

Running head: EASING STRESS: CONTRACT GRADING'S IMPACT

Easing Stress: Contract Grading's Impact on Adolescents' Perceptions of Workload Demand,
Time Constraints, and Challenge Appraisal in High School English

Emily Ward

Lancaster University, UK

Abstract

Mastery-based contract grading is a holistic assessment approach for learning and grading in which students choose their desired effort and outcome by contracting for either an A or B to meet high academic standards. This mixed-methods study examined the impact of mastery-based contract grading on secondary students' (grades 9-12) perceptions of stress and threat appraisal. Participants were 439 adolescents, including 284 returning students and 155 first-year students, completing a high-stakes writing assessment in their required English course. Using 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. The findings revealed that the contract significantly reduced evaluative threat by clarifying expectations and bolstering confidence. 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. The findings of this study make a significant contribution to the field of writing assessment, leading to a call to action for teachers to implement mastery-based contract grading in high school classrooms to create psycho-emotionally healthy learning environments that reduce perceptions of stress and increase challenge appraisal.

Keywords: evaluation stress; stress appraisal; contract grading; high school; adolescents; writing assessment

1. Problem Statement

This research is concerned with a pervasive social force that has bred stress, anxiety, and fear of failure in even the youngest learners: the culture of high-stakes assessment. In January 2020, to mitigate evaluation pressure and seek to emphasize learning over performance, the Good Shepherd High School (GSHS) English Department, where I have taught since 2013, implemented an alternative assessment approach called mastery-based contract grading, in which all adolescents were offered the paths to proficiency (i.e., B grade) or mastery (i.e., A grade). The unorthodox yet democratic contract grading system is uncommon in schools and, in my experience, unfamiliar to most secondary teachers. As evidenced by recent research (Lindemann & Harbke, 2011; Litterio, 2016; Litterio, 2018), Inoue's (2019) recent book, and articles in *Inside Higher Ed* (Warner, 2016, 2017), contract grading may be used more in 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).

Most simply, the contract system contrasts with traditional grading practices by outlining the performance criteria—that is, the specific actions and behaviors that are required to earn each grade (typically, A, B, and C)—and then invites students to participate in their assessment by choosing the criteria that correspond with their desired goal, effort, and energy expenditure. While the traditional 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 and thus ameliorate the source of student stress. Recent work points to its efficacy in contemporary college classrooms (Potts, 2010; Lindemann & Harbke, 2011; Litterio, 2016; Litterio, 2018), yet most empirical work with contract grading has been relegated to the 1970s, limited in size and scope, and focused on college students, with success: under the contract, contemporary college students 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).

The research presented here builds successful pilot study with a high-need group of 12th graders with a history of low grades on a high-stakes research paper (Ward, 2021). Compared to the control group, 12th graders under the contract showed a statistically significant decrease in their perception of stress from workload demands while also earning significantly higher grades. In January 2020, all grade (9-12) and course levels (i.e., regular, AP/honors, and those accommodated for learning disabilities) adopted the grading contract for the five-week assessment, in which all instructional minutes are dedicated to teaching academic research, that is worth 20% of their final grade. The findings of this study will be presented over two articles: first, **the purpose of this article is to examine the impact of contract grading on adolescents' perceptions of academic stress and evaluative threat. A subsequent article will present the findings on academic performance and self-worth protection behaviors under the grading contract.**

The next section outlines the theoretical framework, stress theory, before outlining the study's mixed methods methodology. The findings, presented thematically, precede the discussion, limitations, suggested future work, and conclusion.

2. Theoretical Frameworks

2.1 Stress Theory

In the transactional model of stress and coping, a psychological theory of stress put forth by Lazarus and Folkman (1984), stress is understood as the interplay of a stimulus and the appraisal of the potential stressor. During primary appraisal, the individual examines the potential threat of stimulus, such as high-stakes assessment. During secondary appraisal, the individual examines their available resources to handle the potential threat. 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).

