INVESTIGATING CORTISOL IN THE CLASSROOM: THE ASSOCIATION BETWEEN CORTISOL AND ACADEMIC PERFORMANCE

A Thesis in

Psychology

by

Hyun Joon Park

© 2018 Hyun Joon Park

Submitted in Partial Fulfillment of the Requirements for the Degree of

Master of Science

August 2018
The thesis of Hyun Joon Park will be reviewed and approved* by the following:

Jonathan Cook  
Assistant Professor of Psychology  
Thesis Adviser

Melvin Mark  
Department Head, Professor of Psychology

Martha Wadsworth  
Associate Professor of Psychology

*Signatures are on file in the Graduate School
ABSTRACT

OBJECTIVE: Despite numerous studies suggesting the association between cortisol and cognitive performance, there has been a lack of literature investigating the association between cortisol and academic performance. Further, studies that have examined this association have mainly focused on cortisol response in a very limited setting (e.g., on the day of an examination) and have not considered potential moderators. To address this gap in the literature, using a longitudinal field research design, I investigated students’ cortisol responses and their association with academic outcomes in a challenging science gateway college class. Additionally, I investigated whether students’ cortisol response and its association with performance are moderated by race/ethnicity.

METHOD: A total of 271 students who were enrolled in an Introduction to Molecular and Cellular Biology class participated in this study. Participants submitted two saliva samples per class across four classes (a total of eight possible samples). Participants’ exam scores and whether they dropped out of the course were collected from the instructor and official transcripts. RESULTS: Students’ cortisol levels were not elevated in classes around the exam period. Further, students’ cortisol patterns within class and across classes did not differ as a function of race/ethnicity. Averaged cortisol levels across classes did not have an association with academic outcomes. However, race/ethnicity moderated the association of cortisol with academic outcomes. Cortisol levels were positively associated with performance for negatively stereotyped students (i.e., Black, Hispanic, and Native American). There was no association between cortisol levels and performance for non-stereotyped students (i.e., White and Asian). CONCLUSIONS: Results suggest that cortisol may represent different psychological states and have different associations with academic performance for different racial/ethnic groups.

Keywords: Cortisol, Academic performance, Social identity threat
# Table of Contents

List of Tables .................................................................................................................. vi
List of Figures ................................................................................................................... vii
Acknowledgements ......................................................................................................... ix
CHAPTER 1: Introduction ................................................................................................. 1
CHAPTER 2: Literature Review .......................................................................................... 8
  HPA Axis and Cortisol .................................................................................................... 8
  When is Cortisol Secreted? ............................................................................................ 8
  How Cortisol is Studied in Past Psychology Literature .............................................. 10
  Impact of Elevation of Cortisol on Body and Cognitive Systems ............................. 13
  Cortisol and Cognitive Functioning ............................................................................ 14
  Cortisol and Academic Performance ........................................................................... 15
  Cortisol Representing Engagement ............................................................................. 18
  Interpreting Physiological Arousal ............................................................................. 20
  Malleable Relation between Physiological Arousal (Cortisol) and Performance .... 23
  Interpreting Cortisol in the Classroom: Does the Relation Between Cortisol and
  Academic Performance Differ by Race/Ethnicity? .................................................... 26
    Cortisol levels for negatively stereotyped students ................................................. 30
    Cortisol levels for non-negatively stereotyped students ......................................... 32
  Current Study ............................................................................................................... 34
CHAPTER 3: Method .......................................................................................................... 38
  Participants .................................................................................................................... 38
  Procedure ...................................................................................................................... 44
  Primary Measures ........................................................................................................ 48
    Cortisol ...................................................................................................................... 48
    Race/Ethnicity ........................................................................................................... 49
    Students’ final class score ....................................................................................... 50
    Students’ attrition from the course ........................................................................... 52
  Control Measures ........................................................................................................ 52
    Gender ....................................................................................................................... 52
    Socioeconomic status (SES) .................................................................................... 53
    Prior performance ...................................................................................................... 54
    Variables that may influence cortisol level ............................................................. 54
 CHAPTER 4: Results ........................................................................................................... 56
   Data Management ........................................................................................................ 56
     Data coding strategy ................................................................................................. 56
     Missing data strategy ............................................................................................... 56
     Covariates related to cortisol concentration ............................................................ 56
     Missing prior performance ...................................................................................... 59
   Preliminary Analysis ................................................................................................... 59
     Cortisol ..................................................................................................................... 60
     Academic performance ............................................................................................ 64
   Longitudinal Data Analysis: Cortisol Pattern Within and Between Classes .......... 64
     Cortisol pattern within class .................................................................................. 67
     Cortisol pattern between classes ........................................................................... 71
   Examining the Association Between Cortisol Level and Academic Performance ... 72
     Data analysis plan .................................................................................................... 72
     Cortisol and students’ final scores .......................................................................... 73
     Cortisol and attrition from the class ........................................................................ 77

 CHAPTER 5: Discussion .................................................................................................. 80
   Limitations .................................................................................................................. 87
   Future Directions ...................................................................................................... 90
   Conclusion .................................................................................................................. 90

 Appendix A: Self-Affirmation Intervention ................................................................. 92
   Self-Affirmation Intervention .................................................................................... 92
   Intervention Material (Control Condition) ............................................................... 93
   Intervention Material (Intervention Condition) ....................................................... 94
   Longitudinal Cortisol Pattern Across Classes Moderated by
     Intervention Condition ............................................................................................. 95

 Appendix B: Potentially Mislabeled Saliva Samples .................................................. 97

 Appendix C: Calculating Intra and Inter-Assay Coefficient ......................................... 99

 Appendix D: Predicting Averaged Adjusted Exam Scores ........................................... 100

 References .................................................................................................................... 101
List of Tables

Table 1. Number of Saliva Samples Submitted and Exams Taken ..............................................40
Table 2. Demographic Information of Participants .................................................................42
Table 3. How Students Perceived Class and Attending Medical School .................................43
Table 4. Descriptive Statistics of Unadjusted and Adjusted Exam Scores .................................51
Table 5. Means, Standard Deviations, and Correlations of Primary Study Variables ...............61
Table 6. Descriptive Statistics for Salivary Cortisol ................................................................62
Table 7. Partial Correlations of the Cortisol Level with Demographic Information Controlling for Factors that Influences Cortisol Level (Time Since Awakening, Whether Took Medication, Whether Brushed Teeth) .................................................................64
Table 8. Three-Level Models of Within-Class and Between-Class Cortisol Pattern in 271 Participants (1745 Cortisol Samples) ..........................................................................................70
Table 9. Regression Coefficients Predicting Final Scores Step 1-3 ............................................76
Table 10. Regression Coefficients Predicting Attrition Step 1-3 ................................................79
Table 11. Regression Coefficients Predicting Averaged Adjusted Exam Scores Step 1-3 . 100
List of Figures

Figure 1. Timeline of the study...........................................................................................................................................46

Figure 2. Cortisol pattern within the class. Values in grey at each time in class are predicted cortisol values (log-transformed) for each student averaged across classes and controlling for relevant covariates, with lines connecting values representing individual students’ average within-class slope. The solid line in blue represents the average within-class slope across students averaged across the classes (See Table 8). In a 75-minute class, saliva samples were collected at the beginning of the class (coded -0.5) and at the end of the class (coded 0.5). Class number was centered so that 0 was the theoretical average class (class number coded as following: Class 1 = -1.5, Class 2 = -0.5, Class 3 = 0.5, Class 4 = 1.5). ...........................................................................69

Figure 3. Cortisol pattern across classes. Values in grey at each class are predicted cortisol values (log-transformed) for each student averaged within class controlling for adequate covariates, with lines connecting values representing individual students’ average slope across classes. The solid line in blue represents the average slope across classes averaged within class and across students (See Table 8). Class 1 (Week3) represents an early period of the semester before any exams. Class 2 (Week 5) was the class right before the midterm exam. Class 3 (Week 6) was right after the midterm exam, but before students had received their grades. Class 4 (week 6 -7) was right after students had received their grades on the first midterm. Time within class was centered (start of class = -0.5, end of class = 0.5). The dotted line indicates when the first exam took place. .................................................................................................................................71

Figure 4. Students’ final scores as a function of cortisol levels. The regression line represents predicted scores for each observed value of cortisol controlling for appropriate covariates (all
continuous predictors are mean centered and categorical variables are contrast coded). Displayed points correspond to unadjusted observed data. Students who received over 86% received an A and those who received under 46% failed the class; dotted lines indicate cutoff for those grades.

Figure 5. Students’ final scores as a function of race/ethnicity and cortisol levels. The regression line represents predicted scores for each observed value of cortisol controlling for appropriate covariates (all continuous predictors are mean centered and categorical variables are contrast coded). Displayed points correspond to unadjusted observed data. Students who received over 86% got an A and below 46% failed from the class; dotted lines indicate cutline for those grades.

Figure 6. Students’ probability of dropping out of the class predicted by a function of race/ethnicity and cortisol levels. Values are predicted probability of dropping out from the class controlling for appropriate covariates (all continuous predictors are mean centered and categorical variables are contrast coded).

Figure 7. Cortisol pattern across classes by condition. Values in grey at each class are predicted cortisol values (log-transformed) for each student averaged within class controlling for adequate covariates, with lines connecting values representing individual students’ average slope across classes. The solid line in blue represents the average slope across classes averaged within class for students in the control condition. The dotted line in red represents the average slope across classes averaged within class for students in the intervention condition. Time within the class was centered (start of class = -0.5, end of class = 0.5). Dotted line indicates when the first exam took place.
Acknowledgements

I would like sincerely thank my advisor, Dr. Jonathan Cook. This work would not have existed without his thorough and dedicated guidance over the past two years. Additionally, I would like to thank Dr. Melvin Mark and Dr. Martha Wadsworth for being part of my thesis committee and providing me constructive feedback.

I would like to additionally thank my lab members. Especially, I thank Mikey Pasek, Julia Dahl, and Jackson Harper who were all part of this project and have helped me throughout various phases of this project.

I am very thankful for all my friends in Social area and State College. It was their strong emotional support that motivated me once again when I was uncertain about myself.

Lastly, with love, I would like to thank my parents who were always there for me.

This material is based upon work supported by the National Science Foundation under Grant No. DRL-1109548 awarded to Valerie Purdie-Vaughns, Ph.D., Jonathan Cook, Ph.D., Geoffrey Cohen, Ph.D. Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the National Science Foundation.
CHAPTER 1: Introduction

Cortisol is a hormone produced by the adrenal gland as an end product of the hypothalamic-pituitary-adrenocortical (HPA) axis (Sapolsky, Romero, & Munck, 2000). Cortisol plays a primary role in mobilizing energy in the body to prepare for changes in the environment (Lovallo & Thomas, 2000; Sapolsky et al., 2000). Typically, cortisol is studied as a reaction to negative physiological (Van Uum et al., 2008) and psychological stress (Dickerson & Kemeny, 2004); however, cortisol is also released in situations not necessarily related to negative stress, such as exercising (Jacks, Sowash, Anning, McGloughlin, & Andres, 2002), and engaging in activities one may find exciting (e.g., skydiving) (Chatterton, Vogelsong, Lu, & Hudgens, 1997). This suggests potential ambiguity in the meaning and consequences of cortisol. A primary goal of the research presented here is to probe this ambiguity by testing cortisol as a predictor of academic performance. As described in more detail below, some previous research suggests that elevated cortisol may be associated with worse academic performance, consistent with interpretations of cortisol activity as an indicator of negative stress, while other research suggests that elevated cortisol could be associated with improved performance, consistent with more nuanced interpretations of cortisol activity as an indicator of engagement. It is also possible that cortisol has no association with performance, as some research suggests, which would also be informative with respect to interpreting its role in psychological functioning.

One clue to the role of cortisol may be in its biological effects, which can have pervasive and varied effects on multiple dimensions of the body system. Short-term elevation of cortisol may be beneficial to the body system by gathering resources to perform. Chronic elevation of cortisol, however, is associated with negative health outcomes such as increased risk of diabetes.
and heart disease (Rosmond & Björntorp, 2000), and worsened immune functioning (McEwen, 1998). Interestingly, because cortisol can cross the blood-brain barrier and access parts of the brain related to cognitive functioning (i.e., working memory, memory consolidation, and memory retrieval), it is thought to be closely related to cognitive functioning (Lupien, Maheu, Tu, Fiocco, & Schramek, 2007; Wirth, 2015). Past literature on the relationship between cortisol and cognitive performance has generally found that elevated cortisol can impair cognitive functioning (de Quervain, Roozendaal, Nitsch, McGaugh, & Hock, 2000; Elzinga & Roelofs, 2005; Kirschbaum, Wolf, May, Wippich, & Hellhammer, 1996). More recent literature, however, suggests that elevated cortisol levels can sometimes be beneficial for cognitive functioning, if elevated in moderate levels (Lupien et al., 2007). This suggests that cortisol has a multi-faceted nature. Although cortisol has been primarily been associated with worse cognitive functioning, if elevated in moderate levels, cortisol can be beneficial for cognitive performance.

Although cortisol can influence cognitive functions that would seem closely related to academic performance, very few studies have extended these findings to real-world environments, such as the classroom. The few studies that have examined the relationship between cortisol activity and academic performance have found mixed results, with most finding no relationship between the two (Minkley, Westerholt, & Kirchner, 2014; Pletzer, Wood, Moeller, Nuerk, & Kerschbaum, 2010; Schoofs, Hartmann, & Wolf, 2008), and only one finding a negative relationship between cortisol and academic performance (Ng, Koh, & Chia, 2003). However, the situations in which this association has been tested have been limited. For example, past field research investigating the relationship between cortisol levels and performance has focused exclusively on cortisol levels on the day of an examination (Ng et al., 2003; Pletzer et al., 2010; Schoofs et al., 2008), leaving open the question of whether students’
cortisol levels during typical classes may be a better predictor of performance. Unlike an examination day, a typical class may involve more of the ongoing process of comprehension and consolidation of knowledge that would ultimately predict performance, and which may be undermined by elevated levels of cortisol. Thus, given the relatively small number of studies and methodological limitations, the relation between cortisol and academic performance is not yet clear.

Another reason that the association between cortisol and academic performance may be elusive could be because the interpretation of cortisol is more nuanced than it has often been represented. Based on an evaluation of the literature, I consider it possible that cortisol levels can represent psychological states other than negative stress in academic settings. Although cortisol is typically conceptualized as an outcome of a negative stressful state (Dickerson & Kemeny, 2004), recent research suggests it may indicate other psychological states as well. For instance, in some situations, cortisol may represent engagement with a task (Keller, Bless, Blomann, & Kleinböhl, 2011), which can potentially be positively related to performance.

Indeed, at a general level, physiological responses are known to represent different psychological states (stress versus engagement) depending on individuals’ perception of and reaction to their social environment, such as appraisal of one’s current situation (Blascovich & Mendes, 2010), appraisal of one’s own physiological response (Jamieson, Nock, & Mendes, 2012), and one’s mindset about stressful events (Crum, Salovey, & Achor, 2013). This line of literature potentially indicates that the relation between physiological arousal and performance can differ depending on how individuals interact with and interpret their environment.
In fact, a small body of recent research suggests that the consequences of cortisol on performance may depend on how individuals perceive their own physiological arousal (Akinola, Fridman, Mor, Morris, & Crum, 2016). For instance, in one lab study centered on a mock salary negotiation, participants were either told that physiological arousal was beneficial to one’s negotiation performance or told nothing about the effects of physiological arousal on negotiation performance. It was found that elevated levels of cortisol were associated with better performance (i.e., higher negotiated salaries) only among individuals who had been told that physiological arousal is beneficial (Akinola et al., 2016). This line of work suggests that the relationship between cortisol and academic performance is malleable and may depend on psychological factors such as how students perceive and react to their environment. Therefore, it may be important to consider student characteristics when evaluating the association between cortisol and classroom performance.

I propose that identity factors related to race/ethnicity (or other identities) may suggest conflicting possibilities of what cortisol levels represents in classroom settings. I suspect this because students may perceive and react to academic settings differently according to their race/ethnicity. For instance, students negatively stereotyped in academic settings (e.g., Blacks, Hispanics and Native Americans in the U.S.), compared to non-negatively stereotyped students (e.g., Asians and Whites), may perceive the classroom to be more hostile (Steele, 1997) and thus react to academic environments differently (Major & O’Brien, 2005).

For students negatively stereotyped in academic settings, cortisol levels could represent negative psychological stress, which would in turn undermine performance. This is consistent with a wide body of research showing the relative underperformance of identity-threatened
students (Aud, Fox, KewalRamani, & National Center for Education, 2010; Vanneman, Hamilton, Anderson, Rahman, & National Center for Education, 2009) and would provide a physiological explanation for such effects (Schmader, Johns, & Forbes, 2008). On the other hand, negatively stereotyped students’ cortisol levels can indicate the degree of engagement in the class and be positively associated with performance. Although perhaps counter-intuitive, research has shown that negatively stereotyped students can engage in two coping strategies. Negatively stereotyped students can disengage from and devalue the academic domain (Major & O’Brien, 2005; Schmader, Major, & Gramzow, 2001), which may relate to lower cortisol levels (Olff, Langeland, & Gersons, 2005) and worse performance. Alternatively, negatively stereotyped students can strive to overcome negative stereotypes attached to them (James, Hartnett, & Kalsbeek, 1983), which in turn, can relate to higher cortisol levels (Keller et al., 2011) and better academic performance. As noted above, cortisol levels for negatively stereotyped students may represent degrees of engagement and thus be positively related to academic performance.

