about how students in the Contract Grading Group perceived stress differently during the unit:

Theme 1: Students in the Contract Grading Group viewed the process as significantly less demanding than the previous year.

In the qualitative data, 40% of participants (n=22) referenced the amount of time they had to write the paper and meet the deadlines. From the day it was assigned to the due date, participants in the study had 11 class periods (825 instructional minutes); in the previous year, as juniors, they had 15 (1125 instructional minutes), which many noted. In the qualitative data, 45% (n=9) of the Traditional Grading Group and 39% (n=13) of the Contract Grading Group mentioned the length of time they had to complete the project, yet a key difference emerged in how it impacted them.

For the Contract Grading Group, the project's time-frame was a source of dislike; for the students in the Traditional Grading Group, it was a source of stress. The quantitative analysis revealed that the Contract Grading Group was significantly more likely to agree that the process was "less demanding than last year" (M=4.12, SD=1.08) than the Traditional Grading Group

(M=3.2, SD=1.6), t(51)=2.495, p= 0.04 (see Figure 6). The qualitative data also revealed that 40% (n=8) of the Traditional Grading Group explicitly stated or described that they experienced "more stress" this year than last year, compared to 9% (n=3) of the Contract Grading Group (see Figure 7). For example, one Traditional Grading Group participant wrote that she did "not have enough time to write" while another described the difficulty of balancing schoolwork and then found "little to no time [for] social interaction or relaxation or even sleep" (Johnson). The following statements show the 40% (n=8) of the Traditional Grading Group that experienced heightened stress from workload demands and time constraints:

> This year was more stressful than last year. We had one night to complete a third of our paper three nights in a row for the first rough draft. (Johnson)
>
> Last year, we spent a month and a half on the research paper, but this year, we spent less than a month on the research paper, therefore making the workload more than last year since things went by quicker. (Johnson)
>
> This year, I was more stressed because I felt like I had a lot less time to finish. There was not as much time to get it done. (Von Kleist).

Conversely, while a few in the Contract Grading Group described more stress this year (n=2, 6%) or equal stress both years others (n=2, 6%), the majority (88%, n=29) of students in the Contract Grading Group reported less stress from workload demands under the grading contract. While one student noted that "this year was way better" due to "a sense of safety" that made her less stressed and "more motivated to complete the paper," other students explained what impacted their stress:

> Personally, I was less stressed writing it this year and did not have a lot of questions because my teacher was clear on what to do and gave us work time in class. (Warner).
>
> Last year, I felt as though I got more confused in what I was supposed to be doing and got behind more. This year, I noticed that the standards for getting either an A or B really helped because it clearly showed me what I needed to do to get that grade. (Warner)
>
> Last year, I got behind so this year was a lot better being stress-free. (Warner)
>
> Last year's project was 100% harder than this year. This year was a breeze. (Warner)

Task avoidance, procrastion, and falling behind can stem from confusion about expectations. Under the contract, however, clarity of expectation and reduction of confusion and doubts seemed to maximize class-time to work on the assessment, making the project significantly less demanding.

The Contract Grading Group was also 50% more likely to speak positively about stress and deadlines, even perceiving them as a beneficial challenge, as revealed through these responses:

> The strict deadlines [made the process more stressful] but it was also a positive thing. (Warner)
>
> I wasn't as stressed as I was this year, but a certain amount of stress is good to keep you on track. (Warner)
>
> I liked the shorter turnaround because I feel very relaxed with it done so early. (Warner)
>
> Last year, the whole process took what felt like months as we had 2-3 days for each part, which lightened the workload, but extended the project too long. I wasn't as stressed as I was this year, but a certain amount of stress is good to keep you on track. Compared to last year, finishing it in 2-3 weeks made it more stressful, but the degree of stress was acceptable. (Warner)

Theme 2: The contract's clarity reduced students' academic stress.

No survey question asked students about contract grading, yet when asked what made this year's research paper more or less stressful and what they liked and did not like about it, the Contract Grading Group overwhelmingly (81%, n=27) cited the contract as a critical component in reducing academic stress. The responses below show a sample of students' positive perceptions of the contract, although some called it a "rubric":

> I knew exactly what the teacher was looking for and knew the requirements to receive the grade I wanted. The grading rubric was very helpful and made it a lot less stressful.
>
> There was more freedom and the rubric was easy to understand and follow.
>
> I really liked the contract, actually. It showed me exactly what my paper needed and what I had already completed.