The brain, then, is the primary organ for stress (McEwen, 2012). Appraisal of a stressor determines reality, a point illustrated in studies with first-time and experienced skydivers where 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). Perception, then, determined whether the experience was positive or negative, challenging or threatening. This helps to understand how increased perceptions of stress lowered academic performance for some (Ng, Koh, & Chia, 2003; Spivey, Havrda, Stallworth, Renfro, & Chisholm-Burns, 2020), while the perception of high stress did not impact the performance of those training to be medical doctors (Sanders & Lushington, 2002), a self-selecting program that admits high-ability and self-selecting students who may be better prepared to cope with stress for a variety of reasons. Notably, adolescents are 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). The adolescents in this study are diverse in ability, course type, and grade level.

Additionally, this work explores the impact of examination stress on adolescents. Initially, I sought to examine academic stress, which occurs when a student perceives scholastic demands as taxing or exceeding their resources (Wilks, 2008), such as time, energy, or abilities; however, the survey statements, which I will discuss in the methodology section, 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 this study were asked to take notes, which as a classroom activity is generally low-stress, yet taking notes with the understanding that the notes are for a major assessment, worth 20% of their final grade, can produce examination stress.

3. Contract Grading

Contract grading, a social agreement between the teacher and the class about how grades are constructed (Inoue, 2019), rests on the principle that most students are capable enough to achieve proficiency. The goal is not equality of outcome but the highest possible fulfillment of potential; in other words, “everyone should achieve the best that is possible for them” (Nicholls, 1979, p. 1071). In this study, the B contract served as the minimum threshold for performance, *in an effort 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). Unlike traditional 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, he observed that labor is often taken for granted in traditional grading practices, which is unfair to diverse students. Fortunately, the area most within students’ control—how much time they spend laboring—is most essential to developing their writing skills.*

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 and effort. Twelve GSHS English teachers utilized a hybrid approach (also used by Litterio, 2016, 2018), which was first advocated by Danielewicz & Elbow (2009). Grades up to B are guaranteed for completing learning tasks, while grades higher than B rest

the teacher's subjective estimate of exceptional writing. While Inoue's (2019) labor-based system resists the dominant White language standard, the contract system employed in this study follows Danielewicz and Elbow's (2009) hybrid model, which may not go far enough in offering a socially just way to produce grades but does limit the teacher's judgment for most grades and provides students with a meaningful choice of their desired goal and effort. Compared to traditional practices, the contract invites students to take a more active role in the learning and assessment process. During the first week of the unit, all adolescents were asked to contract for either an A or B (see Figure 1), and throughout the unit, teachers provided feedback aimed at revision for each contract item before the final due date. In this way, this study examined the role of limiting teacher's judgement and restoring students' agency on their perceptions of stress.

4. Methodology

The study followed an explanatory mixed-methods design, as defined by Creswell (2015): data generation occurred in distinct phases with parallel construction, measures, and instruments with the same cohort of 439 participants, who completed a pre-survey in January and post-survey in February. After analyzing the quantitative data, I conducted 40 semi-structured interviews, which lasted between 40-70 minutes, with students of all course types (e.g., accommodated, regular, and honors) and courses (e.g., English 1, Honors English 1, etc.) 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 perceptions of stress.

4.1 Survey Participants

Participants (n=439) were secondary students (grades 9-12, ages 13-19) at a private, religious, college-preparatory institution in an affluent, suburban county on the West Coast of the United States, where 96% attend college after graduation. All were 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.

Of all participants, 226 identified as female, 210 identified as male, one as non-binary, one as transgender male, and two as other. Additionally, unlike the pilot study, which focused on a high-need group of 12th graders with a history of low or failing grades, the adolescents in this study were more diverse in ability and experience: 160 were enrolled in honors or AP English, 241 were enrolled in regular English, and 49 students were diagnosed with learning disabilities and received academic accommodations. The ethnic sample was also more diverse than the predominantly White (72%) institution, with 50% (n=226) identifying as European American; 16.1% (n=72) as Asian or Pacific Islander; 15.9% (n=71) as mixed-race; 9.9% (n=44) as Hispanic or Latinx; 2.7% (n=12) as Middle Eastern; 2% (n=9) as African American; 1.3% (n=6) as Native American; and 1.3% (n=6) as other.