For students less prone to experiencing pervasive negative stereotypes about their intellectual ability (i.e., Whites and Asians), it is less clear what classroom cortisol levels represent. If, as noted above, past research has failed to find an association between cortisol levels and performance because of methodological factors (i.e., the timing of cortisol assessments), examining cortisol in typical classes may yield the expected negative association between cortisol and performance. Although one might expect that cortisol could represent both negative psychological stress and engagement for non-stereotyped students, there may be less reason for there to be a particular pattern for these students, since non-stereotyped students are less likely to need to psychologically protect themselves in academic situations.
Goal of the Current Study

The goal of the current study is to focus on three questions that address a gap in the literature regarding the interpretation of cortisol, specifically in academic settings. First, I will investigate students’ cortisol responses in the academic setting. As our study measures cortisol across multiple classes and at multiple times each class, I can investigate cortisol levels both within and across classes. If levels of cortisol represent stress, I can investigate whether students were stressed during a class or whether students perceived particular classes to be more stressful than others. Second, I will investigate the relationship between cortisol responses during typical classes and students’ academic performance in a challenging college STEM course. As noted above, typical classes are where fundamental learning processes take place and if cortisol responses during such classes represent negative stress, higher levels may impede students’ learning process. Despite the potential influence that cortisol response during classes has on students’ learning process, to our knowledge, there have only been limited studies examining the relation between students’ cortisol response during typical classes and academic performance. Third, I will investigate whether the relationship between cortisol and academic performance is moderated by race/ethnicity. As far as I know, no previous study has investigated whether the association between cortisol and academic performance is moderated by race/ethnicity. As noted above, a number of studies suggest that negatively stereotyped students cope with the academic environment differently than non-negatively stereotyped students, particularly in challenging STEM classes where they are particularly likely to be underrepresented (Syed, Azmitia, & Cooper, 2011). Therefore, it is possible that the psychological meaning associated with cortisol, and thus the association between cortisol and academic performance, differs as a function of race/ethnicity.
In the literature review section below, I will first review the literature on the general physiology of cortisol and how cortisol has been typically studied in psychology. Further, I will describe the association between cortisol and various outcomes such as cognitive and academic performance that have been found in previous research. Then, I will review how and when cortisol may represent other psychological states than negative stress. Lastly, I will interpret what cortisol may represent in the academic setting and how its meaning and association with academic performance can vary by race/ethnicity.
CHAPTER 2: Literature Review

HPA Axis and Cortisol

The hypothalamic-pituitary-adrenocortical (HPA) axis is an adaptive mechanism that readies the body to manage changes in the environment (Dickerson & Kemeny, 2004; Lovallo & Thomas, 2000; Sapolsky et al., 2000; S. M. Smith & Vale, 2006). When changes are detected (e.g., encountering stressful situations), neurons in the hypothalamus secrete corticotropin-releasing hormone (CRH) (Carrasco & Van de Kar, 2003). CRH then triggers the pituitary gland, located in the anterior lobe of the brain, releasing a hormone called adrenocorticotropin (ACTH) (Simpson & Waterman, 1988). Secreted ACTH travels through blood vessels to the adrenal glands, where it triggers the secretion of cortisol.

Once released, cortisol prepares the body for action by helping it generate energy. Cortisol does this by activating the generation of glucose in the liver, breaking down fat and protein in the body, and inhibiting glucose uptake in peripheral tissues, thereby raising the concentration of glucose circulating in blood (Marik & Bellomo, 2013; Sapolsky et al., 2000). Cortisol also regulates other bodily systems, such as the immune system. For example, in the face of acute changes in the environment, cortisol inhibits inflammation to stockpile energy for immediate use (Dedovic & Duchesne, 2012; Segerstrom & Miller, 2004). In sum, cortisol mobilizes resources for immediate use by elevating glucose levels and inhibiting body systems not necessarily needed in the face of a challenging situation.

When is Cortisol Secreted?

Typically, cortisol is thought to be secreted in the face of negative physiological and psychological stress (Dickerson & Kemeny, 2004; Sapolsky et al., 2000). For instance, cortisol
levels rise in response to adverse physical conditions like physical pain (Hannibal & Bishop, 2014; Van Uum et al., 2008) and cold (Pääkkönen & Leppäluoto, 2002). Cortisol is also secreted in response to psychological threats, especially those that are socially evaluative in nature (Dickerson & Kemeny, 2004). For example, the Trier Social Stress Test (TSST) is reliably used to elicit cortisol responses in behavioral research. In this task, participants typically have to give a speech and orally respond to a challenging arithmetic task in front of judges whose demeanor is either neutral or negative (Dickerson & Kemeny, 2004; Kirschbaum, Pirke, & Hellhammer, 1993). Most participants report the TSST to be a highly stressful experience, and cortisol spikes are reliably produced. Overall, these examples illustrate that cortisol can be released in response to situations that are often thought to be negative and stressful.

However, cortisol can also be released in response to situations that are psychologically neutral or positive. For example, cortisol is released in response to emotion-ally-neutral everyday happenings like waking up (Kirschbaum & Hellhammer, 1989), which readies the body for the upcoming activities of the day (Dallman et al., 2002). Also, cortisol can be secreted during activities that one may find exciting and enjoyable, such as during exercise (Jacks et al., 2002), skydiving (Chatterton et al., 1997), and sports competition (Carré, Muir, Belanger, & Putnam, 2006). In the examples described above, the stimuli do not necessarily have to be psychologically negative to elicit a cortisol response; rather, a reaction may be elicited according to whether the body requires additional energy to properly respond to a stimuli or situation, regardless of its valence. Below I will describe how past literature in psychology has investigated cortisol.
How Cortisol is Studied in Past Psychology Literature

As noted above, although cortisol can be secreted in response to stimuli that are psychologically neutral or positive, cortisol is typically thought of as an outcome of stress. As a result, cortisol response to psychologically stressful stimuli has been widely studied in both laboratory and field settings. Many studies have investigated cortisol responses to stressful stimuli in a laboratory setting. As mentioned above, a typical stressor used to elicit cortisol response in the lab setting is the TSST (Kirschbaum et al., 1993). Usually during the TSST, cortisol levels are measured before, during and after a stressful stimulus (i.e., giving a speech to an audience) to identify the cortisol response pattern. Past literature has further investigated whether this cortisol response pattern differed by various social psychological factors such as race/ethnicity (Hostinar, McQuillan, Mirous, Grant, & Adam, 2014), SES (Tackett, Herzhoff, Smack, Reardon, & Adam, 2017), and gender (Liu et al., 2017). These results reveal that negatively stereotyped racial groups and individuals low in SES often show blunted cortisol responses to a stressor (Hostinar et al., 2014; Tackett et al., 2017). These blunted cortisol responses are represented by lower cortisol peaks and they potentially represent dysregulation of the HPA axis rooting from chronic stress (McEwen, 1998). Additionally, men have shown higher levels of cortisol relative to women during the peak of cortisol response to the TSST (Liu et al., 2017). These laboratory studies support further investigation into cortisol responses to stressful stimuli and how various between-person factors may alter the pattern of cortisol responses. Further, these laboratory studies can strictly control for factors that may influence cortisol, such as the time of day when cortisol is collected.
Additionally, past literature has investigated the association between elevated cortisol and cognitive functioning in lab settings. In such cases, cortisol can be elevated exogenously (e.g., administration of glucocorticoid by pharmacological treatment) (de Quervain et al., 2000) or endogenously (e.g., going through a TSST) (Kirschbaum et al., 1993). Once participants’ cortisol levels are elevated, participants’ cognitive performance is assessed. This method also allows researchers to assess how specific types of cognitive functioning (e.g., working memory, encoding information, retrieval of memory) are related to cortisol levels. Overall, despite their advantages, laboratory studies cannot answer whether cortisol levels are elevated in the face of stressful situations in real-life and whether levels of cortisol are associated with performance in real-life tasks.

Although less common than laboratory studies, cortisol responses to stressful situations in real-life settings have also been assessed. Past literature has investigated cortisol responses to real-life stressors such as taking an exam in academic setting (Lacey et al., 2000; Martinke, Oberascher-Holzinger, Weishuhn, Klimesch, & Kerschbaun, 2003; Ng et al., 2003; Preuß, Schoofs, Schlotz, & Wolf, 2010; Schoofs et al., 2008). In this research design, students’ cortisol levels are typically measured before and after a stressor (e.g., exam), as well as on a day when there are no stressors (control day). Cortisol levels assessed on the control day are normally used as a baseline to determine whether cortisol levels are elevated on the day of the stressor. Sometimes in field research, factors that may influence cortisol levels, such as time of day when saliva samples are collected, are not perfectly controlled. For instance, in a study assessing the association between students’ cortisol levels during an exam (cortisol was measured right before and after the exam) and their exam performance, the time of day when students took the exam varied (Schoofs et al., 2008). Despite some limitations, this method grants researchers a better
understanding of cortisol and its impact on performance in real-life settings. For instance, past literature has identified that real-life stressors such as taking an exam can elevate cortisol levels (Lacey et al., 2000). Further, the relationship between cortisol levels on the day of an exam and exam performance has been investigated (Minkley et al., 2014; Pletzer et al., 2010; Schoofs et al., 2008).

Lastly, instead of looking at cortisol response as a reaction to stressful stimuli, diurnal patterns of cortisol have also been investigated. As noted above, cortisol is known to follow a diurnal rhythm characterized by a peak 30 minutes after waking up, which is referred to as the cortisol awakening response (CAR). After this peak in the morning, cortisol levels are known to decline throughout the day and reach their lowest point around midnight (Kirschbaum & Hellhammer, 1989). Typically, cortisol levels are measured at the moment of awakening, 30 minutes after awakening, in the afternoon, and right before going to bed. This approach allows researchers to investigate general patterns of cortisol (i.e., CAR, slopes of cortisol throughout the day) in the natural setting and their association with health outcomes (Doane et al., 2013) or cognitive functioning (Stawski et al., 2011).

As noted above, the current study longitudinally investigates students’ cortisol response in a challenging gateway college science course. In this course, I measure students’ cortisol levels across multiple classes. Although, this research design is less commonly used, my research design has some methodological advantages. My longitudinal research design allows me to investigate cortisol responses in different phases of the semester (e.g., classes around the exam, class in the beginning of the semester), thus allowing me to identify whether students found certain periods of the semester more challenging than others. Additionally, having multiple
cortisol samples in a course throughout the semester allows me to capture students’ comprehensive experience in the course. Further, other studies often use a research design similar to mine. For instance, in a recent longitudinal field study looking at middle school students’ cortisol levels in the academic setting, the research team collected saliva samples from students over six academic days in the school (one sample per student a day) (Lee, Jamieson, Miu, Josephs, & Yeager, 2018). This study, which investigated how adverse events in the academic setting (i.e., declining grades) can affect students’ cortisol levels, demonstrates precedent for the type of research design used in the current research.

Having laid out a general description on what cortisol is, when it is secreted, and how it has been investigated, in the section below, I will narrow my scope to describe the impact of cortisol on various outcomes (e.g., cognitive functioning and academic performance).

**Impact of Elevation of Cortisol on Body and Cognitive Systems**

The elevation of cortisol impacts multiple aspects of the body and cognitive systems in a variety of ways. While short-term elevation may energize and prepare the body for challenging situations, chronically elevated cortisol is associated with various negative health outcomes, such as diabetes, obesity, and heart disease (Lloyd, Smith, & Weinger, 2005; Rosmond & Björntorp, 2000), as well as worsened immune functioning (McEwen, 1998). For example, chronically elevated cortisol increases the risk of Type 2 diabetes by slowing the body’s response to insulin (Sapolsky et al., 2000). Cortisol also impacts cognitive function by binding to glucocorticoid (cortisol) receptors in the prefrontal cortex, hippocampus, and amygdala, which are all primarily in charge of memory and learning processes (Lupien et al., 2007; Wirth, 2015). As cortisol binds to receptors in these areas, it impacts the long-term potentiation (LTP) of neurons. LTP is a
process that is characterized by neurons strengthening connections with each other, and it is known to be a primary neural mechanism for learning and processing information. High elevation of cortisol has been generally found to impede LTP, but moderate levels of cortisol have been found to enhance LTP (Joels, Pu, Wiegert, Oitzl, & Krugers, 2006; Popoli, Yan, McEwen, & Sanacora, 2011; Wirth, 2015). The process described above represents neural mechanisms of how cortisol impacts cognitive functioning. Again, this evidence suggests multifaceted nature of cortisol as it can be both detrimental or beneficial for learning. In the following section, I will further elaborate upon the types of cognitive functions that cortisol can influence and the directions in which they can be influenced.

**Cortisol and Cognitive Functioning**

Past research has generally found that elevated levels of cortisol can impair cognitive functioning in humans (de Quervain et al., 2000; Elzinga & Roelofs, 2005; Kirschbaum et al., 1996). For instance, elevated cortisol has been found to undermine working memory (Braunstein-bercovitz, 2003; Elzinga & Roelofs, 2005; Oei, Everaerd, Elzinga, van Well, & Bermond, 2006), memory retrieval (de Quervain et al., 2000), and the integration of new information into existing memories (Wang, Zhao, Ghitza, Li, & Lu, 2008). Further, chronically elevated cortisol is known to cause permanent damage to the hippocampus, which can impede learning and memory functioning (Lupien et al., 1998).

However, more recent evidences suggest the relation between cortisol and cognitive functioning follows an inverted U-shaped curve, indicating that moderate elevation of cortisol can be beneficial for cognitive functioning (Lupien et al., 2007). For example, an inverted U-shaped relation was found between cortisol and working memory (Moriarty et al., 2014).
memory encoding (Joels et al., 2006), and memory retrieval (Schilling et al., 2013). As noted in the previous section, this evidence once again suggests that cortisol has a multifaceted nature, in which moderate levels of cortisol can benefit cognitive functioning, rather than harm it.

Overall, past literature has generally found that elevated cortisol is detrimental to cognitive performance. However, it is also important to note that cortisol can benefit cognitive performance. As noted above, if cortisol influences basic cognitive functioning, it stands to reason that it possesses some relation to academic performance. In the next section, I will review literature that has investigated the relation between cortisol and academic performance.

**Cortisol and Academic Performance**

A theoretical review paper by (Vogel & Schwabe, 2016) suggests that cortisol can impede students’ academic performance. For example, as explained above, elevated cortisol can potentially prevent students from learning new material by hindering their working memory and encoding and consolidation processes necessary to integrate new information into their existing knowledge. Moreover, impaired memory retrieval caused by cortisol may hinder students’ performance on exams.

Despite the potential influence of cortisol on academic performance, the majority of prior research investigating cortisol in academic settings has investigated cortisol as an outcome indicating stress, rather than as a predictor of performance. This research has typically found cortisol levels to be higher on exam days than on a control day (when students did not have an exam) (Gaab, Sonderegger, Scherrer, & Ehlert, 2006; Lacey et al., 2000; Martinek et al., 2003; Verschoor & Markus, 2011; Weekes et al., 2006).

15
The studies cited above are generally taken as evidence that cortisol can be interpreted as a stress indicator, but only a handful of studies have examined whether cortisol predicts academic performance. These studies have yielded inconclusive results. In fact, although elevated cortisol is thought to be negatively associated with academic performance, most studies have found no association between cortisol and academic performance (Minkley et al., 2014; Pletzer et al., 2010; Schoofs et al., 2008). Schoofs et al. (2008), for example, examined cortisol levels of students immediately before and after an oral exam, which was a requirement for graduation. They found that cortisol levels were elevated on the day of the exam compared to a control day, which was within 7 days before or after the exam. Results revealed that there was no association between absolute cortisol concentration on the day of the exam and students’ exam scores. Additionally, Pletzer et al. (2010) looked at students’ cortisol levels 15 minutes before and after a real-life statistics exam in an introductory college statistics class. They did not find an overall pattern of elevated cortisol on the exam day relative to a control day, which was a typical day in the course without an exam. Further, they did not find any association between absolute cortisol levels on the day of the exam and students’ academic performance. Additionally, academic performance did not have any association with difference in cortisol levels between the exam and control day.

To my knowledge, only one study has found a negative relation between cortisol and academic performance (Ng et al., 2003). In this study, the cortisol levels of graduate students were assessed immediately before and after a final examination, which students were required to take for graduation. Ng et al. (2003) found that cortisol levels assessed immediately before the exam were negatively correlated with students’ scores. Although this study found a negative association between cortisol and performance, it is contrary to most of the previous findings,
which indicate a null relation between the two. Further, all studies noted above do not primarily focus on the relationship between cortisol levels on typical classes and academic performance. Although, Pletzer et al. (2010) measured cortisol during a typical class, they only used it as a baseline measure and did not investigate its direct relation with performance. Further, Schoofs et al. (2008) did not measure students’ baseline cortisol during a typical class. Instead, Schoofs and colleagues (2008) measured cortisol levels on a normal day (regardless of whether students had class on that day or not) within 7 days before or after the exam. Overall, the literature investigating the relationship between cortisol and academic performance is quite limited, having only produced mixed results while restricting its scope primarily to exam days.

More research is needed to clearly identify the relationship between cortisol and academic performance. Further, research that captures students’ comprehensive experience in the classroom setting is needed. As noted above, researchers have primarily used cortisol response on exam days as a predictor of performance. However, such an approach does not ask whether students’ cortisol responses during typical classes are associated with students’ academic performance. Typical classes are when fundamental learning processes take place, from learning new material to integrating content into existing knowledge, and these processes could ultimately relate to academic performance. It is possible that cortisol levels during these typical classes are a better predictor of students’ academic performance, as elevation in cortisol could impede students’ learning processes. Perhaps past literature has not found any relation between cortisol and academic performance is because it did not investigate how cortisol responses during typical classes relate to academic performance. However, I acknowledge the alternative that elevated cortisol could be beneficial, as some elevation of cortisol has been found to be beneficial for cognitive performance. Additionally, as I will explain in the next section, recent studies suggest
that elevated levels of cortisol can represent engagement with a task, which in turn may positively be related to performance.

**Cortisol Representing Engagement**

Based on past literature, I believe it is possible that cortisol responses in an academic setting can not only represent negative stress, but also other psychological states such as engagement. Contrary to the common perception that cortisol solely represents negative psychological states (Dickerson & Kemeny, 2004), cortisol can rise in reaction to novel activities that individuals find enjoyable—cortisol can even be associated with positive affect (Hoyt, Zeiders, Ehrlich, & Adam, 2016). More specifically, recent literature suggests that cortisol can indicate engagement with a task. Keller and colleagues (2011) instructed participants in a lab study to play the computer game Tetris for 15 minutes. The primary goal of Tetris is to arrange falling objects to completely fill the bottom lines of a playing field. Participants were assigned to a boredom, fit, or overload condition. In the boredom condition the game difficulty was very low, while in the overload condition the game was overwhelmingly difficult. In the fit condition, the game difficulty was consistently adjusted throughout the game based on participants’ performance. Baseline saliva samples were collected before participants engaged in the task, and three extra saliva samples were taken 20, 25, and 30 minutes after they started playing the game. Involvement in the task was assessed after participants finished playing. Participants in the fit condition reported the highest level of involvement with the game and had the highest cortisol level increase from baseline. The association of higher levels of cortisol with greater engagement with the game suggests that elevated cortisol can represent task engagement rather than just negative psychological states.
More recent findings also suggest that elevated cortisol is, when in moderate elevations, associated with engagement. In one study, participants were brought into the lab and engaged in the TSST. After the TSST, participants engaged in a complicated computer task using a program called Cabin Air Management System (CAMS), which simulates the complex environment of a spaceship. Participants’ perception of involvement in the task was assessed right after the CAMS task. Cortisol levels were assessed 5 times after the TSST task. Moderate levels of cortisol right after the TSST (the first two samples taken after the TSST averaged together) were associated with highest level of involvement relative to lower or higher levels of cortisol (Peifer, Schulz, Schächinger, Baumann, & Antoni, 2014). Indeed, the studies noted above suggest that elevated cortisol, to some degree, can be associated with one’s engagement with a task.