> I loved the grading rubric. Knowing what exactly was expected to get the letter grade I wanted was very helpful.
>
> I liked how the rubric gave me the structure to my paper and I knew what I needed to have on my paper to achieve the grade I wanted.
>
> This year, you chose how much work and effort you wanted to put in and it was on you to get things done.

Furthermore, the quantitative data revealed that the Contract Grading Group was significantly more likely to disagree that the "teacher's expectations were unclear" (M=1.48, SD=.795) than the Traditional Grading Group (M=2.2, SD=1.005),t(51)=-2.87, p=0.03 (see Figure 8). The contract's "easy to understand" clarity told students "exactly what to do." To understand how this impacted their performance, the final grade of all participants (n=119) who signed the consent form and released their grades were analyzed (see Figure 9). The Contract Grading Group earned significantly higher grades on the final paper (M=86.23, SD=7.88) than those in the Traditional Grading Group (M=78.09, SD=7.84), t(117)=5.35, p=0.00.

5 Discussion

5.1 Perceptions of Academic Stress

The qualitative and quantitative analyses revealed that participants in the Contract Grading Group perceived the process as less stressful and significantly less demanding than the control group, despite the fact that both groups completed less work than the previous year as a result of the English Department's workload reduction to combat stress. Notably, the majority of the Contract Grading Group (81%) brought up the contract's clarity of expectation as key components that lowered their level of stress by providing a clear path to success While many participants had prior experiences that resulted in low or failing grades, the grading contract appeared to shift

achievement motivation, orienting adolescents toward success and "alleviat[ing] the threat and stigma of failure" (Smith & Lerch, 1973, p. 82).

Additionally, the objective measures of wellbeing revealed no statistical significance between the treatment or control group, yet the qualitative data helped explain why the subjective perception of stress varied significantly between the group. First, the Traditional Grading Group's responses show their focus on their perceived powerlessness toward factors outside of their control: timeframe and deadlines. Whether the project felt too drawn out or too quick, increased stress appraisal accompanied the perception of low control. The contract restored a sense of control that helped students take ownership of their work and grade, which corroborates the aforementioned research of Polczynski & Shirland (1977) and Parks & Zurhellen (1978). The choice of contracts, clear path to desirable outcomes, and subsequent sense of agency triggered a positive response that helped students mitigate stress and look back on the unit with increased satisfaction.

Two active ingredients, both of which are often absent in traditional assessment, appeared to shift students' appraisal of workload demands and ultimately their perception of academic stress: control and choice. Providing the contracts at the start of the unit not only gave students a clear sense of the expectations but choice over how much energy to exert to reach a desirable goal, either an A or B. This choice appeared to be psychologically motivating for the adolescents in the Contract Grading Group. Through the enumeration of the project's expectations combined with the contract's unusual grade guarantee upon completion of the project, students were oriented toward success, which led them to reappraise the workload demands, particularly the timeframe and deadlines, as less stressful and demanding. This work points to the contract grading's ability

to promote eustress and transform stress triggers into a source of motivation, thereby improving students' experience in the writing classroom.

5.2 Academic Achievement

To add further insight into perceptions of stress and to understand if the survey participants (n=53) differed from those who signed the consent form (n=119), all students' grades were analyzed (See Figure 9). There was no significant difference between the grades of those who completed the survey and those who did not. Both assessment approaches were impactful at increasing academic performance outcomes, as measured by the number of Ds and Fs on the project. No student in either group earned a failing grade on the project, and the number of Ds decreased in both groups. In the Traditional Grading Group, 20% earned a D or an F last year, which lowered to 4.3% earning a D this year. In the Contract Grading Group, 24% earned a D or an F on the paper last year, which lowered to 1.4% earning a D this year. Furthermore, many participants who earned low grades on the paper last year failed to complete and finish it on time, yet all participants (n=119) in both groups submitted the paper on time this year.