4.2 Ethical Consideration

This research was conducted with full institutional consent from the administration and formal ethical clearance from the Department of Educational Research's Research Ethics Officer. The name of the school and the names of all participants have been anonymized to protect participants' confidentiality and anonymity. To mitigate the impact of my positionality and the unbalanced power relationships as a teacher who is researching the adolescents at her school, all of my current students were excluded from the study, as well as students in my 2018-2019 courses while I was conducting the pilot study. Care was taken to ensure that no student was pressured to participate. To minimize pressure to provide socially desirable answers, identities were anonymous unless participants volunteered for an interview and supplied their email address.

4.3 Research Instruments

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 adolescents and their alignment to the transactional model of stress (Carpenter, 2016). First-year participants (n=155) completed an adapted version of the Stress Appraisal Measure (SAM) (Peacock & Wong, 1990). Students with prior experience (n=284) completed adapted versions of the Perceptions of Academic Stress Scale (PASS) (Bedewy & Gabriel, 2015); Self-Worth Protection Scale (SWPS) (Thompson & Dinnel, 2003); and Primary Appraisal Secondary Appraisal (PASA) (Gaab, 2009). When applicable, "research paper" replaced a generic term, and a seven-point Likert scale was employed to align responses. Each scale appears to be psychometrically sound after item-selection,

analysis, and validation procedures. This article, however, only presents the findings on the dimensions related to stress appraisal (see Table 1).

4.4 Interview Guides

Two interview guides, one for First Timers and another for Prior Experiencers, were developed based on quantitative findings. Both guides 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). The interview questions were organized around the survey dimensions survey, perceptions of and experiences with the grading contract, and their overall learning experience. In line with semi-structured interviews, questions were used flexibly (Merriam, 2009), although all dimensions were covered with all participants. Regular follow-up questions were asked regularly to yield rich descriptive data.

4.5 Interview Sampling

Stratified sampling was used as a systematic method to choose interview participants from each grade level, English class, and class type. Interviews took place until English course, course type, gender, and ethnicity were as balanced as possible (see Figure 3.2). Of all interview participants, 47.5% (n=19) identified as female, 47.5% (n=19) as male, one as transgender male, and one as non-binary; 45% (n=18) were enrolled in regular-level courses, 37.5% (n=15) were in honors or AP courses, and 17.5% (n=7) had a diagnosed learning need. Finally, sixty-percent (n=24) had prior experience with the

research unit, while 40% (n=16) were first-year students completing the paper for the first time.

Additionally, two interviewees shared that they were English Language Learners studying abroad.

4.6 Data Analysis

While the data generation methods have analytical integration (Yin, 2006), each has a distinct and preferred analytical technique. The quantitative data was analyzed using IBM SPSS Statistics version 25. Paired samples t-test compared the mean of individual scores on the pre- and post-survey, and to control for the familywise error rate, *p*-values were adjusted using the Benjamini-Hochberg procedure. Descriptive statistics tests were used to understand the population of the sample.

I transcribed all interviews in full using Trint and coded them in NVivo 12. First, I 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). In this way, I took a deductive stance to generate initial codes and subcodes, which then led to the generation of broad themes.

The next section presents the findings, which are organized around the research questions.