Furthermore, past literature suggests that low levels of cortisol can be associated with disengagement from stressful social environments. One such study worked with soldiers who were on active duty. These soldiers were going through a survival course at US Army survival school, and the experiment took place during a training session in which soldiers had to go through a mock interrogation process that was both physically and psychologically demanding. Soldiers’ cortisol levels and dissociative experiences (e.g., “Do you feel as if you are watching the situation as an observer or a spectator?”) were assessed right after the interrogation process. The results revealed that cortisol levels during the interrogation were negatively associated with dissociation (Morgan et al., 2001). This study suggests that low levels of cortisol can represent disengagement from stressful situations. Overall, the studies described in this section suggest that cortisol can indeed represent engagement with the task at hand or an encountered social environment.
As these studies suggest, if cortisol represents engagement, especially in the classroom setting, it is possible that students’ cortisol levels could be positively associated with academic performance. So far, I have illustrated that cortisol can represent psychological states (stress and engagement) that are quite different from each other. If cortisol can represent different psychological states, it is important to identify the conditions that shape its psychological meaning. Indeed, in the next section I will describe factors that are known to influence the meaning of physiological arousal, specifically in regards to cortisol.

### Interpreting Physiological Arousal

Indeed, at a general level, research has long suggested that physiological arousal can have different psychological meanings depending on how individuals perceive and react to their social environment (e.g., Schacter & Singer, 1962). For example, imagine two students about to take an exam, one who has studied very hard and the other who has not. Both students could feel aroused before taking an exam, yet while one student’s heart may be beating fast from excitement about taking the exam, the other is experiencing the same feelings because of a fear that lack of preparation will result in failure. This simple example illustrates how similar experiences of physiological arousal can indicate different psychological states depending on how individuals perceive the situation. As I discuss below, the psychological meanings of physiological arousal can vary as a function of how people appraise a situation (Blascovich & Mendes, 2010), how they appraise their own physiological arousal (Jamieson et al., 2012), and their stress mindset (Crum, Akinola, Martin, & Fath, 2017).

One factor that may shape the psychological meaning of physiological arousal is one’s appraisal of an encountered situation. For instance, people can perceive a stressful situation to be
challenging or threatening (for review see Blascovich & Mendes, 2010). Situations are perceived as challenging when a person perceives him or herself as possessing sufficient resources to cope. On the other hand, situations are perceived as threatening when a person perceives him or herself as lacking the resources to cope. On the surface, both challenging and threatening appraisals are associated with increased heart rate, yet each is associated with contrasting patterns of cardiovascular activity. If one appraises the situation as challenging, one’s heart beats fast and efficiently: one’s arteries expand and thus, provide more blood flow (and thus more oxygen) to peripheral sites, such as the brain. This pattern of cardiovascular activity can be adaptive, as it mobilizes energy in the body by providing more oxygen to peripheral sites. Also, this pattern is known to be associated with an approach motivation (Tomaka, Blascovich, Kelsey, & Leitten, 1993). On the other hand, if one appraises the situation to be threatening, one’s heart beats in an inefficient way: one’s arteries are contracted and thus, not able to provide as much blood flow to the peripheral sites. This pattern of cardiovascular activity can be detrimental to performance as it can limit the supply of oxygen to the brain and is associated with an avoidance motivation (Mendes, Blascovich, Hunter, Lickel, & Jost, 2007; Tomaka et al., 1993). This line of literature suggests that the same pattern of physiological arousal (e.g., increased heart rate) can represent different psychological states depending on how people appraise a situation.

Another factor that can color the psychological meaning of physiological arousal is how one appraises his or her own physiological arousal. In a lab experiment, participants were randomly assigned to two different conditions before completing a TSST. Before engaging in the TSST, participants in the intervention condition were told by the research team that physiological arousal is functional and adaptive, and they were asked to read articles that explain the adaptive role of physiological arousal. On the other hand, students in the control condition were told by
the researchers that ignoring physiological arousal is beneficial to performance, and they were asked to read three articles that explained why ignoring one’s own arousal may be beneficial for one’s performance. As described in the previous section, participants in the intervention condition showed adaptive cardiac activity, which was observed when individuals appraise their situation to be challenging. On the other hand, participants in non-intervention condition showed maladaptive cardiac activity, which is observed when individuals appraise their situation to be threatening (Jamieson et al., 2012). Thus, the same increased cardiac activity had different psychological and physiological implications.

In the previous two paragraphs, I described studies that demonstrate how a given measure of physiological arousal can represent different psychological meanings, focusing primarily on cardiac reactivity. It is possible that the same ambiguity in interpretation can apply to cortisol. Indeed, in the next paragraph I will explain an example that illustrates that the same level of cortisol elevation can have different psychological and physiological implications depending on one’s mindset about stress.

A person’s mindset about the role of stress can affect that person’s psychological meaning of physiological arousal (Crum et al., 2017; Crum et al., 2013). Depending on whether an individual perceives stress to be associated with positive or negative outcomes, a person can experience one of two types of stress mindset. A stress-is-enhancing mindset is characterized by a belief that stress can augment various outcomes such as performance, health, and well-being. On the other hand, a stress-is-debilitating mindset is typified by a belief that stress deters those outcomes. A recent lab study investigated whether participants’ stress response differed as a function of stress mindset. Individuals were brought into the lab to engage in a modified TSST,
which was described to them as a mock interview. Before participants began, they were randomly assigned to a manipulation (watching a 3-minute long video clip) that induced either a stress-is-enhancing mindset or a stress-is-debilitating mindset. In both conditions, participants’ cortisol levels were elevated during the TSST. Additionally, participants’ cortisol level did not differ by condition, although the physiological and psychological implications of elevated cortisol did. Dehydroepiandrosterone-sulfate (DHEAS), which is an anabolic hormone known to play a protective and regenerative role such as regeneration of cells (Theorell, 2008), was also measured during the TSST. Participants in the stress-is-enhancing condition showed a higher level of DHEAS and more positive affect than those in the stress-is-debilitating condition (Crum et al., 2017). This result suggests that cortisol response may represent different psychological states depending on the condition. Despite having the same levels of elevated cortisol as their control counterparts, participants in the stress-as-enhancing condition experienced the cortisol increase in association with more adaptive physiological functioning (increased levels of DHEAS) and positive psychological states.

As noted in the current section, past literature suggests the representation of physiological arousal can be malleable as a function of individuals’ perception of and reaction to their situation. The next question that would follow is whether these different experiences of physiological arousal are associated with performance outcomes.

**Malleable Relation between Physiological Arousal (Cortisol) and Performance**

Indeed, for instance, research suggests that physiological arousal can improve performance when individuals believe that this arousal should benefit their performance. In one study, students who were planning to take the Graduate Record Examination (GRE) in the near
future were brought into the lab and assigned to either a control or an intervention condition. Students assigned to the intervention condition were told by the experimenter that physiological arousal was beneficial for their performance, whereas students in the control condition were not given any information about the effect of physiological arousal on performance. After the condition manipulation, students completed a mock GRE test in a setting that resembled an actual GRE test site. Students’ salivary alpha amylase (sAA) was measured 3 times: on the day before the experiment (T0), on the day of the experiment before the mock GRE test (T1), and after the mock GRE test. Physiological arousal is represented by activation of sympathetic nervous system and can be measured by sAA (Nater & Rohleder, 2009). For students who were instructed to think that physiological arousal was beneficial to their performance, an increase in sAA during the mock GRE test (sAA level at T1 subtracted from T2) was positively associated with a better math score. This relation held after controlling for participants’ baseline sAA measured at T0. There was no association between elevation in sAA level and students’ verbal score. For participants in the control condition, there was no association between sAA elevation and performance in both sections. Interestingly, only for participants in the intervention condition did elevated sAA during the mock GRE positively predict actual GRE math score 1-3 months after the experiment (Jamieson, Mendes, Blackstock, & Schmader, 2010). This study suggests that reappraising one’s physiological response can not only alter the psychological representation of physiological response but also influence its association with performance. Further, this research illustrates the salient impact of appraisal of one’s own physiological arousal on performance, as sAA positively predicts actual GRE math scores taken 1-3 months after the experiment for individuals in the reappraisal condition. If the relation between arousal, represented by sAA, and performance is moderated by reappraising one’s own physiological
response, it is possible that the relation between cortisol and performance will be also moderated by psychological factors.

Indeed, reappraising one’s physiological response is known to alter the relation between cortisol reactivity and performance. As described in the earlier section, in a lab experiment, participants engaged in a salary negotiation task and their cortisol levels were assessed before and after the negotiation. Before negotiations began, participants were either told that physiological arousal was beneficial to one’s negotiation performance (intervention condition) or told nothing about the effect of physiological arousal on negotiation performance (control condition). For participants in intervention condition, there was a positive relation between negotiation performance and cortisol response. In contrast, there was a negative relationship between cortisol and negotiation performance for participants in the control condition (Akinola et al., 2016). This work is important because it is one of the first studies to demonstrate that cortisol and its relation to performance can be molded depending on one’s appraisal of his or her own physiological arousal.

So far, I have illustrated that the psychological meaning of physiological response and its relation with performance can be moderated by how individuals perceive and react to their environment. To more clearly understand the representation of cortisol and its association with performance in the classroom setting, it is important to identify a potential moderator that may alter students’ experience in the classroom setting. Further, it is important to understand whether such a moderator may alter the psychological meaning of cortisol and its relation with academic performance.
Interpreting Cortisol in the Classroom: Does the Relation Between Cortisol and Academic Performance Differ by Race/Ethnicity?

Beyond examining the pattern of cortisol response and its relationship with academic performance, I further consider the possibility that the psychological meaning of cortisol and its association with performance can be moderated by students’ race/ethnicity. In the section below, I will describe why I consider race/ethnicity as a potential moderator of the psychological meaning of cortisol and its relationship with academic performance.

First, students may perceive and react to the academic setting differently depending on their race/ethnicity. For example, students from racial/ethnic groups that are associated with negative stereotypes about intellectual ability (e.g., Black, Hispanic and Native American students) may perceive the academic environment to be more hostile and threatening compared to White and Asian students whose race/ethnicity is not associated with these stereotypes (Steele, 1997). This social identity threat can lead students from groups associated with negative intellectual stereotypes to worry that there could be a kernel of truth to the stereotypes, or that their performance will be judged as representative of their entire group (Steele, Spencer, & Aronson, 2002). They may be concerned, understandably, that they will be treated with bias (G. L. Cohen & Garcia, 2005). These worries and concerns may be amplified in a STEM classroom, such as where the research reported here was conducted, because STEM domains are particularly associated with perceptions that intellectual giftedness is required (Leslie, Cimpian, Meyer, & Freeland, 2015). This helps explain the finding that negatively stereotyped students are pronouncedly less represented in STEM fields compared to other departments (Syed et al., 2011). Given the past literature, negatively stereotyped students are more likely to perceive the classroom setting to be threatening than non-negatively stereotyped students. As noted above, if
the situation is appraised as threatening, physiological arousal can be associated with an avoidance motivation. If so, for negatively stereotyped students, cortisol can represent negative stress and be negatively associated with academic performance. Further, if cortisol represents stress, negatively stereotyped students may have higher levels of cortisol compared to non-negatively stereotyped students. On the other hand, as non-negatively stereotyped students do not have to face the added burden of social identity threat in the academic setting, they may be more likely to perceive the classroom as challenging rather than threatening. If so, cortisol can possibly be associated with an approach motivation for those students. In this case, cortisol for non-negatively stereotyped students may represent engagement and be positively related to performance.

Further, negatively stereotyped students are known to respond to academic environments differently from their non-negatively stereotyped counterparts. As a way of coping with identity threat, negatively stereotyped students may disengage their self-worth from the academic domain (Crocker & Major, 1989; Major & O’Brien, 2005), which can be associated with lower levels of cortisol (Mason et al., 2001; Morgan et al., 2001; Olff et al., 2005), and possibly worse performance. On the other hand, negatively stereotyped students may respond by striving particularly hard to overcome the stereotypes attached to their group (James et al., 1983). This strategy can be associated with higher levels of cortisol (Keller et al., 2011), and potentially better performance. If this is the case, levels of cortisol can represent degrees of engagement for negatively stereotyped students and have a positive association with performance. As noted above, non-negatively stereotyped students are less likely to engage in the aforementioned two forms of coping strategy than negatively stereotyped students. If true, it is less likely that cortisol
will represent engagement for non-stereotyped students. Thus, for non-negatively stereotyped students, cortisol may be negatively associated with performance.

Additionally, students’ motivations and goals for pursuing an academic career are known to differ by race/ethnicity. Differences in motivation may impact how students perceive their academic environment, which in turn, can shape how cortisol relates to performance. For instance, relative to White students, it has been suggested that Black, Hispanic, and Native American students place greater importance on altruistic goals as a motivation for their education, such as helping others through one’s work and contributing to one’s community (Fryberg & Markus, 2007; Torres, 2009). Endorsing altruistic goals is positively associated with motivation for higher levels of academic pursuit, particularly for negatively stereotyped students (J. L. Smith, Cech, Metz, Huntoon, & Moyer, 2014; Thoman, Brown, Mason, Harmsen, & Smith, 2015). Overall, the past line of literature suggests that altruistic goals (e.g., helping others through one’s work) are particularly important for motivating negatively stereotyped students to pursue their academic career. However, academic institutions in North America are known to emphasize independent goals such as individual success and development more than altruistic goals (Stephens, Fryberg, Markus, Johnson, & Covarrubias, 2012). Therefore, negatively stereotyped students may perceive an incongruency between their own goals and the goals vaunted by their academic institution. This perceived incongruency may lead negatively stereotyped students to view their academic setting as a hostile environment. If so, as noted above, cortisol may represent negative psychological states and may be negatively associated with academic performance. On the other hand, as non-negatively stereotyped students may perceive less incongruency between their goals and goals valued by academic institution, cortisol
may more likely to represent engagement for non-negatively stereotyped student and thus be positively related to performance.

Lastly, there is evidence that the HPA axis functions differently across races. As described earlier, the HPA axis is a system in the body that regulates secretion of cortisol in the face of the stressful or challenging situations. Past literature suggests that recurring activation of the HPA axis due to chronic stress, such as discrimination and poverty, is known to dysregulate individuals’ HPA axis (Adam et al., 2015; Bush, Obradovic, Adler, & Boyce, 2011; Hajat et al., 2010; Lupie, King, Meaney, & McEwen, 2001; McEwen, 1998). This dysregulation is thought to be a symptom of the cumulative wear and tear of the body’s stress system (i.e., allostatic load) from its chronic and inefficient activation. In negatively stereotyped racial groups, a flattened diurnal pattern of cortisol (DeSantis et al., 2007; Hajat et al., 2010) and a blunted cortisol response to stress in a laboratory setting (Hostinar et al., 2014), has been interpreted as dysregulation. It is also possible that negatively stereotyped students can have lower levels of cortisol, as chronic stress is known to reduce overall secretion of cortisol throughout the day, a state known as hypocortisolism (Miller, Chen, & Zhou, 2007).

Overall, this dysregulated cortisol pattern potentially indicates an inability to activate the HPA axis in times of need. For example, flatter diurnal cortisol patterns for negatively stereotyped students arise from a lower cortisol peak in the morning and a higher level of cortisol during the evening. The lower peak of cortisol in the morning may represent an inability to fully activate the HPA axis to ready the body for the demands of the day (Dallman et al., 2002). Further, racial/ethnic minority children have been shown to have a blunted cortisol response to the TSST compared to White children (Hostinar et al., 2014).
These dysregulated cortisol patterns are known to be negatively associated with various health outcomes and cognitive functioning; therefore, it is possible that cortisol can have a different association with performance for negatively stereotyped students. For example, flattened diurnal cortisol patterns have been associated with negative cognitive functioning (Rickenbach, Almeida, Seeman, & Lachman, 2014; Stawski et al., 2011), depression (Doane et al., 2013) and cardiovascular disease (Kumari, Shipley, Stafford, & Kivimaki, 2011). Furthermore, hypocortisolism has been found to be negatively associated with memory task performance (Tiemensma, Andela, Biermasz, Romijn, & Pereira, 2016). Given this past literature it is possible that the relationship between cortisol and performance may vary as a function of race/ethnicity. For example, low levels of cortisol for negatively stereotyped students may not only represent low levels of stress but also represent hypocortisolism or a dysregulated cortisol response that may be negatively associated with students’ academic performance.

In the section above, I described reasons why I think race/ethnicity will alter the meaning of cortisol and its association with academic performance. Below, summing up information from this section, I will further discuss what cortisol may represent and how it may relate to performance for different races/ethnicities in the academic setting.

**Cortisol levels for negatively stereotyped students.** As noted above, cortisol may indicate psychological stress for students who belong to a negatively stereotyped ethnic/racial group. In this case, cortisol may be negatively associated with academic performance. In fact, a negative association of cortisol with performance would provide a physiological explanation for how race-based stress undermines the performance of negatively stereotyped students (Levy, Heissel, Richeson, & Adam, 2016; Schmader et al., 2008). This finding would also be consistent
with the idea that negatively stereotyped students may appraise the classroom setting as threatening rather than challenging. As noted in the previous section, threat is associated with an avoidance motivation. It is thus possible that cortisol may represent negative psychological states and be negatively associated with academic performance.

As briefly noted above, another possibility is that cortisol levels during a class can represent engagement for negatively stereotyped students, and hence relate positively to academic performance. Past literature suggests that negatively stereotyped students may cope with identity threat in academic settings in two widely disparate ways. As noted previously, one possible coping strategy is to put in less effort or separate academic performance from one’s self-worth (Crocker & Major, 1989; Major & O’Brien, 2005; Major, Spencer, Schmader, Wolfe, & Crocker, 1998; Schmader et al., 2001; Steele et al., 2002; Woodcock, Hernandez, Estrada, & Schultz, 2012). This disengagement strategy may relate to low levels of cortisol (Keller et al., 2011; Mason et al., 2001; Morgan et al., 2001; Olff et al., 2005) and worse performance. Negatively stereotyped students have been found to lessen their effort or deprioritize the importance of academics to their self-worth in both field and experimental settings. In a longitudinal study of 10th-grade students, it was found that self-esteem was positively related to grades only for White students, although there was a weak association for Black female students (Osborne, 1995). Additionally, recent research by Woodcock et al. (2012) investigated deidentification process among Black and Hispanic college students in science majors. Black and Hispanic students who perceived greater stereotype threat had lower identification with science, which was associated with less intention to pursue a scientific career. A similar idea was tested in an experimental study in which both Black and White college students were brought into a lab to take a bogus intelligence test, which was presented as a standardized test of intelligence.
Students were given either positive or negative feedback about their test results, and their self-esteem was assessed before and after taking the test. While White students had higher post-test self-esteem when they received positive feedback than when they received negative feedback, Black students’ self-esteem did not differ as a function of feedback received (Major et al., 1998). This suggests that Black students were more likely than White students to separate their self-esteem from the outcome of a test. Overall, these examples suggest the possibility that low cortisol can represent disengagement for negatively stereotyped students and is associated with worse performance. Further, as noted previously, low cortisol levels for negatively stereotyped students can potentially signal dysregulated HPA axis which has been found to negatively associated with cognitive performance.