As measured by the number of grades higher than C, the Contract Grading Group reached higher performance outcomes. In the Contract Grading Group, 84% of participants (n=61) met the contract requirements for either an A or B, while 16% (n=12) earned a C or lower. In the Traditional Grading Group, 60% of participants (n=28) earned a C or lower (see Figure 11). Based on Callahan's aforementioned study (1979), I hypothesized that students in the Contract Grading Group would have lower perceptions of stress by earning more Bs yet less Cs and As. The Contract Grading Group, however, not only earned 2.5 times as many Bs but six times as many As as their peers in the control group. The contract outlined the performance criteria, yet the grade of A ultimately rested on the teacher's subjective evaluation of exceptionally high-level work. This

work corroborates the finding of Johnston & O'Neill (1973). When given clear criteria for a high level of achievement accompanied by a valued outcome, the Contract Grade Group fulfilled the requirements and then earned higher grades, despite past performance. As Hiller & Hietapelto (2001) observed in their courses with contract grading, "Most of our students have risen admirably to the learning challenges they contracted to undertake" (p. 674). While some fell below the contract's requirements, the Traditional Grading Group earned two times as many Cs on the final paper. This finding of increased academic performance also corroborates Lindemann & Harbke's (2011) research of college students who were three times more likely to earn an A and one-third as likely to fail or withdraw than their peers in a traditional grading course. By offering choice and a sense of control, the contract not only significantly reduced perceptions of stress and workload demands but led students to significantly higher achievement.

6 Study Limitations and Further Research

The pilot study had several limitations, the first of which is a potential Pygmalion effect, or other-imposed prophecy, as the teacher of the treatment group emphasized her commitment to her students' well-being, which may have impacted, consciously or unconsciously, their perceptions and behavior. The survey, however, was careful not to mention the grading contract, which makes it significant that the majority of the treatment group cited the contract, by themselves, as a key element that reduced stress and improved their learning experience. 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 psycho-emotional health among

adolescents. The impact of the grading contract may have been augmented by the school's culture, which may limit the generalizability of this research.

Additionally, the potential for self-report bias in objective and subjective measures of academic stress and socio-emotional well-being is a design limitation, as students may have offered socially desirable answers. To encourage authentic responses, students were informed twice, first on the consent form and then on the survey, that their responses were confidential and would be anonymized by the researcher. The study also attempted to minimize bias by including both objective and subjective measures. Aside from the questions from the Perceptions of Academic Stress Scale, six retrospective questions attempted to gather objective measures of well-being related to relaxing, sleeping, and doing homework. The phrasing of those six questions may have encouraged students to presume differences that did not exist; however, no retrospective self-report finding was statistically significant in this study.

Additionally, while examining statistically similar populations of students, the use of three teachers is a design limitation that may account for some of the variance between the treatment and control groups. As previously mentioned, care was taken to minimize the influence of variance in instruction or expectations by ensuring all teachers worked together and had comparable instruction, teaching materials, and expectations for the final paper. Compared to her students' grades from last year, the teacher of the treatment group, Warner, saw a significant increase in academic performance, as measured by the final grade, under the grading contract.

Additionally, a data limitation is the 44% participation rate. This may have been due to the fact that students received a printed consent form to complete in class, while the online survey required out-of-class time. In an effort to mitigate the influence of my positionality, I took great care to ensure that no student was pressured, either formally or informally, to participate, and no

incentives were offered. As a result, the participation rate for the survey was lower than desired. Notably, however, the quantitative data analysis revealed no statistically significant difference between the grades of those who signed consent forms and those who completed the survey.

While no survey question asked about race or socioeconomic status, one impact limitation of this study is that the school's study body is upper-middle class and predominantly white (67%). While participants were identified as a high-need group, their socio-economic status and private school education have given them access to privileged discourses that may limit the generalizability of the results for classrooms with a wider spectrum of abilities and what Shor (2009) calls "the cultural clash of working-student identity with middle-class academic culture" (p. 11).

Further research in this area should also explore how contract grading impacts students with little to no prior experience with the research paper units and the contract items, such as high school freshmen. Additionally, since many of the students in the study were medium-to-low level academic achievers, future research should also explore how contract grading impacts the academic stress and performance of high-achieving high school students who consistently earn grades higher than B when the outcome is not guaranteed. Finally, future work should also explore how contract grading impacts adolescents' self-worth protection behaviors, achievement motivation, and affective responses to academic writing.

7 Conclusion

This study examined the impact of contract grading on high school seniors' academic stress during a five-week research paper unit. Although students in the study were identified as a high-

need population due to a history of low or failing performance during the research unit, students with the contract perceived the project as significantly less demanding of their time and resources compared to their peers in the control group. While reappraising the deadlines and workload as less demanding, 84% of students with the grading contract earned either a B or an A on the paper, compared to only 40% of students in the control group. Notably, even before learning of their high grade, the Contract Grading Group pointed to the contract's clear performance standards as instrumental in reducing their perceptions of stress related to the workload. Providing students control over their work and choices of energy expenditure can be psychologically motivating, helping students reappraise potential stressors as motivation.