5. Findings

5.1 How does contract grading affect students' perceptions of academic stress?

5.1.1 Reduced Perception of Stress

Prior-Experiencers were significantly less likely to report that the workload was “very stressful” ($M=4.17$, $SD=1.622$) compared to the previous year ($M=4.82$, $SD=1.445$), $6.070(283)$, $p=.000$. They were also significantly less likely to say they “feared failing the research paper this year” ($M=4.4$, $SD=1.931$) compared to the previous year ($M=4.85$, $SD=1.952$), (279) , $p=.000$. Notably, the size of the workload increased for Prior-Experiencers (see Table 2), yet all five statements in the dimension of “workload stress” yielded statistically significant results (see Table 3). Prior-Experiencers were significantly less likely to find the workload demands “excessive” ($M=3.40$, $SD=1.516$) compared to the previous year ($M=4.62$, $SD=1.497$), $7.619(282)$, $8.247(281)$, $p=.000$; “too much” ($M=3.26$, $SD=1.349$) compared to the previous year ($M=4.07$, $SD=1.392$), $p=.000$; or “unusually difficult” ($M=3.30$, $SD=1.474$) compared to the previous year ($M=3.77$, $SD=1.486$), $4.276(284)$, $p=.000$. Under the contract, they were also significantly less likely to report that were “unable to catch up” if they got behind ($M=3.11$, $SD=1.726$) compared to the previous year ($M=3.37$, $SD=1.716$), $2.004(282)$, $p=.046$. For this dimension, the effect size ranged from small for statements about stress and difficulty to medium for perceptions of the workload demands, which will be explained in the limitations section.

The qualitative data then revealed how and why adolescents’ perceptions of stress shifted: the majority of adolescents interviewed ($n=30$, 75%) 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.

For adolescents in this study, “confusion” often 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.

5.1.2 Reduced Stress from Time Constraints

The quantitative data revealed that Prior-Experiencers were significantly less likely to experience stress from time constraints under the grading contract, although the effect size ranged from small to medium (see Table 4). More specifically, these adolescents reported significantly less stress from deadlines under the contract ($M=4.15$, $SD=1.874$) compared to the previous year ($M=4.97$, $SD=1.605$) $5.006(283)$, $p=.000$, and significantly less likely to perceive the timeframe as “too short” under the contract (from $M=3.63$, $SD=1.56$ to $M=3.31$, $SD=1.56$) $2.705(284)$, $p=.006$. Additionally, they were also significantly more likely to say they had “enough time to relax after working on the paper” under the contract (from $M=3.59$, $SD=1.805$ to $M=4.0$, $SD=1.789$) $-2.894(283)$, $p=.006$.

For students of all grade levels and course types, the qualitative data revealed that the presentation of the contract, particularly its brevity, led students to reappraise time constraints. For Jax (PE

accommodated 11th) “*the wording of the contract*” and “*the simplicity of the work*” reduced stress from time constraints. He said—

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:

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

The qualitative analysis revealed that the majority of students (80%, n=32), including 70% of First-Timers and 87% of Prior-Experiencers, perceived having enough time to meet workload demands and deadlines this year. While Prior-Experiencers had the same amount of time to complete a longer project both years, many (n=8, 33%) perceived having more time this year, as Mia (PE accommodated 11th) revealed:

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.

Two (8%) recognized that they completed more work this work 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.*" *In this way, the contract's organization and utility as a learning tool appeared to reduce the task's overall cognitive load, as defined by Clark, Nguyen, and Sweller (2006), which then maximized students' efficiency.* In fact, the workload not only *felt* like less, although it was a longer project; many (50%, n=12) reported doing less work this year, 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.

Droit-Volet and Meck (2007) revealed that emotions color one's perception of time, worries, doubts, and fear not only make the time *feel* longer (Lay & Schouwenburg, 1993), yet the contract mitigated these issues by providing clarity that enabled students to work confidently and 'do it right' the first time, as Evan (PE AP 11th) 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.

The size of the task, though larger this year, felt smaller as a result of the contract's presentation.

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): 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.

The perception of 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.

5.1.3 Increased Challenge Appraisal

Under the contract, Prior-Experiencers appraised the project as significantly more challenging and less threatening than their prior experience (see Table 5). They also 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 $M=3.86$, $SD=1.585$) $-3.599(274)$, $p=.000$; significantly less likely to perceive the paper as “very unpleasant” (from $M=5.15$, $SD=1.5$ to $M=4.74$, $SD=1.656$) $3.936(277)$, $p=.000$; 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 $M=5.16$, $SD=1.178$) $-2.415(273)$, $p=.036$. They were also significantly less likely to report that the research paper “scares” them (from $M=3.77$, $SD=1.456$ to $M=3.30$, $SD=1.574$)

3.237(274), $p=.001$. Finally, they were significantly more likely to report that they “know what [they] can do to succeed” (from $M=4.86$, $SD=1.297$ to $M=5.17$, $SD=1.223$), $-3.428(274)$ $p=.021$. The effect size, however, was small.