On the other hand, it is possible that negatively stereotyped students may strive harder to overcome salient negative stereotypes in a valued domain. This coping strategy can be characterized by John Henryism, an active coping strategy (e.g., working extremely hard) invoked in the face of chronic stressors like social discrimination or poverty (James et al., 1983). Indeed, it has been found that African Americans reported higher John Henryism than Whites (James, Strogatz, Wing, & Ramsey, 1987). These results suggest that negatively stereotyped students are more likely to go to extra lengths to do better in the class. This coping strategy may thus be associated with more engagement in class, higher levels of cortisol (Keller et al., 2011), and better performance.

**Cortisol levels for non-negatively stereotyped students.** Cortisol also may represent negative psychological states for non-negatively stereotyped students. The current study took place in a challenging gateway college science course which is known to be difficult and
important as grades from this course could determine students’ eligibility to continue in their selected major. Given the importance and difficulty of the class, it is possible that all students, regardless of their race/ethnicity, may have perceived the classroom setting to be hostile and threatening. Additionally, cortisol may be more likely to represent negative stress than engagement for non-stereotyped students as these students are less likely to engage in the aforementioned two forms of coping strategy (i.e., disengagement and striving) than negatively stereotyped students.

However, it is also possible that cortisol may represent engagement for non-negatively stereotyped students, and thus positively correlate with academic performance. As described previously, non-negatively stereotyped students, may perceive academic settings to be challenging more so than threatening. As mentioned above, physiological arousal has been associated with motivation related to approach when the situation was appraised to be challenging. If this is the case, cortisol for non-negatively stereotyped students can potentially represent engagement and positively related to performance.

In this section, I investigated possible interpretations of cortisol in academic settings and found that multiple interpretations of cortisol are possible for both negatively and non-negatively stereotyped students. Overall, as noted above, the relationship between cortisol and academic performance is complicated due to the multi-faceted and often malleable nature of cortisol. If only cognitive functioning influences academic performance, it is likely that cortisol may be negatively related to performance. However, complicating the issue, cortisol also can represent engagement with the task which can be positively associated with academic performance. Indeed, as noted above, there has been evidence that cortisol can be beneficial for performance.
Additionally, the fact that psychological meanings of cortisol can be altered by social psychological factors further complicates the relation between cortisol and academic performance. In my current study, I aim to untangle these ambiguities. In the section below, I will describe the current study in more detail and ways that I am planning to untangle the ambiguity of what cortisol may represent in the academic setting.

**Current Study**

The goal of the current study is to identify students’ cortisol response patterns and their association with students’ academic performance in a gateway college science course. Further, I investigate whether cortisol and its association with performance varies by race/ethnicity.

First, my research will examine students’ cortisol responses within and across classes. If cortisol represents stress, this could inform us of whether students experienced stress during the class and whether students perceived some classes to be particularly stressful relative to others. Cortisol was measured by collecting saliva samples at the start and end of each class across four classes during the first half of the fall semester. If students found classroom experiences to be stressful, I expected that cortisol levels would rise between the start and end of each class. However, if students were more stressed from anticipating a challenging class, cortisol level assessed at the beginning of the class might be higher than cortisol level at the end of the class (Preuß et al., 2010). Furthermore, by measuring students’ cortisol levels across four different
classes, I can identify whether students perceived certain classes to be more stressful than others.¹

My research broadens previous literature by examining the association between classroom cortisol levels and academic performance in a gateway college class. Based on past literature, I suggest that there can be three possible relationships between cortisol and academic performance. First, students’ cortisol levels could be negatively associated with students’ academic performance. As noted above, elevated cortisol has been found to be associated with impaired cognitive functioning (e.g., working memory, executive functioning, memory retrieval), which may hinder students’ learning abilities. Second, cortisol levels could have no relation to academic performance. This would be in line with past literature that suggests that cortisol levels assessed in the academic setting are not associated with students’ performance. Last, cortisol levels could be positively related to academic performance. Moderately elevated cortisol has been associated with better cognitive functioning, which can be positively associated with academic performance. Furthermore, elevated cortisol can indicate engagement in an academic setting, which in turn may be positively related to academic performance.

Additionally, I explore whether 1) the pattern of cortisol responses and 2) the association of cortisol with academic performance may differ as a function of race/ethnicity. As noted above, negatively stereotyped students may perceive the classroom to be more stressful. In such a case,

¹ This type of analysis of cortisol over days is unusual but should nonetheless be appropriate because saliva samples were collected at consistent class times and because I control for time since awakening in analyses described below.
negatively stereotyped students can have higher average levels of cortisol and experience a greater increase in cortisol during the class than non-negatively stereotyped students. Similarly, longitudinal patterns of cortisol may differ by race/ethnicity. For instance, negatively stereotyped students might exhibit a sharper increase in cortisol approaching the exam period, potentially when they experience the greatest amount of social identity threat regarding their intellectual ability.

Furthermore, I investigate the possibility that the relation between cortisol and performance can be moderated by race/ethnicity. As noted above, cortisol may represent negative stress and maladaptive physiological arousal for negatively stereotyped students, and thus be negatively associated with performance. Interestingly, the opposite pattern is also possible. Negatively stereotyped students may cope with stereotype threat in an academic setting differently than non-negatively stereotyped students. To cope with negative stereotypes, negatively stereotyped students may disengage from their classes, which would lead to lower cortisol levels and worse performance. Alternatively, negatively stereotyped students could strive to overcome stereotypes, which would lead to higher engagement, increased cortisol, and, ultimately, better performance. For non-negatively stereotyped students, cortisol can represent negative stress and be negatively related to performance. As this class was known to be difficult and important, it is possible that cortisol may represent negative stress for non-negatively stereotyped students. On the other hand, as non-negatively stereotyped students may perceive the classroom setting to be less threatening, cortisol to them may represent engagement or adaptive physiological responses. If so, it is possible that cortisol will be positively associated with academic performance for non-negatively stereotyped students.
Overall, I believe the greatest asset of our research is its examination of students’ cortisol in a real-life context with pivotal consequences. This class was important for many students, including the pre-medical students required to take it, as grades received in this course could determine their eligibility to continue in their selected major.
CHAPTER 3: Method

The research presented here is taken from a larger project that examined students’ performance and experiences in Introduction to Molecular and Cellular Biology, a gateway college class for students in many STEM majors, such as pre-med, biology and chemistry. Especially, as this class was a required course for pre-med majors, the majority of enrolled students (more than 75%) were planning to apply to medical school in the future.

One reason that I conducted my research in this class is because this class is known to be particularly challenging. As will be noted below, the majority of students perceived this class to be extremely difficult. As such, this class would be ideal for investigating students’ cortisol responses. Further, it is important to identify how stress responses in such a challenging and important class may influence students’ performance. Also, I selected this course because of its overall size. This class was one of the STEM courses that had the largest number of students ($N = 552$), and thus granted me a big enough sample size for my analyses. Additionally, I selected this class as it had the highest probability of providing me with a large-enough sample of negatively stereotyped students. As noted above, African American, Hispanic, and Native Americans are known to be underrepresented in STEM domains (Syed et al., 2011), and thus it was important to select a class that was large in its size as it would increase our probability of recruiting underrepresented ethnic minorities students.

Participants

A total of 552 students were enrolled in the Introduction to Molecular and Cellular Biology class at Columbia University in the fall of 2014. Of these, 328 students (59%) consented
to participate in the study.² Among the 328 students who consented, 53 did not submit any saliva samples, and were thus excluded from the analyses presented here.³

Of the remaining 275 consenting participants, two were excluded because they did not take any exams and thus did not have a performance outcome. Consistent with recommended guidelines (Granger, Shirtcliff, Booth, Kivlighan, & Schwartz, 2004), another two participants were excluded because their average cortisol levels were above 3 standard deviations from the sample mean, 5 and 14 standard deviations, respectively. After exclusions, the final sample size was 271. Most participants (77.4%) submitted at least 75% of the requested saliva samples (i.e., ≥ 6 out of 8 samples total). Further, most participants (87.5%) completed all four exams (see Table 1).

---

² As I did not have any demographic information on non-consenting students, I could not identify whether consenting and non-consenting students differed in their demographics.

³ Of these 53, about a quarter (23%, n = 12) dropped the class, and thus, were not part of the study anymore. Most of the others who did not provide saliva samples (53%, n = 28) did not complete other parts of the study as well, and thus, the absence of saliva was part of a larger pattern of incomplete participation. The others (24%, n = 13) had missing saliva samples for unknown reasons.
Table 1. Number of Saliva Samples Submitted and Exams Taken

<table>
<thead>
<tr>
<th>Measure</th>
<th>Number of participants (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(N = 271)</td>
</tr>
<tr>
<td>Number of submitted saliva sample</td>
<td></td>
</tr>
<tr>
<td>1-2</td>
<td>26 (9.6%)</td>
</tr>
<tr>
<td>3-4</td>
<td>33 (12.1%)</td>
</tr>
<tr>
<td>5-6</td>
<td>62 (22.8%)</td>
</tr>
<tr>
<td>7-8</td>
<td>150 (55.3%)</td>
</tr>
<tr>
<td>Number of exams taken</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>11 (4.1%)</td>
</tr>
<tr>
<td>2</td>
<td>15 (5.5%)</td>
</tr>
<tr>
<td>3</td>
<td>8 (3.0%)</td>
</tr>
<tr>
<td>4</td>
<td>237 (87.5%)</td>
</tr>
</tbody>
</table>

Participants were 21 years old on average (M = 21.18, SD = 3.91; Range = 18 – 48), and primarily female (65.3%). The class was diverse with respect to race/ethnicity. From a list provided by the research team, the largest self-identified racial/ethnic group was White (34.7%), followed by Asian (27.3%), multi-racial (18.1%), Hispanic (9.2%), Black (7.7 %,), Native American (0.7%) and other (2.2%). As noted above, students who belong to a negatively stereotyped racial/ethnic group (i.e., Black, Hispanic, Native American) may have qualitatively different experiences in the classroom compared to those who belong to a non-negatively stereotyped racial/ethnic group (i.e., White, Asian) (Major & O’Brien, 2005; Schmader et al., 2008; Steele et al., 2002). For this reason, in the current study, I categorized students based on whether they belonged to a negatively stereotyped (N = 77) or non-negatively stereotyped (N = 194) racial/ethnic group. As it will be described in more details below, for students who

---

4 Three students did not answer demographic questions and were identified through phenotypical traits.
identified themselves as multi-racial, I decided their group assignment based on the following set of rules. Those who indicated that at least one of their racial/ethnic groups was Black, Latino, or Native American were categorized as negatively stereotyped. Those who did not identify themselves as Black, Latino, and Native American but identified themselves as Asian and/or White were assigned to the non-negatively stereotyped group.

Participants tended to come from families with high levels of education and income (see Table 2). Non-negatively stereotyped students were more likely to come from families who had high education and income compared to negatively stereotyped students. Almost all the students were either in their second (54.6%) or third (15.9%) year of college, with the rest in their first (3.3%), fourth (2.2%), or fifth (0.4%) year. The remaining 23.6% were postbaccalaureate (post-bac) students. These are students who already have an undergraduate degree but are taking courses to bolster their academic credentials before attending professional (typically medical school) or graduate school.

As noted earlier in this section, this class was perceived as both difficult and important. Prior to the class starting or in the first weeks of the semester, over half of the participants (51%) reported expecting the class to be very difficult, but the vast majority (87%) strongly agreed that it was important for them to do well in the class. Over half of participating students (54%) reported that they were “definitely” planning to apply to medical school in the future and most (58%) strongly agreed that getting into medical school was important to them (see Table 3).
Table 2. Demographic Information of Participants

<table>
<thead>
<tr>
<th>Measure</th>
<th>Total (%) (N = 271)</th>
<th>Non negatively stereotyped students (N = 194)</th>
<th>Negatively stereotyped students (N = 77)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M (SD)</td>
<td>21.18 (3.91)</td>
<td>21.45 (4.20)</td>
<td>20.51 (2.96)</td>
</tr>
<tr>
<td><strong>Gender</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>94 (34.7%)</td>
<td>69 (35.6%)</td>
<td>25 (32.5%)</td>
</tr>
<tr>
<td>Female</td>
<td>177 (65.3%)</td>
<td>125 (64.4%)</td>
<td>52 (67.5%)</td>
</tr>
<tr>
<td><strong>Parents’ education level</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than high school</td>
<td>8 (30%)</td>
<td>1 (0.5%) -</td>
<td>7 (9.1%) +</td>
</tr>
<tr>
<td>High school diploma</td>
<td>28 (10.3%)</td>
<td>16 (8.2%) -</td>
<td>12 (15.6%) +</td>
</tr>
<tr>
<td>Two-year college degree</td>
<td>16 (5.9%)</td>
<td>8 (4.1%) -</td>
<td>8 (10.4%) +</td>
</tr>
<tr>
<td>Four-year college degree</td>
<td>45 (16.6%)</td>
<td>31 (16.0%) -</td>
<td>14 (18.2%) +</td>
</tr>
<tr>
<td>Master’s degree</td>
<td>64 (23.6%)</td>
<td>45 (23.2%) -</td>
<td>19 (24.7%) +</td>
</tr>
<tr>
<td>Ph.D. or professional degree</td>
<td>110 (40.6%)</td>
<td>93 (47.9%) +</td>
<td>17 (22.1%) -</td>
</tr>
<tr>
<td>(MD, MBA, JD)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Self-reported parent’s</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>household income</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Below $30,000</td>
<td>21 (7.7%)</td>
<td>8 (4.1%) -</td>
<td>13 (16.9%) +</td>
</tr>
<tr>
<td>$30,001-$50,000</td>
<td>29 (10.7%)</td>
<td>16 (8.2%) -</td>
<td>13 (16.9%) +</td>
</tr>
<tr>
<td>$50,001-$90,000</td>
<td>47 (17.3%)</td>
<td>28 (14.4%) -</td>
<td>19 (24.7%) +</td>
</tr>
<tr>
<td>$90,001-$150,000</td>
<td>60 (22.2%)</td>
<td>43 (22.2%) -</td>
<td>17 (22.1%)</td>
</tr>
<tr>
<td>More than $150,000</td>
<td>78 (28.8%)</td>
<td>66 (34.0%) +</td>
<td>12 (15.6%) -</td>
</tr>
<tr>
<td>Unsure</td>
<td>36 (13.3%)</td>
<td>33 (17.0%) +</td>
<td>3 (3.9%) -</td>
</tr>
<tr>
<td><strong>Self-reported year at</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Columbia</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1st year</td>
<td>9 (3.3%)</td>
<td>8 (4.1%)</td>
<td>1 (1.3%)</td>
</tr>
<tr>
<td>2nd year</td>
<td>148 (54.6%)</td>
<td>103 (53.1%)</td>
<td>45 (58.4%)</td>
</tr>
<tr>
<td>3rd year</td>
<td>43 (15.9%)</td>
<td>30 (15.5%)</td>
<td>13 (16.9%)</td>
</tr>
<tr>
<td>4th year</td>
<td>6 (2.2%)</td>
<td>4 (2.1%)</td>
<td>2 (2.6%)</td>
</tr>
<tr>
<td>5th year</td>
<td>1 (0.4%)</td>
<td>0 (0.0%)</td>
<td>1 (1.3%)</td>
</tr>
<tr>
<td>Postbac</td>
<td>64 (23.6%)</td>
<td>49 (25.3%)</td>
<td>15 (19.5%)</td>
</tr>
</tbody>
</table>

*Note. Group differences in age were tested by one-way analysis of variance. Group differences in categorical variables were tested with chi-squared tests (+/- signs indicate that cell proportions were bigger/smaller than expected, based on adjusted standardized residuals).*

*p ≤ .05. **p ≤ .01. ***p ≤ .001.
### Table 3. How Students Perceived Class and Attending Medical School

<table>
<thead>
<tr>
<th>Measure</th>
<th>Number of participants (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>((N = 271))</td>
</tr>
<tr>
<td><strong>Expected difficulty of the course</strong></td>
<td></td>
</tr>
<tr>
<td>(How difficult do you expect this course to be?)</td>
<td></td>
</tr>
<tr>
<td>Very easy</td>
<td>1 (0.4%)</td>
</tr>
<tr>
<td>Somewhat easy</td>
<td>5 (1.8%)</td>
</tr>
<tr>
<td>Neither easy nor difficult</td>
<td>15 (5.5%)</td>
</tr>
<tr>
<td>Somewhat difficult</td>
<td>111 (41.0%)</td>
</tr>
<tr>
<td>Very difficult</td>
<td>139 (51.3%)</td>
</tr>
<tr>
<td><strong>Importance of course to participant</strong></td>
<td></td>
</tr>
<tr>
<td>(It is important to me to do well in this course)</td>
<td></td>
</tr>
<tr>
<td>Strongly Disagree</td>
<td>3 (1.1%)</td>
</tr>
<tr>
<td>Disagree</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Neither Agree nor Disagree</td>
<td>1 (0.4%)</td>
</tr>
<tr>
<td>Agree</td>
<td>32 (11.8%)</td>
</tr>
<tr>
<td>Strongly Agree</td>
<td>235 (86.7%)</td>
</tr>
<tr>
<td><strong>Whether participant was planning on applying to medical school at that time</strong></td>
<td></td>
</tr>
<tr>
<td>(Are you planning on applying to medical school?)</td>
<td></td>
</tr>
<tr>
<td>Definitely No</td>
<td>18 (6.6%)</td>
</tr>
<tr>
<td>Probably No</td>
<td>26 (9.6%)</td>
</tr>
<tr>
<td>Undecided</td>
<td>23 (8.5%)</td>
</tr>
<tr>
<td>Probably Yes</td>
<td>57 (21%)</td>
</tr>
<tr>
<td>Definitely Yes</td>
<td>147 (54.2%)</td>
</tr>
<tr>
<td><strong>Importance of medical school to participants</strong></td>
<td></td>
</tr>
<tr>
<td>(It is important to me to get into medical school)</td>
<td></td>
</tr>
<tr>
<td>Strongly Disagree</td>
<td>12 (4.4%)</td>
</tr>
<tr>
<td>Disagree</td>
<td>14 (5.2%)</td>
</tr>
<tr>
<td>Neither Agree nor Disagree</td>
<td>32 (11.8%)</td>
</tr>
<tr>
<td>Agree</td>
<td>55 (20.3%)</td>
</tr>
<tr>
<td>Strongly Agree</td>
<td>158 (58.3%)</td>
</tr>
</tbody>
</table>
Procedure

Participants were recruited first with an email sent just before the beginning of the semester and then again in class during the first week of the semester. The first wave of recruitment began a week prior to the start of the class in the fall. With assistance from the course instructor, we sent an email to all students that described the study as aiming to improve students’ experiences in university biological sciences classes, and invited them to participate. The email informed students that they could begin their participation by completing a set of questionnaires online (baseline measures), accessible by a link included in the recruitment email. Upon opening the link, participants were presented with an online consent form and a set of baseline measures (described below).