Ultimately, assessment shapes teaching and learning as well as students' perceptions of workload demands and even their own capabilities. By inviting students into the learning and grading process and allowing them to choose their desired level of effort, the results of this study reveal that contract grading can reduce students' perceptions of stress and appraise workload demands as facilitative. With a choice of paths, each with a desirable outcome, students' perceptions of academic stress decreased as more students rose to a high standard of achievement. The results of this study led to the full launch of contract grading in all courses in the GSHS English Department for next year's research paper unit and should inspire confidence for using contract grading in contemporary secondary classrooms.

8 **Funding**

The author received no funding for this work.

References

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

American Psychological Association (APA) (2018). "Stress in America: Generation Z." Stress in America™ Survey. Retrieved from https://www.apa.org/news/press/releases/stress/2018/stress-gen-z.pdf

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

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.

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

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 Okely, J., & Callaway, H., (eds.,) *Anthropology and Autobiography* (29–49). New York: Routledge.

CASEL (2018). Respected: Perspectives of Youth on High School & Social Emotional Learning. Retrieved 14 January 2019, from https://casel.org/wp-content/uploads/2018/11/Respected.pdf

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.

Creswell, J., & Guetterman, Timothy 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.

Docksai, R. (2009). Opening Up the Shut-Down Learner: More visuals and less text is the key to helping many underperforming students learn better, says a school psychologist. *The Futurist, 43*(5), 54.

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

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

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

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.

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

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. (2012). *Grading contracts: Assessing their effectiveness on different racial*

formations. In A. B. Inoue & M. Poe, Race and writing assessment (pp. 79-94) New York: Peter Lang.

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 https://wac.colostate.edu/books/perspectives/labor/

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

James, S. (1977). The Evolution of a Language Arts Program for Pre- and Early Adolescent Students. *English Journal, 66*(4), 47-51.

Johnston, J., & 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. 112-3.

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

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.

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. M. (2018). Contract grading in the technical writing classroom: Blending community-based assessment and self-assessment. *Assessing Writing, 38*, 1-9.

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

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

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*.

O'Hanlon, J., & Bock, R. (1973). Teaching Responsibility through Student Commitment. *Theory Into Practice, 12*(4), 242-245.

Parks, A.G., & Zurhellen, H.S. (1978). Student Attitudes toward the Grade Contract. *Improving College and University Teaching, 26*(4), 239-42.

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

Potts, G. (2010). A simple alternative to grading. *Inquiry: The Journal of the Virginia Community Colleges*, *15*(1), 29–42.

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

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

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

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

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

Yin, R. K. (2006) "Mixed Methods Research: Are the Methods Genuinely Integrated or Merely Parallel," *Research in the Schools*, *13(1)*, 41-47.

Appendix A

Part 1: Objective Measures of Well-Being

- Sleep
  - On an average weeknight, how many hours do you sleep?
  - During the Research Paper process, how many hours did you sleep?
- Homework
  - On an average weeknight, how many hours do you spend completing homework?
  - During the Research Paper process, how many hours did you spend completing homework on an average weeknight?
- Relaxation
  - On an average weeknight, how many hours do you spend relaxing (e.g., watching Netflix, listening to music, or going outside)?
  - During the Research Paper process, how many hours did you spend relaxing (e.g., watching Netflix, listening to music, or going outside) on an average weeknight?

Part 2: Perceptions of Academic Stress (Randomized)

Ability Doubts

- My final grade is important to me.
- Throughout the process, I was confident I would succeed.
- Based on the instruction, I felt capable of meeting all of the requirements.
- I am disappointed with my final work.
- Throughout the Research Paper process, I feared I would fail.

Workload Demands

- For me, the process was mostly stress-free.
- The size of the workload was excessive.
- My teacher's expectations were unclear.
- The process was less demanding than last year.
- The expectations for the project are too demanding.

Time Constraints

- I was unable to catch up when I got behind.
- I did not have enough time to complete all the work.
- On average, I had enough time to relax after completing my work.
- It was difficult to stay motivated.

Part 3: Free-response Questions about Academic Stress

- Compare last year's Research Paper process and this year's research paper process. Describe the difference in your personal stress level.
- What, if anything, made the Research Paper process more stressful this year?
- What, if anything, made the Research Paper process less stressful this year?
- What, if anything, did you like about this year's Research Paper process?
- What, if anything, did you not like about this year's Research Paper process?