Adolescents 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/or 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) said, “A challenge pushes my potential and kind of drives me.”

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.

Overwhelmingly, students across all grade levels and course types ($n=39$, 98%) appraised the paper as challenging, not threatening, as Olivia (PE regular 10th) and Caleb (PE honors 11th) explained:

Olivia: 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: This year felt more challenging, and I didn't really feel this year that I was unable to achieve the grade of the 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 Prior-Experiencers (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.

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.

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

5.1.4 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, particularly sports teammates, had given them a "warning" about the assessment. In spite of what they had heard, however, the majority of First-Timers' (68%, n=11) appraisal positively shifted, 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 First-Timers were significantly more likely to report a challenge appraisal after their experience with the grading contract; however, the effect size remained small (see Table 6). 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 ($M=5.37$, $SD=1.247$), $t(154)=2.185$, $p=.05$.

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, the majority of First-Timers (88%, $n=14$) also credited their perception of "*an abundance*" of help and resources to assist them with the paper (Malcolm, FT honors sophomore). All First-Timers ($n=16$, 100%) felt supported, with sufficient resources to succeed. As Sienna (FT regular freshman) said, "*There was a lot*

of help everywhere,” which made the process easier, better, or not as “*bad*” as expected for many

First-Timers (65%, n=11).

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.

While the workload was reduced the previous year, Prior-Experiencers 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, many (69%, n=11) First-Timers 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.

Gabriel (FT regular junior): 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 Prior-Experiencers alike perceive the project as easier and more manageable.

5.1.5 Reduced Stress

Overall, the clarity of the grading contract alleviated many worries, doubts, and fears about assignment expectations that are often associated with traditional grading practices and thus significantly reduced academic stress. At all grade levels and course types, the majority of adolescents (n=30, 75%) interviewed cited the contract as central to reducing stress this year, alongside teacher support (n=19, 48%) and prior experience (n=10, 25%) at any grade level, including middle school (grades 6-8). Academic stress, 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. As students prepare for assessments, Thomas & Rowler (1986) observed that "clarity of purpose is pronouncedly atypical" as students "rarely know with any precision what they are preparing themselves to do" (p. 30). However, Ella, whose favorite subject is science, the contract "*adapted to [her] skills and then broke [the paper] down more for a different kind of learner,*" explaining—

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 prior experience or maturity may have altered her perception, she credited the contract as critical to reducing her stress and making the process “*easier*.” Like Ella, many adolescents I interviewed (n=23, 56%) 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*.” Aside from the contract, the pacing and structure of the course were comparable in both years, yet the contract made the process less demanding, which points again to contract’s impact on reducing extraneous cognitive load through simple, accessible directions. Additionally while GSHS has taught this research unit for decades, and thus no lesson was created from scratch, the grading contract was constructed following the principles of backward curriculum design; in this way, the creation of the contracts may have improved teaching by focusing and refining existing lessons, even without teachers’ conscious knowledge, thereby improving students’ experiences.

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.

Luck and mystery surround the traditional grading process, where 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 was more direct, revealing

under such conditions, students couldn't "*have as much confidence.*" To describe the grading contract's impact on their psycho-emotional well-being during the unit, many (35%, n=14) students I interviewed 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.

While traditional assessment often obscures the evaluation standards, the contract offered procedural fairness in the form of clear learning goals, which brought adolescents relief and reduced their perception of stress.

6. Discussion

Under the grading contract, adolescents reported a statistically significant reduction in stress and evaluative threat. 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). Students must regularly ascertain implicit expectations, a taxing task that can exacerbate stress levels and lower confidence, according to the adolescents in this study. 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). For adolescents in the study, the task’s newfound clarity of purpose increased their 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 less evaluative threat and more confidence.