The second wave of recruitment took place during the first week of the semester, when teams of research assistants distributed consent forms in class while one of the project’s principal investigators made an announcement about the study. Students who had not already consented to participate online were encouraged to sign the forms and return them in class. These students were subsequently sent an email with a link to the baseline measures without the online consent. All consenting participants were asked to complete and sign a transcript release form for obtaining registrar records at this time as well. Most participants who consented to participate in the study (95%) signed a transcript release form.\textsuperscript{5}

\textsuperscript{5} This study was part of a larger project, in which participants were randomly assigned to complete a values-affirmation or control writing exercise in their recitation section in Week 3 of the semester. Results reported here were not moderated by condition and thus, the affirmation intervention is not described further in the main text.
Saliva samples were collected in a single class during Weeks 3, 5, 6, and 7 of the 16-week semester (see Figure 1). These weeks were selected to incorporate different experiences that students might perceive around the first midterm of the semester. Class during Week 3 (Class 1) represents an early period of the semester before students’ comprehension had been tested. Week 5 (Class 2) samples were collected one class before the midterm exam. Week 6 (Class 3) was right after the midterm exam, but before students had received their grades. Weeks 5 and 6 incorporated students’ experience during the exam period. The last class when saliva was collected (Week 6 - 7) (Class 4) was right after students had received their grades on the first midterm.\(^6\) For each day of class, participants could choose to attend either a morning session (10:10 – 11:25 AM) or an afternoon session (4:10 – 5:25 PM). As will be noted below, to eliminate any spurious effect rooting from the time of the day, time since awakening at the moment of saliva collection was controlled in all analyses.

\(^6\) There is variability in when the last saliva samples were collected because the post-exam class was rescheduled to two different days due to an overlap with a religious holiday on the scheduled day of class. These classes straddled Weeks 6 and 7. The time of the morning and evening lectures on these revised dates was unchanged.

However, the intervention did moderate the longitudinal pattern of cortisol. See Appendix A for these results and more detailed information about the intervention.
Within each class, participants submitted saliva samples at the beginning and end of the class, yielding a total of eight possible samples across the four classes. This procedure allows both within-class and between-class analysis of cortisol. Saliva assessed at the beginning of the class can potentially provide information about students’ anticipation of challenging stimuli (Preuß et al., 2010), whereas change in saliva from the beginning to the end of the class can provide information about how students respond to the lecture.

On the day before each saliva sample collection, the research team sent an email to participants that reminded them of the saliva collection the next day and provided instructions for picking up the saliva collection materials on their way into class. The email also asked participants to refrain from activities that may influence their cortisol level.7 On days of saliva sample collection, a team of researchers waited in front of the classroom before the class started

---

7 In particular, participants were asked to avoid the following activities 60-minutes before the class: Brushing teeth, eating or drinking (particularly avoiding caffeinated drinks and dairy products), and chewing gum. They were also asked to avoid alcohol, tobacco, and marijuana in the 12 hours before class.
and students picked up plastic bags as they entered. These plastic bags had been prepared in advance by the research team and included two saliva sample tubes and a half-sheet of paper with instructions for providing the saliva samples and a short set of questions to assess potential behaviors or activities that have been shown to affect saliva (e.g., time of waking, medications taken). A sticker was attached outside of the plastic bag where students could write their school ID for tracking purposes. For the first class when saliva samples were collected, the researchers also made an announcement to the class that explained the saliva collection procedures.

Participants were instructed to provide the first saliva sample as soon as they sat down in class, and the second right before leaving class. To distinguish the samples, stickers were attached on top of saliva tubes. An orange sticker with the number 1 written on it was attached on the saliva tube that was designated for the beginning of the class, and a blue sticker with the number 2 written on it was attached to the saliva tube for the end of the class. For each saliva sample, participants were asked to remove a cotton swab from the sample tube and to chew it for one minute and to keep it in their mouth for another minute without chewing, consistent with recommended guidelines (Salimetrics & SalivaBio, 2015). Then, participants put the cotton swab back into the tube and replaced the cap. Once done with sample collection, participants were asked to place tubes and the sheet with the questions answered back into the plastic bag. At the end of class, the team of researchers waited outside of the classroom to collect the plastic bags from students.

Immediately after the class, saliva samples were brought back into the lab and the research team coded whose were submitted based on students’ school ID, which were matched to students’ study ID number (PPID). The research team also noted the class saliva samples were
collected from and whether students submitted their saliva questionnaire with their sample. Additionally, the research team noted the same information on the saliva questionnaires to verify who had completed each saliva questionnaire. Following this process, a label was put on each tube indicating PPID number, the class when saliva was collected, and the time within the class when saliva was collected (i.e., at the start or at the end of the class).\textsuperscript{8} Immediately after this process, student’s school ID was deidentified from the samples by discarding the plastic bags that had students’ school ID written on it. Then, the saliva samples were immediately frozen at or below \(-25^\circ C\) (Garde & Hansen, 2005) until they were assayed.

Students took four exams throughout the semester in Weeks 5, 8, 12, and 16 (see Figure 1). Students’ scores were provided by the instructor at the end of the semester, along with students’ final score in the class (a score assigned to students at the end of the semester by the instructor based on their exam scores). At the end of the semester, students were asked to complete post-class measures that were similar in format and content to the baseline measures. As results from the post-class measures will not be presented in the current study, they are not discussed further.

**Primary Measures**

**Cortisol.** Saliva samples were delivered to the Biomarker Core Lab at the Pennsylvania

\textsuperscript{8} For 12 samples, the research team potentially mislabeled tubes as there were discrepancies between information coded right after the class (e.g., PPID number, class samples were submitted) and information on the label attached to the tube. More details on how we dealt with these samples are presented in Appendix B. My findings do not change in its significance and direction after excluding participants who had discrepancy in information from the label and from the information coded after class.
State University in an insulated box with temperature below freezing to prevent samples from melting. Once samples arrived at the lab, they were immediately transferred to a -70°C freezer until the day of immunoassay. The samples were removed from the freezer on the morning of analysis, allowed to thaw at room temperature (approximately 20°C) and centrifuged (5000 x g) for 10 minutes. Samples were assayed using a highly sensitive salivary cortisol immunoassay (Salimetrics, State College, PA). Participants submitted a total of 1,753 saliva samples, but eight samples were not able to be assayed due to insufficient saliva in the sample, leaving 1,745 samples from 271 participants. Our samples had acceptable averaged intra- and inter-assay coefficients (5.7% and 5.8%).

Race/Ethnicity. Students described their race/ethnicity by using both a free response item and a checklist of pre-defined racial categories. First, students were asked to complete a free response item describing in their own words their race/ethnicity. Then, students were asked to identify their race from a list of pre-defined categories (choosing multiple options was possible). The racial categories were White, Black, Asian, Latino, Native American, Multi-racial, and Other. If students selected Multi-racial or Other, students could also describe additional information about their race/ethnicity. I defined students’ race/ethnicity based on following rules.

---

9 Ninety percent of samples were analyzed in singlets and 10% (selected randomly) were analyzed in duplicate (one sample tested twice) for assessing intra-assay coefficients. Duplicate cortisol results were averaged, and averaged values were used in all analyses.

10 Intra-assay and inter-assay coefficients represent the precision of salivary cortisol analysis. Details of how coefficients were calculated are provided in Appendix C.
First, I categorized students’ race/ethnicity based on how students responded to the free response item describing their own race/ethnicity. If students described themselves to be only one race/ethnicity and identified themselves to be White or Asian, I categorized them as non-negatively stereotyped. On the other hand, if students described themselves to be only one race/ethnicity and identified themselves to be Black, Latino, or Native American, I categorized them as negatively stereotyped. Second, if students described themselves to be more than one race/ethnicity, I used the following rule to define their race/ethnicity. If students described themselves to be at least one of Black, Hispanic, and Native American as their racial/ethnic group, they were categorized as negatively stereotyped. If they described themselves as Asian and/or White, but not Black, Hispanic, and Native American, they were categorized as non-negatively stereotyped. Lastly, there were cases where students did not indicate their race/ethnicity on the free response item. In this case, I utilized both the initial checklist of pre-defined racial categories students used to describe their race/ethnicity and the additional information students filled out once they indicated that they are multi-racial. Students were categorized into either the negatively or non-negatively stereotyped group following the same logic noted above.

**Students’ final class score.** Students took four exams in the course. For each exam, the maximum raw score that students could achieve was 100 points. However, each exam varied in its difficulty. To adjust for differences in mean scores across exams, the instructor curved scores. Following this procedure, the mean of each exam was brought into a similar range (see Table 4 for averaged exam scores for each exam before and after the adjustment).
Table 4. Descriptive Statistics of Unadjusted and Adjusted Exam Scores

<table>
<thead>
<tr>
<th>Measure</th>
<th>Number of participants who took exam (Total N = 271)</th>
<th>Unadjusted exam scores Mean (SD)</th>
<th>Adjusted exam scores Mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exam 1</td>
<td>269</td>
<td>59.76 (13.09)</td>
<td>73.76 (13.09)</td>
</tr>
<tr>
<td>Exam 2</td>
<td>261</td>
<td>68.26 (16.44)</td>
<td>70.26 (16.44)</td>
</tr>
<tr>
<td>Exam 3</td>
<td>245</td>
<td>72.59 (13.23)</td>
<td>74.59 (13.23)</td>
</tr>
<tr>
<td>Exam 4</td>
<td>239</td>
<td>75.63 (15.34)</td>
<td>75.63 (15.34)</td>
</tr>
</tbody>
</table>

*Note.* Fourteen points were added to exam 1 (maximum score 114). Two points were added to exam 2 and 3 (maximum score 102). Zero points were added to exam 4 (maximum score 100).

Based on the adjusted exam scores, the instructor calculated students’ final score for the class in two ways and used the higher score of the two when assigning a final letter grade. In the first method, the instructor dropped the lowest of students’ first three exams and then added the remaining two exam scores to the fourth exam score. For instance, using this method, if a student’s exam scores were, in chronological order, 90, 80, 70, and 60, the lowest score of the first three (70) would be dropped and the final score would be 230 (90 + 80 + 60). In the second method, the instructor counted the fourth exam with a 50% weight and the lowest of the first 3 exams with a 50% weight and then added to these the full score of remaining two exams. For instance, using this method, the same exam scores as above would result in a final score of 235 (90 + 80 + 70*0.5 + 60*0.5). Because the second method results in the higher score, this student’s final class score would be 235. If students did not take an exam, the instructor assigned a score of 0 for that exam, and calculated the final score in the same manner. The maximum score students could get for the semester was 317 (i.e., this would be the case if students received the maximum adjusted score for each exam, which were 114, 102, 102, 100; 114 + 102 + 102*50 + 100*50 = 317). On average, students scored 217 (68%) (SD = 54.57; Range = 29 – 295 or 9% -
93%). Students received an A grade if they scored more than 260 (82%) on their final score. The final numeric score for the class is the dependent variable in the current research.\textsuperscript{11} I used this rather than students’ final letter grades because 29 students who had taken some of the exams dropped out of the class, and thus, did not receive a final letter grade. I believed it was important to include these participants, as one might imagine that these students were struggling with the material, which might reveal itself in the association of cortisol with performance.

**Students’ attrition from the course.** Whether students dropped out of the course was identified from participants’ transcripts. If students dropped out from the class, either the class name was removed from the transcript or a brief note indicating that student dropped the course was present on the transcript. However, transcript data was missing for 10 students as they did not consent to the release of the transcript. For these students, I was able to identify whether they dropped the course from the instructor’s gradebook, where the instructor indicated students’ grade and whether they dropped the course.

**Control Measures**

**Gender.** As gender may influence cortisol responses to stressful situations (Hostinar et al., 2014; Liu et al., 2017), and because previous research has found gender differences in performance in STEM classes (Brewe et al., 2010), gender is included as a control variable.

\textsuperscript{11} One could imagine using a simple average of available exam scores as another approach to assessing performance. Using this approach, the direction and significance of results do not change. See Appendix D.
Participants could identify their gender as Male, Female or Other. If they reported their gender as Other, they could describe their gender.\textsuperscript{12}

\textbf{Socioeconomic status (SES).} As socioeconomic status (SES) can potentially influence cortisol patterns (Chen, Sheldon, & Miller, 2009; Stephens, Townsend, Markus, & Phillips, 2012) and performance (Reardon, 2011), I included SES as a covariate. Additionally, as noted above, I observed a difference in SES between negatively and non-negatively stereotyped students. This further suggests that SES needs to be controlled as SES can confound the effect of race/ethnicity on cortisol, academic performance, and their relationship. I calculated SES based on three items: participant’s self-reported parental average household income, highest level of education of either parent, and perceived social status relative to other people in the United States. These variables were standardized and averaged to form a single measure of SES. When missing data were present, the average is based on the subset of available items.

Participants were asked to report their parents’ average household income on a scale from (1 = $0-$30,000, 2 = $30,001-$50,000, 3 = $50,001-$70,000, 4 = $70,001-$90,000, 5 = $90,001-$110,000, 6 = $110,001-$130,000, 7 = $130,001-$150,000, 8 = $150,001-$250,000, 9 = $250,000+, 10 = I do not know). For the 10 participants who did not know, this item was coded as missing.

Participants were also asked to report the highest level of education of either parent on a scale from 1 to 6 (1 = Less than high school, 2 = High school diploma, 3 = Two-year college

\textsuperscript{12} Four participants identified themselves as Other and described their gender as “transgender”, “agender or genderfluid”, and “genderqueer”. To avoid creating an additional gender category for this small number of individuals, their gender was coded based on their phenotypical presentation as female.
degree, 4 = Four-year college degree, 5 = Master’s degree, 6 = PhD or professional degree (MD, MBA, JD)).

To assess subjective SES, participants were presented with a picture of a ladder with 10 rungs. Participants were asked to place themselves on the ladder based on their income, education, and occupation compared to others in the United States. The score range was from 1 (lowest on the ladder) to 10 (highest on the ladder) (Adler, Epel, Castellazzo, & Ickovics, 2000). Lower numbers indicate that individuals perceived themselves to be lower in status than others in the United States with higher scores.

Prior performance. I wanted to control for the influence of students’ prior knowledge and achievement in biology and chemistry, which could affect performance in the current course, their assessed cortisol level, and the association of cortisol with performance. Students’ prior performance from other classes was collected from official transcripts. I used students’ prior averaged GPA from biology and chemistry classes as a covariate in analyses presented below. I selected GPA from these classes because performance from biology and chemistry classes potentially is more related to students’ performance in Introduction to Molecular and Cellular Biology than grades from other classes would be. Students’ prior performance could range from 0 (F) to 4.33 (A+).

Variables that may influence cortisol level. As cortisol is known to follow a circadian rhythm, such that it peaks in the morning shortly after awakening and declines throughout the day (Debono et al., 2009; Krieger, Allen, Rizzo, & Krieger, 1971), time since awakening is important for ruling out potential spurious effects due to the time of day on cortisol levels. I used participants’ report of their time of awakening assessed in the saliva questionnaire (i.e., “what
time did you wake up today?”) to calculate time since awakening at the time of saliva sample collection.

Also assessed on the in-class questionnaire that accompanied saliva sample collection was whether participants were taking medication such as sedatives or barbiturates, antibiotics, anti-inflammatory medicines, and steroids that can potentially alter cortisol levels (Granger, Hibel, Fortunato, & Kapelewski, 2009). Participants were given a checklist of the aforementioned medications and asked to check if they have taken any in the 24 hours before the saliva collection. We also assessed whether participants brushed their teeth in the hour before saliva collection with a yes/no question because brushing teeth can cause bleeding in the gums and blood can contaminate saliva and elevate salivary cortisol level (Malamud & Rodriguez-Chavez, 2011).

Other health questionnaires potentially associated with cortisol (Fukuda & Morimoto, 2001; Gibson et al., 1999; Jacks et al., 2002) were also assessed at the time of saliva collection, but they are not controlled in the models below, as in preliminary analyses they were not associated with students’ cortisol levels. Variables that were assessed but not incorporated in the model include consumption of caffeinated and/or alcoholic drinks, smoking cigarettes and/or marijuana, having recently eaten certain types of food (i.e., a major meal, sour food, food with a lot of sugar, dairy products), having recently chewed gum, and having recently exercised.
CHAPTER 4: Results

In this section, I will first present my coding strategies and how I handled missing data. Second, I will present preliminary data analyses showing the intercorrelations of my primary variables of interest (e.g., cortisol, academic performance, race/ethnicity). Third, I will present longitudinal analyses investigating cortisol patterns within and between classes. Lastly, I will present analyses investigating the relation between cortisol and academic performance.

Data Management

Data coding strategy.

Race/ethnicity was contrast coded (Rosentahal & Rosnow, 1985; Cohen, Cohen, West, & Aiken, 2003), such that non-negatively stereotyped students = -1 and negatively stereotyped students = 1. Gender was contrast coded: males were coded -1 and females were coded 1. I created a dummy variable based on whether participants had taken any medication before the saliva collection (coded as 1) or not (coded as 0). Another dummy variable indicated whether students had brushed their teeth (coded as 1) or not (coded as 0) before saliva collection.

Missing data strategy.

Covariates related to cortisol concentration. A total of 1,745 saliva samples were collected from the 271 participants. Of these, a total of 133 samples from 48 participants were missing data on some or all of the questions that were assessed at the time of saliva collection. As noted earlier, these questions pertained to behaviors that can affect cortisol, which were assessed as potential covariates to include in statistical analyses. In particular, five participants did not submit completed questionnaires with any of their 23 saliva samples. Of the remaining 43 students, all submitted at least one questionnaire, but had missing data on others, totaling 110
saliva samples. Among these 43 students, 40 did not submit saliva questionnaires at all for a total of 104 saliva samples. The other three students submitted their saliva questionnaires but had partially missing data across six saliva samples. One participant had missing data on tooth brushing for two samples. Another participant had missing data on both whether they had brushed their teeth and whether they had taken medication. The third participant had missing data on their time of awakening. Overall, time since awakening was missing from 129 measurement occasions, tooth brushing was missing across 131 measurement occasions, and whether students had taken medication was missing across 129 measurement occasions.

Below, I describe the strategies I implemented to handle these missing data so as to minimize listwise deletion and maximize statistical power. In each case, I replaced missing data with a reasoned approach based on patterns of available data and logical consideration of the variables being assessed. I also considered implementing a multiple imputation strategy. However, upon reflection, this seemed less necessary, given that data were missing on covariates rather than critical predictor or outcome variables, and that manual consideration of the nature of the variables and patterns of data would provide what I considered to be reasonable approximations.

In imputing missing data, I used a separate strategy for the 43 students who had submitted at least one saliva questionnaire and the five students with no saliva questionnaire data at all. For the former group, I inferred missing data based on these students’ answers to other questions and the time of the class they attended. Time of class (i.e., morning or afternoon) is relevant because it relates to time since awakening and the likelihood that students had recently brushed their teeth. To impute missing data, I used data from each student’s other questionnaire
assessed at a different class. Especially, an existing questionnaire from a class that was on the same time of day as the missing class was imputed. Additionally, if a student had submitted questionnaires on multiple classes all on the same time of day as missing class, I prioritized the questionnaire from the most adjacent class from the missing class. For example, a student may have a missing saliva questionnaire on a morning session of Class 2, but still have submitted saliva questionnaires on morning sessions on Class 3 and 4. In this case, the questionnaire from Class 3 was implemented as it is more proximal to missing data from Class 2. On the other hand, if existing questionnaires were not from classes that were on the same time of day as the missing class, existing data from the most adjacent class was used.

For the five participants who did not have any data on saliva questionnaires, I replaced missing values either by using information from their baseline survey or imputation based on the probability observed for missing variables. From the baseline survey, I used information that students had reported about their typical wake time on school days. To calculate time since awakening, I used the time of the class that they attended (noted by the research team on saliva samples) and their typical awakening time they reported on the baseline survey. As the probability of having recently brushed their teeth likely differs based on which section of the class students attended, I imputed missing data on this variable with randomly generated values based on probabilities of brushing teeth in each section of the class (calculated from existing data on brushing teeth). Similarly, missing data on whether students took medication 24 hours before the class were replaced with randomly generated values based on the probability of all students’ having taken medication from our existing data.
**Missing prior performance.** Among our final sample of 271 students, 33 did not have any data on their prior GPA in biology and chemistry classes. This is mainly due to these students (1) not having taken any biology and chemistry classes, (2) not having taken any prior classes, or (3) not having consented to the release of their transcript. Nine of the 33 students had not previously taken chemistry or biology classes, but had a GPA from other classes. For these students, I used their prior GPA from classes other than biology and chemistry for their prior performance. Of remaining 24 students, 14 did not have any prior GPA, as this was their first semester at Columbia. For these students, I used their GPA in other classes from the same semester as the Introduction to Molecular and Cellular Biology class (i.e., fall) excluding the grade for the current class. For the remaining ten students who did not consent to the release of their transcript, I used the mean GPA of prior biology and chemistry classes from the other students. To further reduce the possibility of misspecification, I included an indicator variable for these 10 participants, such that 0 corresponded to students with intact prior performance data and 1 corresponded to those whose prior performance was mean imputed (J. Cohen, Cohen, West, & Aiken, 2013). Although mean imputation is less recommended than newer analytic strategies (Graham, 2009), I deemed it acceptable in this case because of the relatively small number of cases (< 4%) and the fact that prior performance was a covariate rather than a primary predictor variable.

**Preliminary Analysis**

I begin by providing descriptive information about cortisol and academic performance, the primary study variables, and how these are associated with other measured variables (see Table 5).
**Cortisol.** Students’ cortisol concentration, averaged across their 8 samples (or available subset) ranged from 0.04 – 0.99 µg/dL with a grand mean of .24 µg/dL ($SD = .15$). Breaking this down by class time, saliva samples collected during morning and evening classes had average values of 0.31 ($n = 1158, SD = .25$) and 0.12 ($n = 587, SD = .10$), respectively. To put this in context, past literature has suggested that normal cortisol concentration in the morning (8 AM) is around 0.12 – 0.98 µg/dL (Aardal & Holm, 1995) and in the late afternoon (5-7 PM) is 0.06 – 0.37 µg/dL (Hansen, Garde, Christensen, Eller, & Netterstrømt, 2003). Thus, average cortisol values from the morning and evening sections were within a typical range (see Table 6 for descriptive statistics of cortisol at each time point).
Table 5. Means, Standard Deviations, and Correlations of Primary Study Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>$M$</th>
<th>$SD$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Cortisol level$^1$</td>
<td>0.24</td>
<td>0.15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Final scores$^2$</td>
<td>217.00</td>
<td>54.57</td>
<td>.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Gender$^3$</td>
<td>0.31</td>
<td>0.95</td>
<td>.11</td>
<td>-.16$^{**}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Race/ethnicity$^4$</td>
<td>-0.43</td>
<td>0.90</td>
<td>.07</td>
<td>-.27$^{***}$</td>
<td>.03</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. SES$^5$</td>
<td>0.01</td>
<td>0.83</td>
<td>.002</td>
<td>.23$^{***}$</td>
<td>-.01</td>
<td>-.38$^{***}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Prior performance</td>
<td>3.40</td>
<td>0.49</td>
<td>-.06</td>
<td>.49$^{***}$</td>
<td>-.18$^{**}$</td>
<td>-.26$^{***}$</td>
<td>.28$^{***}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Time since awakening$^6$</td>
<td>4.41</td>
<td>2.34</td>
<td>-.58$^{***}$</td>
<td>-.08</td>
<td>-.07</td>
<td>-.07</td>
<td>.01</td>
<td>-.04</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Medication$^7$</td>
<td>0.14</td>
<td>0.29</td>
<td>.14$^*$</td>
<td>-.002</td>
<td>.10</td>
<td>-.03</td>
<td>.14$^*$</td>
<td>-.14$^*$</td>
<td>0.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Brush Teeth$^8$</td>
<td>0.21</td>
<td>0.32</td>
<td>.42$^{***}$</td>
<td>.07</td>
<td>.04</td>
<td>.06</td>
<td>.06</td>
<td>.11</td>
<td>-.44$^{***}$</td>
<td>-.01</td>
<td></td>
</tr>
</tbody>
</table>

*Note. N = 271. $^1$ Raw cortisol concentration averaged across four classes (µg/dL). $^2$ Final score assigned to students by the instructor. $^3$ Male = -1, Female = 1. $^4$ Non-stereotyped racial group = -1, Negatively stereotyped racial group = 1. $^5$ An average of standardized SES variables (parent’s income, parent’s education level, social status). $^6$ Time since awakening averaged across four classes (in hours). $^7$ Whether students took medication 24 hours before saliva collection averaged across four classes. $^8$ Averaged times of whether students brushed teeth an hour before the saliva collection.

$p \leq .05$. $^{**}p \leq .01$. $^{***}p \leq .001$. 

61
As noted in the Method section, a number of covariates were tested for their association with cortisol. As shown in Table 5, brushing teeth an hour before the class, taking medication 24 hours before the saliva collection, and time since awakening were all associated with levels of cortisol. Brushing teeth an hour before the saliva collection, and taking medication 24 hours before the saliva collection were both associated with higher cortisol concentrations. Further, consistent with diurnal patterns of cortisol established by previous research (e.g., Kirschbaum & Hellhammer, 1989), students’ cortisol concentrations were negatively associated with time since awakening.
Demographic factors such as race/ethnicity, SES, and gender were not found to be associated with cortisol levels of students in either zero-order correlation analysis presented in Table 5, or in partial correlation analysis controlling for variables that can influence cortisol concentration (i.e., time since awakening, medication intake, and tooth brushing) (see Table 7). Although lab research has implicated these variables as being associated with cortisol (Fiocco, Joober, & Lupien, 2007; Liu et al., 2017), there have been relatively few studies examining whether students’ cortisol responses differ by demographic variables in an academic field setting. My finding of a null association of cortisol with gender is consistent with a recent study done in a science lab class, which likewise found a null association between cortisol and gender (Minkley, Ringeisen, Josek, & Kärner, 2017). Although past theory has suggested that negatively stereotyped students and individuals from low SES groups may have higher cortisol levels in the academic setting due to perceiving discrimination or social identity threat (Levy et al., 2016; Schmader et al., 2008), we did not find this association. To my knowledge no empirical studies have investigated this association, and these data raise questions about the association of cortisol with race/ethnicity in an academic setting.
Table 7. Partial Correlations of Cortisol Levels with Demographic Variables Controlling for Factors that Influence Cortisol Level (Time Since Awakening, Whether Took Medication, Tooth Brushing)

<table>
<thead>
<tr>
<th>Variable</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Cortisol level</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Gender</td>
<td></td>
<td>.08</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Race/ethnicity</td>
<td></td>
<td>.03</td>
<td>.03</td>
<td></td>
</tr>
<tr>
<td>4. SES</td>
<td></td>
<td>-.03</td>
<td>-.03</td>
<td>-.38***</td>
</tr>
</tbody>
</table>

Note. N = 271. ¹ Raw cortisol concentration averaged across four classes (µg/dL). ² Male = -1, Female = 1. ³ Non-stereotyped racial group = -1, Negatively stereotyped racial group = 1. ⁴ An average of standardized SES variables (parent’s income, parent’s education level, social status). ⁵ Time since awakening averaged across four classes (in hours). ⁶ Whether students took medication 24 hours before saliva collection averaged across four classes. ⁷ Whether students brushed teeth an hour before the saliva collection averaged across four classes.

*p ≤ .05. **p ≤ .01. ***p ≤ .001.

Academic performance. Students’ could score up to 317 points on their final score for exams and their observed average was 217 (68%) (SD = 54.57; Range = 29 - 295). As shown in Table 5, demographic factors such as race/ethnicity, gender and SES were associated with students’ final class score. Students who belonged to a negatively stereotyped racial group had a lower final class score (M = 194.01, SD = 56.52) than students who belonged to a non-negatively stereotyped racial group (M = 226.13, SD = 51.12). Also, male students had higher final scores (M = 228.63, SD = 51.64) than female students (M = 210.83, SD = 55.22). Higher SES was also associated with better performance (r = 0.23, p < .001). Students’ prior performance was also positively associated with students’ final class score (r = 0.49, p < .001).

Longitudinal Data Analysis: Cortisol Pattern Within and Between Classes

In my first set of primary analyses, I examined cortisol as an outcome, testing whether students’ cortisol concentrations varied within and across classes. Within-class analyses provide
a test of whether students’ cortisol concentration at the beginning of the class differed from their cortisol concentration at the end of the class. This provides information about how students physiologically responded within the class. For instance, if students found the class period itself stressful, we might expect cortisol concentrations to be higher at the end of the class than the beginning of class. Between class analyses provide a test of whether students’ cortisol concentration varied across classes. This provides information about whether students perceived some classes as more challenging or stressful than others.

To test these questions, I conducted multilevel regression analyses using the nlme package (Pinheiro, Bates, DebRoy & Sarkar, 2017) available within the R statistical programming software (R Core Team, 2018). In particular, I specified 3-level models, in which students’ within-class cortisol concentrations and their time since awakening were modeled at Level 1. Students’ between-class cortisol concentrations were modeled at Level 2, along with relevant between-class covariates (i.e., whether students took medication or brushed their teeth before the class), as predictors of the Level-1 intercept and within-class slope. At Level 3, between-person variables (i.e., race/ethnicity, SES, gender) were added to predict Level 2 intercepts and slopes.

Intercepts were allowed to vary across all levels, as was the Level-1 coefficient for the within-class difference in cortisol. The coefficient representing the trajectory of cortisol level across class was also allowed to vary across individuals (Level 2).

I tested models using a hierarchical model-building approach in order to separately examine two different research questions. First, in the simple model, I aim to test the main effect of time within class and between classes on cortisol. In the first model, I included demographic
variables only as covariates (i.e., predictors of the intercept). In the second model, I tested whether cortisol patterns differed by demographic factors by including interactions of within-class time and between-class time by demographic variables (i.e., race/ethnicity, SES, gender).

Below, I provide the specification of the complete statistical model. In this model, the subscript \( t \) represents when cortisol was collected in the class (i.e., at the beginning or end of the class), the subscript \( d \) represents the four classes when saliva samples were collected, and the subscript \( i \) represents individuals. Equations written in bold are only included in the second model, which has all interaction terms included.

Level 1: \( \text{Cortisol value}_{tdi} = \beta_{0di} + \beta_{1di}(Time \text{ in class}_{tdi}) + \beta_{2di}(Time \text{ since awakening}_{tdi}) + e_{tdi} \)

\[ \begin{align*}
\beta_{1di} &= \gamma_{10i} + u_{1di} \\
\beta_{2i} &= \gamma_{20i} \\
\end{align*} \]

Level 2: \( \beta_{0di} = \gamma_{00i} + \gamma_{01i}(Class \text{ number}_{di}) + \gamma_{02i}(Class \text{ number}_{di})^2 + \gamma_{03i}(Class \text{ number}_{di})^3 + \\
\gamma_{04i}(Medication_{di}) + \gamma_{05i}(Brush \text{ teeth}_{di}) + u_{0di} \)

\[ \begin{align*}
\gamma_{10i} &= \pi_{100} + \pi_{011}(Race_i) + \pi_{012}(SES_i) + \pi_{013}(Gender_i) + r_{01i} \\
\gamma_{20i} &= \pi_{010} + \pi_{011}(Race_i) + \pi_{012}(SES_i) + \pi_{013}(Gender_i) + r_{02i} \\
\gamma_{02i} &= \pi_{020} + \pi_{021}(Race_i) + \pi_{022}(SES_i) + \pi_{023}(Gender_i) \\
\gamma_{03i} &= \pi_{030} + \pi_{031}(Race_i) + \pi_{032}(SES_i) + \pi_{033}(Gender_i) \\
\gamma_{04i} &= \pi_{040} \\
\gamma_{05i} &= \pi_{050} \\
\gamma_{10i} &= \pi_{100} + \pi_{101}(Race_i) + \pi_{102}(SES_i) + \pi_{103}(Gender_i) + r_{10i}
\end{align*} \]
\[ y_{11t} = \pi_{110} \]
\[ y_{12t} = \pi_{120} \]
\[ y_{13t} = \pi_{130} \]

Time when cortisol was assessed in the class was centered and coded such that -0.5 = cortisol assessed at the beginning of the class and 0.5 = cortisol assessed at the end of the class. Class numbers were centered and coded as -1.5 = first class, -0.5 = second class, 0.5 = third class, and 1.5 = fourth class. I centered both time in the class and class number so that intercept would indicate overall averaged cortisol level throughout the semester. Additionally, as students may have higher levels of cortisol on classes around the time of exam, cortisol patterns can follow a quadratic or cubic pattern. To identify this pattern, class numbers were squared and cubed to examine quadratic and cubic trend of cortisol (Biesanz, Deeb-Sossa, Papadakis, Bollen & Curran, 2004). Time since awakening (coded in hours) and whether individuals took medication and brushed their teeth were grand mean centered. Unlike time since awakening, which is incorporated as a Level 1 covariate, whether students took medication and brushed their teeth were incorporated as Level 2 covariates, as they would not have varied within classes. SES was also grand mean centered. As mentioned above, race/ethnicity and gender were contrast coded. Race/ethnicity, gender and SES were included as Level 3 covariates as they varied across individuals. Additionally, cortisol values were log-transformed, and this successfully reduced skewness (skewness was reduced from 3.93 to -0.03 after transformation) (Schultheiss, Schiepe, & Rawolle, 2012). This transformation satisfied the assumptions of multi-level model that dependent variable has to be normally distributed (Snijders & Bosker, 1999).
**Cortisol pattern within class.** As seen in Figure 2, averaging across classes, students had higher cortisol level at the start of the class compared to the end of the class, $\beta = -0.14$, $t(865) = -15.50$, $p < .001$, 95% CI [-0.16, -0.13] (see Table 8). As individuals’ cortisol levels are known to follow a diurnal pattern which peaks in the morning and declines throughout the day (Kirschbaum & Hellhammer, 1989), some decline in cortisol level might be expected simply due to the passage of time during the class. When interaction terms were included in the model, students’ cortisol patterns within the class was not moderated by race/ethnicity, $\beta = -0.01$, $t(862) = -0.80$, $p = .43$, 95% CI [-0.03, 0.01], gender, $\beta = 0.0004$, $t(862) = 0.04$, $p = .97$, 95% CI [-0.02, 0.02], or SES, $\beta = 0.002$, $t(862) = 0.19$, $p = .85$, 95% CI [-0.02, 0.02].
Figure 2. Cortisol pattern within the class. Values in grey at each time in class are predicted cortisol values (log-transformed) for each student averaged across classes and controlling for relevant covariates, with lines connecting values representing individual students’ average within-class slope. The solid line in blue represents the average within-class slope across students averaged across the classes (See Table 8). In a 75-minute class, saliva samples were collected at the beginning of the class (coded -0.5) and at the end of the class (coded 0.5). Class number was centered so that 0 was the theoretical average class (class number coded as following: Class 1 = -1.5, Class 2 = -0.5, Class 3 = 0.5, Class 4 = 1.5).
Table 8. Three-Level Models of Within-Class and Between-Class Cortisol Pattern in 271 Participants (1745 Cortisol Samples)