Under the contract, adolescents perceived that they had more or enough time to complete the task, even when compared to previous years when they completed shorter papers. Distress depends upon one’s perception of the potential stressor (Lazarus & Folkman, 1984), and as Droit-Volet and Meck (2007) reported, “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) documented with several pioneering studies: “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. For students of all levels and course types, however, the contract positively changed students’ appraisal of the task by explicitly connecting the due dates, classwork, and expectations in the student’s minds, which reduced the task’s cognitive load, decreased evaluative threat, and increased adolescents’ sense of control and confidence.

7. Limitations

This study had several limitations, the first of which is a design limitation. The absence of a control group in the pre-post test design increases the potential for internal validity threats. Adolescents with prior experience 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 perception of task demands 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, historical events serve as another threat: 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. In the pilot study, however, there was no difference in the objective measures of well-being, yet the subjective measures were statistically significant. Perception 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. Connected to this is the limitation is self-report bias as students may have offered

socially desirable answers, yet as previously mentioned, responses were anonymous unless participants volunteered for an interview.

Additionally, a data limitation is the participation rate of 45% as a result of the study design, which required participants to complete both the pre- and post-surveys. 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.

An impact limitation is the small effect sizes of many of the significant quantitative results. This, I surmise, 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 writing process (Murray, 1972). Fortunately, the in-depth interviews helped to understand the significant results, including “what is” through the distribution of variables across multiple dimensions through the quantitative strand and “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 naturally lingering fear and uncertainty that may diminish over time, thereby increasing the effect of the grading contract. The qualitative dataset

revealed the impact of contract grading on adolescents' psycho-emotional well-being. A limit to the generaliability of this research, however, is the institution's proclaimed commitment to adolescent well-being: direct and palliative changes to address students' mental health may have strengthened the impact of the contract grading, the first pedagogical change to address evaluative stress from high-stakes assessment.

8. Future Work

Socially just assessment is fair to students as well as academic and administrative staff (McArthur, 2018), who are burdened by the time-consuming labor of grading. Future work should examine how grading impacts English teachers' perceptions of their 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.

9. Conclusion

The findings of this study can have rebounding effects on writing assessment in secondary schools. In this context, contract grading reduced evaluative threat for adolescents during high-stakes writing assessment, even when employed for a single unit, to create a psycho-emotionally healthy learning environment. 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. As Riley (PE 11th regular) said, the contract was "*just a piece of paper*" but it offered "*comfort*." While using the letters and numbers of the traditional grading, the contract grading system

stands in contrast to conventional practices by de-emphasizing the product (Inoue, 2019) and alleviating the looming shadow of the grade (Danielewicz & Elbow, 2009). Contract grading, if used more widely, can promote adolescents' sense of agency and control while improving psycho-emotional conditions.

References

- 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, 2*(2), 205510291559671-2055102915596714.
- Burstein, B, Agostino, H, 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.
- Carpenter, R. (2016). A Review of Instruments on Cognitive Appraisal of Stress. *Archives of Psychiatric Nursing, 30*(2), 271-279.
- Creswell, J.W. (2015). *A Concise Introduction to Mixed Methods Research*. Los Angeles: Sage.
- Danielewicz, J., & Elbow, P. (2009). A Unilateral Grading Contract to Improve Learning and Teaching. *College Composition and Communication, 61*(2), 244-268.
- 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.
- Fairbanks, P. 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 Adolescence*, 41(C), 31-42.
- 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.
- 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>
- 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.
- 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.
- 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.

- 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
- Nicholls, J. (1979). Quality and equality in intellectual development: The role of motivation in education. *American Psychologist, 34*(11), 1071-108
- Parks, A.G. & Zurhellen, H.S. (1978). Student Attitudes toward the Grade Contract. *Improving College and University Teaching, 26*(4), 239-242.
- Peacock, E.J. & Wong, P.T. (1990). The stress appraisal measure (SAM): A multidimensional approach to cognitive appraisal. *Stress Medicine, 6*(3), 227-236.
- 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.
- 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.
- 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

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.

<https://doi.org/10.1016/j.cognition.2019.104116>

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.

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

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.

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/>

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.