<table>
<thead>
<tr>
<th>Fixed effects (Intercepts, Slopes)</th>
<th>Estimate (SE)</th>
<th>95% CI</th>
<th>t</th>
<th>p</th>
<th>df</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Level 1 (within-class)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average cortisol level (intercept)</td>
<td>-0.76*** (0.01)</td>
<td>[-0.78, -0.73]</td>
<td>-53.26</td>
<td>&lt; .001</td>
<td>865</td>
</tr>
<tr>
<td>Time in class¹</td>
<td>-0.14*** (0.01)</td>
<td>[-0.16, -0.13]</td>
<td>-15.50</td>
<td>&lt; .001</td>
<td>865</td>
</tr>
<tr>
<td>Time since awakening²,⁶</td>
<td>-0.06*** (0.003)</td>
<td>[-0.07, -0.06]</td>
<td>-21.76</td>
<td>&lt; .001</td>
<td>865</td>
</tr>
<tr>
<td><strong>Level 2 (between-class)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Class (linear)³</td>
<td>0.03 (0.02)</td>
<td>[-0.01, 0.07]</td>
<td>1.64</td>
<td>.10</td>
<td>602</td>
</tr>
<tr>
<td>Class (quadratic)⁴</td>
<td>0.01 (0.01)</td>
<td>[-0.003, 0.02]</td>
<td>1.47</td>
<td>.14</td>
<td>602</td>
</tr>
<tr>
<td>Class (cubic)⁵</td>
<td>-0.03*** (0.01)</td>
<td>[-0.05, -0.02]</td>
<td>-3.83</td>
<td>&lt; .001</td>
<td>602</td>
</tr>
<tr>
<td>Medication⁶</td>
<td>0.02 (0.02)</td>
<td>[-0.02, 0.06]</td>
<td>0.87</td>
<td>.39</td>
<td>602</td>
</tr>
<tr>
<td>Brush teeth⁶</td>
<td>0.06** (0.02)</td>
<td>[0.02, 0.09]</td>
<td>2.89</td>
<td>.004</td>
<td>602</td>
</tr>
<tr>
<td><strong>Level 3 (between-person)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Race/ethnicity⁷</td>
<td>0.01 (0.01)</td>
<td>[-0.01, 0.04]</td>
<td>0.98</td>
<td>.33</td>
<td>267</td>
</tr>
<tr>
<td>Gender⁸</td>
<td>0.01 (0.01)</td>
<td>[-0.01, 0.03]</td>
<td>0.95</td>
<td>.34</td>
<td>267</td>
</tr>
<tr>
<td>SES⁹</td>
<td>0.001 (0.01)</td>
<td>[-0.03, 0.03]</td>
<td>0.07</td>
<td>.94</td>
<td>267</td>
</tr>
</tbody>
</table>

*Note.* This table presents the model without interaction terms between variables.¹ Time when saliva samples were collected in the class coded as following: -0.5 = collected at the start of the class, 0.5 =collected at the end of the class. ² Time since awakening coded in hours. ³ Class numbers were coded to examine linear trend across classes: first class of saliva collection = -1.5, second class of saliva collection = -0.5, third class of saliva collection = 0.5, fourth class of saliva collection = 1.5. ⁴ Class numbers squared: Class 1 = (-1.5)², Class 2 = (-0.5)², Class 3 = (0.5)², Class 4 = (1.5)². ⁵ Class numbers cubed: Class 1 = (-1.5)³, Class 2 = (-0.5)³, Class 3 = (0.5)³, Class 4 = (1.5)³. ⁶ Time since awakening, whether taken medication 24 hours before the class, whether brushed teeth an hour before the class were grand mean centered. ⁷ Non-stereotyped racial group = -1, Negatively stereotyped racial group = 1. ⁸ Male = -1, Female = 1. ⁹ An average of standardized SES variables (parent’s income, parent’s education level, social status).

*p ≤ .05. **p ≤ .01. ***p ≤ .001.
Cortisol pattern between classes. Students’ cortisol level followed a downward cubic slope across the four classes, $\gamma = -0.03$, $t(602) = -3.83$, $p < .001$, 95% CI [-0.05, -0.02] (see Table 8). As Figure 3 shows, students’ cortisol levels were highest during the first class of saliva collection and lowest during the last class of saliva collection. Students’ cortisol pattern across classes did not differ by race/ethnicity, $\gamma = -0.01$, $t(593) = -0.51$, $p = .61$, 95% CI [-0.03, 0.01], gender, $\gamma = -0.002$, $t(593) = -0.18$, $p = .86$, 95% CI [-0.02, 0.02], or SES, $\gamma = -0.01$, $t(593) = -0.98$, $p = .33$, 95% CI [-0.03, 0.01].

Figure 3. Cortisol pattern across classes. Values in grey at each class are predicted cortisol values (log-transformed) for each student averaged within class controlling for adequate covariates, with lines connecting values representing individual students’ average slope across classes. The solid line in blue represents the average slope across classes averaged within class and across students (See Table 8). Class 1 (Week 3) represents an early period of the semester before any exams. Class 2 (Week 5) was the class right before the midterm exam. Class 3 (Week 6) was right after the midterm exam, but before students had received their grades. Class 4 (week 6-7) was right after students had
received their grades on the first midterm. Time within class was centered (start of class = -0.5, end of class = 0.5). The dotted line indicates when the first exam took place.

Examining the Association Between Cortisol Level and Academic Performance

**Data analysis plan.** Next, I investigated the association between students’ averaged cortisol level across the four classes and students’ final score assigned by the instructor. Rather than using cortisol levels from each class, I used an average of cortisol levels, as it may comprehensively and stably represent students’ cortisol response to the course throughout the semester.

To test the overall association between cortisol and students’ final scores, in *Step 1* I regressed students final scores for the class on averaged cortisol with no additional predictor variables. In *Step 2*, I added relevant covariates to the model (i.e., demographic variables, factors that influence cortisol level, prior performance) to test whether any association between cortisol and final scores changes after taking into account other potential confounding explanations. To test the interaction between cortisol and race/ethnicity in predicting students’ final scores, in *Step 3*, I added to the model the interaction between cortisol and race/ethnicity.

As in the longitudinal analysis of cortisol, I used log-transformed cortisol (our main predictor) in following analyses. For these analyses, factors that can affect cortisol levels (i.e., time since awakening, whether medication was taken within 24 hours, and whether students had brushed their teeth) were also averaged across classes as we are using averaged cortisol level across four classes.
Cortisol and students’ final scores. When examining the bivariate association between cortisol levels and final scores in Step 1, cortisol was not associated with students’ final scores, $b = 6.12, F(1, 269) = 0.20, p = .66, 95\% \text{ CI } [-21.13, 33.38], \eta^2_p = .001$. Moreover, controlling for relevant covariates, in Step 2, students’ average cortisol levels continued to not be associated with their final scores, $b = 9.75, F(1, 260) = 0.32, p = .58, 95\% \text{ CI } [-24.42, 43.92], \eta^2_p = .001$ (see Figure 4). Cortisol concentration did not predict students’ academic performance, consistent with past research that has found a null relation between cortisol and academic performance (Pletzer et al., 2010; Schoofs et al., 2008). Table 9 provides regression coefficients.
Figure 4. Students’ final scores as a function of cortisol levels. The regression line represents predicted scores for each observed value of cortisol controlling for appropriate covariates (all continuous predictors are mean centered and categorical variables are contrast coded). Displayed points correspond to unadjusted observed data. Students who received over 86% received an A and those who received under 46% failed the class; dotted lines indicate cutoff for those grades.

When I included the interaction between cortisol levels and race/ethnicity in Step 3, there was a significant interaction between cortisol and race/ethnicity in predicting students’ final scores, $b = 40.33, F(1, 259) = 8.42, p = .004, 95\% \text{ CI} [12.95, 67.70], \eta_p^2 = .03$. The results of this analysis are presented in Table 9. Simple effects tests revealed that for negatively stereotyped students, cortisol predicted higher final scores, $b = 71.31, F(1, 259) = 6.84, p = .01, 95\% \text{ CI} [17.64, 124.99], \eta_p^2 = .03$. However, among non-stereotyped students, there was not an
association of cortisol with final scores, $b = -9.34$, $F(1, 259) = 0.26$, $p = .61$, 95% CI [-45.44, 26.76], $\eta^2_p = .001$ (see Figure 5).

Figure 5. Students’ final scores as a function of race/ethnicity and cortisol levels. The regression line represents predicted scores for each observed value of cortisol controlling for appropriate covariates (all continuous predictors are mean centered and categorical variables are contrast coded). Displayed points correspond to unadjusted observed data. Students who received over 86% got an A and below 46% failed from the class; dotted lines indicate cutline for those grades.
### Table 9. Regression Coefficients Predicting Final Scores Step 1-3

<table>
<thead>
<tr>
<th>Variable</th>
<th>Step 1</th>
<th>Step 2</th>
<th>Step 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>b</td>
<td>se(b)</td>
<td>CI(95%)</td>
</tr>
<tr>
<td>Intercept</td>
<td>217.00***</td>
<td>3.32</td>
<td>210.47 - 223.54</td>
</tr>
<tr>
<td>Gender 3</td>
<td>-4.95</td>
<td>3.07</td>
<td>-11.00 - 1.11</td>
</tr>
<tr>
<td>SES 4</td>
<td>4.03</td>
<td>3.88</td>
<td>-3.62 - 11.68</td>
</tr>
<tr>
<td>Prior performance</td>
<td>47.99***</td>
<td>6.45</td>
<td>35.28 - 60.69</td>
</tr>
<tr>
<td>Time since awakening 5</td>
<td>-1.19</td>
<td>1.76</td>
<td>-4.66 - 2.28</td>
</tr>
<tr>
<td>Medication 6</td>
<td>9.96</td>
<td>10.17</td>
<td>-10.06 - 29.98</td>
</tr>
<tr>
<td>Brush Teeth 7</td>
<td>-2.18</td>
<td>10.43</td>
<td>-22.72 - 18.36</td>
</tr>
<tr>
<td>Missing prior performance 8</td>
<td>20.02</td>
<td>15.38</td>
<td>-10.27 - 50.31</td>
</tr>
<tr>
<td>Cortisol X Race/ethnicity</td>
<td>20.02</td>
<td>15.38</td>
<td>-10.27 - 50.31</td>
</tr>
</tbody>
</table>

N = 271  \[R^2 = .001\]  \[R^2 = .282***\]  \[R^2 = .304***\]

Note. \(R^2\) change between steps significant at \(p < .001\). \(^1\) Log-transformed cortisol level averaged across four classes and mean centered. \(^2\) Non-stereotyped racial group = -1, Negatively stereotyped racial group = 1. \(^3\) Male = -1, Female = 1. \(^4\) An average of standardized SES variables (parent’s income, parent’s education level, social status). \(^5\) Time since awakening averaged across four classes (in hours). \(^6\) Whether students’ took medication 24 hours before saliva collection averaged across four classes. \(^7\) Whether students brushed teeth an hour before the saliva collection averaged across four classes. \(^8\) Participants who had prior performance data = 0, Participants who did not have prior performance data due to not consenting to release their transcript = 1.

\(^* p \leq .05. \,** p \leq .01. \,*** p \leq .001.\)
I conducted a residual analysis to test for any possible violation of model assumptions. Results revealed that there were 3 participants who had potentially problematic centered leverage values (data points with leverage values that exceed a value of $3p/n$, where $p$ is the number of parameters including the intercept and $n$ is the sample size) (Hoaglin & Welsch, 1978) and 5 participants who had studentized deleted residuals exceeding $|3|$. There was no outlying value in Cook’s D exceeding 1. Removing these 8 potential outliers did not affect the direction of coefficients or pattern of statistical significance reported above.

**Cortisol and attrition from the class.** I further examined if cortisol predicted students’ attrition from the class and whether any potential association was moderated by race/ethnicity. In *Step 1*, I regressed whether students dropped out on students’ averaged cortisol level. Cortisol levels were not associated with whether students dropped the class, OR = 0.52, Wald $\chi^2 (1, N = 271) = 0.61$, $p = .43$, 95% CI [0.10, 2.64]. Even after including all relevant covariates in *Step 2*, cortisol was not associated with students’ dropping out of the course, OR = 0.53, Wald $\chi^2 (1, N = 271) = 0.23$, $p = .63$, 95% CI [0.04, 7.29] (see Table 8). When incorporating the interaction between cortisol levels and race/ethnicity in *Step 3*, a log-likelihood test revealed that adding the interaction term in Step 3 significantly improved the model from Step 2, $\chi^2 (1, N = 271) = 4.84$, $p = .03$. Simple effects clarified the pattern of results. For negatively stereotyped students, cortisol predicted marginally lower odds of dropping the course, OR = 0.03, Wald $\chi^2 (1, N = 271) = 3.18$, $p = .08$, 95% CI [.001, 1.41]. For non-negatively stereotyped students, cortisol was not significantly associated with dropping the course, OR = 2.28, Wald $\chi^2 (1, N = 271) = 0.30$, $p = .58$, 95% CI [.12, 43.05] (see Figure 6).
Figure 6. Students’ probability of dropping out of the class as a function of race/ethnicity and cortisol levels. Values are the predicted probability of dropping out from the class controlling for appropriate covariates (all continuous predictors are mean centered and categorical variables are contrast coded).

An analysis of residuals revealed 12 participants who had potentially problematic centered leverage values (data points with leverage values that exceed a value of $3p/n$, where $p$ is the number of parameters including the intercept and $n$ is the sample size) (Hoaglin & Welsch, 1978), but none with studentized deleted residual value > $|3|$ or outliers in Cook’s D exceeding 1. Although this suggests little reason to be concerned with outliers, I nonetheless conducted the analysis again excluding the 12 cases with potentially problematic leverage values ($N = 259$). Doing so did not change the pattern of results, but the significance of the interaction between cortisol and race/ethnicity was reduced to marginal significance, $OR = 0.13$, Wald $\chi^2 (1, N = 259) = 2.96$, $p = .09$, 95% CI [.01, 1.34].
Table 10. Regression Coefficients Predicting Attrition Step 1-3

<table>
<thead>
<tr>
<th>Variable</th>
<th>Step 1</th>
<th>Step 2</th>
<th>Step 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Odds ratio</td>
<td>CI(95%)</td>
<td>Odds ratio</td>
</tr>
<tr>
<td>Intercept</td>
<td>0.12</td>
<td></td>
<td>3.02</td>
</tr>
<tr>
<td>Cortisol level&lt;sup&gt;1&lt;/sup&gt;</td>
<td>0.52</td>
<td>0.10</td>
<td>2.64</td>
</tr>
<tr>
<td>Race/ethnicity&lt;sup&gt;2&lt;/sup&gt;</td>
<td></td>
<td></td>
<td>1.46</td>
</tr>
<tr>
<td>Gender&lt;sup&gt;3&lt;/sup&gt;</td>
<td></td>
<td></td>
<td>1.45</td>
</tr>
<tr>
<td>SES&lt;sup&gt;4&lt;/sup&gt;</td>
<td></td>
<td></td>
<td>1.28</td>
</tr>
<tr>
<td>Prior performance</td>
<td></td>
<td></td>
<td>0.38&lt;sup&gt;*&lt;/sup&gt;</td>
</tr>
<tr>
<td>Time since awakening&lt;sup&gt;5&lt;/sup&gt;</td>
<td></td>
<td></td>
<td>1.03</td>
</tr>
<tr>
<td>Medication&lt;sup&gt;6&lt;/sup&gt;</td>
<td></td>
<td></td>
<td>0.29</td>
</tr>
<tr>
<td>Brush Teeth&lt;sup&gt;7&lt;/sup&gt;</td>
<td></td>
<td></td>
<td>0.84</td>
</tr>
<tr>
<td>Missing prior performance&lt;sup&gt;8&lt;/sup&gt;</td>
<td></td>
<td></td>
<td>0.00</td>
</tr>
<tr>
<td>Cortisol X Race/ethnicity</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

N = 271
Nagelkerke<em>pseudo</em> R<sup>2</sup> = .005
Nagelkerke<em>pseudo</em> R<sup>2</sup> = .122<sup>∗</sup>
Nagelkerke<em>pseudo</em> R<sup>2</sup> = .156<sup>∗</sup>

<sup>Note</sup>. 1 Log-transformed cortisol level averaged across four classes and mean centered. 2 Non-stereotyped racial group = -1, Negatively stereotyped racial group = 1. 3 Male = -1, Female = 1. 4 An average of standardized SES variables (parent’s income, parent’s education level, social status). 5 Time since awakening averaged across four classes (in hours). 6 Whether students’ took medication 24 hours before saliva collection averaged across four classes. 7 Averaged times of whether students brushed teeth an hour before the saliva collection. 8 Participants who had prior performance data = 0, Participants who did not have prior performance data due to not consenting to release their transcript = 1.

<sup>*</sup>p ≤ .05. <sup>∗</sup>p ≤ .01. <sup>∗∗</sup>p ≤ .001.
Chapter 5: Discussion

Using a longitudinal field study design in a challenging gateway STEM course, Introduction to Molecular and Cellular Biology, my research aimed to accomplish three goals: 1) to investigate students’ cortisol responses both within and across classes, 2) to examine the relationship between students’ cortisol levels and academic performance, and 3) to identify whether cortisol response patterns and the relationship between cortisol and academic performance differ by race/ethnicity. Results revealed that students had higher levels of cortisol at the start of class than at the end. Furthermore, students’ cortisol levels were lower during the examination period than during the introductory weeks of the course. Students’ overall levels of cortisol, their cortisol pattern within the class, and their longitudinal pattern across classes did not vary as a function of race/ethnicity. I found no relation between average levels of cortisol and students’ performance, except when moderated by race/ethnicity. For negatively stereotyped students, cortisol levels were positively associated with exam scores and negatively associated with the probability of dropping out of the course. There was no relation between cortisol levels and academic performance for non-negatively stereotyped students.

Investigation into cortisol response patterns within the class revealed that students had higher cortisol concentrations in the beginning of the class compared to the end of the class. It is possible that students may have experienced a salient cortisol response from anticipating a challenging class, and this is why levels were higher at the beginning. Indeed, past literature looking at cortisol response in the academic setting found similar results: expecting a challenging situation produced a stronger cortisol response than the situation itself (Lacey et al., 2000; Martinek et al., 2003; Ng et al., 2003). For example, past literature found that students had higher concentrations of cortisol at the start of an examination than at the end (Ng et al., 2003). Further,
anticipatory response to the class rather than cortisol response during the class may be a more salient predictor of academic performance. Indeed, Ng and colleagues (2003) found anticipatory cortisol response measured right before the exam predicted students’ performance. In an ancillary analysis, I found that in a regression with averaged cortisol levels from the start and the end of the class as simultaneous predictors (instead of using an overall average), only cortisol levels assessed at the beginning of the class interacted with race/ethnicity to predict academic performance. This result mirrors past literature that suggests that the anticipatory response to stressful stimuli is a predictor of performance (Ng et al., 2003). Ultimately, however, I cannot definitively conclude what cortisol levels at the beginning of the class represent. For example, as an alternative possibility, students may have been more physically active right before class (i.e., as they were getting to the class). As increased levels of physical activity are known to elevate cortisol (Jacks et al., 2002), this could be a simple alternative explanation to an anticipatory stress response. Future research that investigates longitudinal patterns of cortisol would benefit from a more detailed assessment of activities and physiological/psychological states that could be affecting cortisol immediately preceding baseline assessments.

The fact that students’ cortisol concentration decreased during class may, in part, be a function of time, since cortisol levels decline throughout the day (Kirschbaum & Hellhammer, 1989). Also, students’ cortisol levels may decline during the class simply because they are passively sitting rather than doing something active. However, results provide reasonably strong evidence that either the class period itself was not experienced as stressful or that cortisol did not capture the stress students experienced. As students’ performance in this class can determine their eligibility to continue pursuing their selected career (i.e., medical school), stress would seem to be a normative experience, suggesting that the latter explanation may be more accurate.
The cortisol pattern across classes suggests a similar conclusion. Students’ cortisol levels were not elevated during the examination period, and indeed showed a small decline across the classes. My results are contrary to past literature which found that exams elicit a strong cortisol response. For instance, cortisol levels were found to be higher on the day of the exam relative to days when there were no exams (Lacey et al., 2000; Martinek et al., 2003; Merz & Wolf, 2015; Preuß et al., 2010; Schoofs et al., 2008). Anticipating an exam has also been found to be a significant stressor. For example, previous research found that students had higher levels of cortisol on the day right before the exam than the following days (Preuß et al., 2010). However, I found that individuals’ cortisol levels collected during Class 2 (a class that was two days prior to the exam) were lower than cortisol levels measured during Class 1 (a class during an earlier period of the semester). I also initially expected that cortisol levels measured during Class 3 (the first class after the exam) would be elevated relative to cortisol levels on Class 1. As students did not know about their exam grade at Class 3, their cortisol levels may have been elevated due to uncertainty (de Berker et al., 2016). However, results revealed that cortisol levels from Class 3 were lower than those at Class 1.

Past literature has suggested that cortisol is secreted in the face of a stressful situation when it is novel, uncontrollable, and social-evaluative. Further, cortisol is known to be secreted in a situation that involves active performance (e.g., giving a speech) (Dickerson & Kemeny, 2004). It is possible that a typical academic classroom does not have the characteristics of a situation that elevates cortisol. For example, the classroom could have lacked novelty because students became accustomed to the classroom over the course of 75 minutes, resulting in a diminished stress response to the class (Al'Absi & Lovallo, 1993). Also, as students did not take tests during typical classes, the classroom setting may not have been uncontrollable and social-
evaluative in its nature. Further, as students were unlikely to engage in active performance activities during class, it is possible that merely listening to a lecture may not have elicited a salient cortisol response. Similarly, it is possible that classes during the exam period may not have been novel to students as they had become accustomed to the class over the course of the semester. Also, it is possible that classes around the exam period were not particularly more uncontrollable, and social-evaluative compared to other classes. In fact, it is possible that the class in the earlier phase of the semester may have been the most stressful for students because they were still adjusting to the course and were uncertain about what the course would be like. Admittedly, however, such an explanation would contradict previous research and the theory underlying the methodology of the current research.

Taken together, results suggest that students’ anxiety or stress in a college class may be better captured by other physiological measures. In the face of a challenging or stressful situation, both the sympathetic adrenal medullary (SAM) and HPA axes can be activated. Typically, the activation of SAM axis is thought to be more spontaneous compared to HPA axis, as the SAM axis can respond to a stimuli and peak within seconds. In comparison, the HPA axis is known to reach its peak response in 20 - 40 minutes (Dickerson & Kemeny, 2004). The HPA axis is known to be a more conservative stress response system compared to the SAM axis, as it needs longer exposure to stress to be activated. (Blascovich & Mendes, 2010). Overall, stress perceived in the academic setting may be better captured by the SAM axis because it is more easily activated by stressful stimuli.

It is possible that cortisol can represent other psychological states such as engagement, which makes its meaning in real-world settings murkier than has previously been assumed. If
cortisol sometimes represents stress and other times represents engagement, there may be less reason to expect it to fluctuate throughout a course. As a general conclusion, results here suggest the importance of investigating the association of cortisol with performance more regularly, as past research has largely considered cortisol to be negatively associated with performance and has not focused on positive aspects of cortisol.

Interestingly, averaged cortisol levels, cortisol response patterns within class and across classes did not differ by race/ethnicity. I initially expected negatively stereotyped students would have a greater cortisol response within class compared to non-negatively stereotyped students due to social identity threat (Levy et al., 2016; Schmader et al., 2008). Alternatively, I also acknowledged the possibility that cortisol levels may be lower for negatively stereotyped students, as they may be disengaged from the academic domain (Mason et al., 2001) or have a dysregulated HPA axis due to chronic stress (i.e., discrimination and social identity threat) (McEwen, 1998). I expected negatively stereotyped students’ cortisol levels would be elevated during the examination period, as it would be a period when social identity threat would be more salient to negatively stereotyped students. However, my results did not support this hypothesis. These results suggest two possibilities. First, as noted above, it is possible that cortisol response did not capture stress in the academic context. Second, cortisol could represent different psychological meanings for different races/ethnicities. Indeed, I discuss this possibility in more detail later in this section. Also, this result is, to my knowledge, one of the first studies to investigate whether students’ cortisol response pattern differed by race/ethnicity in an academic setting.
My results expand the literature by investigating the relationship between cortisol concentrations during typical classes and academic performance. My results revealed that there was a null relation between students’ cortisol levels during typical classes and academic performance. This result mirrors past studies that found a null relationship between students’ cortisol levels on exam day and their performance outcomes (Pletzer et al., 2010; Schoofs et al., 2008). One explanation for my null finding is that students may not have had a particularly strong cortisol response in the classroom, as past literature suggests that certain levels of cortisol are necessary for hindering cognitive performance (Joels et al., 2006; Lupien et al., 2007). Past literature that found a negative association between cortisol and cognitive performance often used acute applications of cortisol (i.e., administration of glucocorticoid by pharmacological treatment). Typically, 25 mg Cortisone is known to increase saliva levels up to 1.81µg/dL (de Quervain et al., 2000), which is much higher than cortisol levels observed in most of our participants. Alternatively, the positive relationship between cortisol and performance, even the inverted U-shaped relationship where moderate elevation of cortisol is beneficial for performance, was not observed. This apparent lack of a relationship between cortisol and academic performance could, alternatively, be attributed to the influence of a moderator.

Indeed, there was a significant moderation effect by race/ethnicity on the association between cortisol and academic performance. For negatively stereotyped students, cortisol levels were positively related to final scores and negatively related to the probability of dropping out of the class. For non-negatively stereotyped students, there was a null relation between cortisol and academic performance. My findings suggest that, at least for negatively stereotyped students, cortisol may represent engagement. Perhaps this finding can be explained by particular coping strategies that negatively stereotyped students are known to use. Because negatively stereotyped
students are more likely to use two extreme coping mechanisms in the classroom (i.e., disengaging and striving) (James et al., 1983; James et al., 1987; Major & O'Brien, 2005; Schmader et al., 2001), it is possible that low levels of cortisol for negatively stereotyped students actually represent disengagement. Higher levels of cortisol for negatively stereotyped students, then, represent a striving strategy and thus are connected to better performance. However, these results raise the question of whether higher levels of cortisol for negatively stereotyped students are beneficial in the long run. For example, John Henryism, in the long run, is known to be associated with high blood pressure, which could be detrimental to individuals’ health (James et al., 1983). Similarly, higher levels of cortisol may be beneficial for negatively stereotyped students in the short term, but can be detrimental for them in the long term. Indeed, chronically elevated cortisol levels have been found to be associated with worse cognitive functioning (Lupien et al., 1998) and negative health outcomes such as diabetes and obesity (Lloyd et al., 2005). Overall, it is possible that higher levels of cortisol may be positive in the short term but associated with negative outcomes in the long term.

Interestingly, my results revealed that low levels of cortisol were particularly detrimental for negatively stereotyped students. Negatively stereotyped students who had lower levels of cortisol were found to have the worst academic performance. Non-negatively stereotyped students performed better than negatively stereotyped students on average, and cortisol did not predict their performance. Negatively stereotyped students who had higher levels of cortisol had outcomes comparable to non-negatively stereotyped students. Indeed, past literature suggests that lower levels of cortisol for negatively stereotyped student can point to a dysregulated HPA axis (Hostinar et al., 2014; Miller et al., 2007), which impairs cognitive functioning (Tiemensma et al., 2016). Taken together, these results suggest that low levels of cortisol for negatively
stereotyped students may be associated with a dysregulated HPA axis. Although past literature has speculated that a dysregulated HPA axis represented by hypocortisolism may be associated with disengagement from the academic domain for negatively stereotyped students (Mendes & Jamieson, 2011), to my knowledge, my result is first to identify such a relationship.

For non-negatively stereotyped students, cortisol levels were not associated with performance, leaving the interpretation of cortisol ambiguous. Perhaps non-negatively stereotyped students are less likely than stereotyped students to engage in extreme coping strategies, so their cortisol levels are more likely to represent general stress rather than engagement. Alternatively, it is possible that there is a restricted range effect at work for non-negatively stereotyped students. Restricted range effect refers to a situation where the sample has limited variability, and thus the possibility for an association between variables is mathematically reduced. As non-negatively stereotyped students performed significantly better and had less variability in their performance than negatively stereotyped students, the overall detrimental or beneficial effect of cortisol on performance may not have been captured for non-negatively stereotyped students.

**Limitations**

Perhaps the largest limitation is that I do not have a clear non-class cortisol measure to use as a baseline against which to compare cortisol levels during the class. Although I can compare cortisol levels across classes, I cannot determine if students’ cortisol levels were higher on class days than on other days with no classes. Furthermore, due to the lack of a non-class baseline measurement, I cannot clearly identify whether the effect of cortisol on performance is from 1) physiological responses to the class, 2) individual differences in base cortisol level, or 3)
both. However, I can consider students’ cortisol levels assessed during the earlier phase of the semester (Class 1) as a baseline cortisol level. This was the intended goal of the research design, as Class 1 was near the beginning of the semester, after students had adjusted to the class, but before concerns about the exam were likely to become pressing. In an ancillary analysis, I included Class 1 cortisol as a baseline in the analysis (i.e., using it as a covariate) and used cortisol levels around the exam period (i.e., averaged cortisol levels from Class 2, Class 3, Class 4) to predict students’ final scores. Results from this analysis revealed that controlling for baseline cortisol, cortisol levels in Classes 2 – 4 combined continued to be associated with performance in the same pattern as reported in the Results section. That is, cortisol levels around the exam period positively predicted final scores for negatively stereotyped students and had a null association with performance for non-negatively stereotyped students. This analysis partially clarifies the ambiguity of whether physiological response to the class or individual differences in baseline cortisol affected performance. This result suggests that students’ cortisol response during the examination could be driving the effect of cortisol on performance. Whether the first class was representative of a typical cortisol level for students at that time of day remains unknown.

Another limitation is that for each class, students could choose to attend a morning (10:10 – 11:25 AM) or afternoon session (4:10 – 5:25 PM). This could be problematic in interpreting longitudinal patterns of cortisol, as the time of the day can affect cortisol. To eliminate any spurious effect from the time of the day on cortisol levels, I controlled for students’ time since awakening at the moment of the saliva collection throughout my analyses. To further address this issue, I tested in a series of ancillary analysis whether the pattern of cortisol responses and the association between cortisol and performance systematically differed
depending on the time of the class that students attended. They did not. In longitudinal analyses, I looked at morning and afternoon sessions separately and the same patterns reported in the text emerged for both. I also tested whether within-class cortisol was moderated by the time of class and it was not. In examining performance, I included an interaction between time of class and cortisol, and this also did not change the pattern of reported results. These analyses help alleviate the concern that the time of class was driving the reported results.

Another limitation is that I could not measure students’ cortisol levels on the day of the exam. Because the instructor worried that collecting cortisol on exam day could distract students, saliva samples were not collected during the exam. Students might have experienced their highest cortisol response on the day of the exam. It would have been informative to measure whether students’ cortisol level on the day of the exam was higher than normal classes and whether this impacted performance. Measuring cortisol in the middle of class may have also been helpful; such a measure could have captured students’ stress in the class better. Perhaps there could have been a cortisol peak in the middle of the class. This would also provide more nuanced information regarding which pattern of cortisol response was most beneficial to the learning process. Maybe having a moderate peak of cortisol during the class would aid performance, while having an intense peak would be detrimental. Unfortunately, measuring cortisol during the middle of the class was not possible, since assessing cortisol in the middle of the class could have distracted students.

Lastly, I did not measure students’ psychological states at the moment of saliva collection. The relation between students’ psychological state and their cortisol level could have been identified more directly if I had done so.
Future Directions

Future studies can address the methodological limitations described above. In particular, collecting saliva samples on a day without the class would provide important information regarding whether students’ cortisol levels were elevated on any class days (relative to “normal life”) and whether this elevation drove the effect that cortisol had on performance. Additionally, connecting individuals’ psychological states to cortisol response would further the understanding of cortisol. For instance, a recent study by Hoyt et al. (2016) found that elevated levels of cortisol were associated with the momentary positive affect that followed soon after the measurement of cortisol.

Furthermore, more lab studies are needed to identify how the psychological meaning of cortisol can be altered by social psychological factors such as race/ethnicity in the face of social identity threat. Additionally, other physiological measures, like cardiovascular activity and DHEA, could be measured alongside cortisol, which could further clarify the psychological meaning of cortisol. If, for example, elevated cortisol was accompanied by more adaptive cardiovascular activities and increased levels of DHEA, it would suggest that cortisol is related more to engagement than stress.

Conclusion

Despite the limitations noted above, the current study contributes to the understanding of the interpretation of cortisol in the real-life academic setting. The present findings suggest that the psychological meanings of cortisol can differ by race/ethnicity. More specifically, cortisol in the academic setting may represent engagement for negatively stereotyped students but not for non-negatively stereotyped students. My study also further illustrates the multi-faceted nature of
cortisol as I found cortisol can be positively associated with performance for negatively stereotyped students. Finally, my result suggests that low levels of cortisol, which may represent a dysregulated HPA axis, can have deleterious effects on academic performance for negatively stereotyped students.
Appendix A

Self-Affirmation Intervention

During Week 3 of the semester, participants completed an affirmation or control writing intervention exercise in their recitation. Participants were randomly assigned to each condition using the R package blockTools (Moore, 2012). In both conditions, participants ranked the personal importance of 11 values (e.g., membership in social group, relationships with friends or family). In the affirmation condition, participants next wrote about why their #1 ranked value was important to them. Participants in the control condition wrote about why their #9 ranked value might be important to someone else. A second “booster” version of the intervention was administered online in Week 8. In the section below, I will present the intervention materials for both control and affirmation condition. Lastly, I will present Figure 7, which illustrates that longitudinal cortisol patterns across classes differ by intervention condition.
Intervention Material (Control Condition)

WHAT ARE YOUR PERSONAL VALUES?
Below is a list of characteristics and values, some of them may be important to you; some may be unimportant to you. Please rank them from 1 to 11 according to how important they are to you (“1” being the most important item, “11” being the one that is least important to you). Use each number only once.

________ Being Good at Art
________ Physical Attractiveness
________ Creativity
________ Independence
________ Membership in a Social Group (such as your community, racial group, or school club)
________ Music
________ Politics
________ Relationships with Friends or Family
________ Religious Values
________ Sense of Humor
________ Sports Ability

Directions
1) Look at the value you ranked as #9 on the previous page.
2) Think about times when this value would be important to someone else (like another student or a person you’ve heard about).
3) Describe why this value would be important to someone else.
Focus on your thoughts and feelings and don’t worry about spelling, grammar, or how well written it is.

____________________________________________________________________________
____________________________________________________________________________
Again, look at your #9 value. List the top two reasons why someone else would pick this as their most important value.
1. __________________________________________________________________________
2. __________________________________________________________________________

1. This value has influenced some people. Yes No
2. This value is important to some people. Yes No
Intervention Material (Intervention Condition)

(see above for the material which asks students to rank the importance of one’s value as this section is identical across conditions)

Directions
1) Look at the value you picked as most important to you (the value you ranked as #1 on the previous page).
2) Think about times when this value was or would be very important to you.
3) Describe why this value would be important to you.
Focus on your thoughts and feelings and don’t worry about spelling, grammar, or how well written it is.
___________________________________________________________________________
___________________________________________________________________________

Again, look at the value you picked as most important. List the top two reasons why this value is important to you.
1. ___________________________________________________________________________
2. ___________________________________________________________________________

1. This value has influenced my life. Yes No
2. This value is an important part of who I am. Yes No
Longitudinal Cortisol Pattern Across Classes Moderated by Intervention Condition

My analysis revealed that students’ cortisol pattern across classes was moderated by intervention condition, $\beta = 0.02$, $t(590) = 1.99$, $p = .047$, 95% CI [0.0004, 0.03] (see Figure 7 below). It is unclear what this result may represent. This graph suggests that individuals who were in intervention condition may have had less fluctuation in cortisol across classes compared to control group.

Figure 7. Cortisol pattern across classes by condition. Values in grey at each class are predicted cortisol values (log-transformed) for each student averaged within class controlling for adequate covariates, with lines connecting values representing individual students’ average slope across classes. The solid line in blue represents the average slope across classes averaged within class for students in the control condition. The dotted line in red represents the
average slope across classes averaged within class for students in the intervention condition. Time within the class was centered (start of class = -0.5, end of class = 0.5). Dotted line indicates when the first exam took place.
Appendix B

Potentially Mislabeled Saliva Samples

As noted above, there were 12 samples that had discrepancies between information coded on a sheet right after the class and information noted on the label on the tube. For 6 samples, I did not have samples from participants who the research team coded as having received samples. The research team coded that PPID 467 submitted two samples on Class 4. However, I did not have samples from PPID 467 on that class, but rather had samples from PPID 497, who had no record of submitting saliva samples for that class. Similarly, the research team has coded that PPID 398 submitted two samples for Class 3. However, I did not have samples from PPID 398 for that class, but rather had samples from PPID 328, who I did not have any record of receiving saliva samples for that class. Lastly, research team coded that PPID 58 submitted 2 saliva samples for Class 2. I did not receive two samples from PPID 58 for that class, but received samples from PPID 56, who was not recorded as submitting their samples for Class 2. As the labeling process was done based on coded information, I inferred the characteristics of these samples based on the information coded right after the class. As the discrepancies were from having a label that is only one digit different from what is coded, it is very likely that errors happened due to attaching the wrong labels onto the tubes. Additionally, as noted above in the Method section, the research team briefly noted who submitted their samples and for which class they submitted them on the saliva questionnaire. I also utilized this information on the saliva questionnaire and the information aligned with the information that the research team noted right after class.

Additionally, for four samples, the class when samples were submitted was noted differently on the sheet and the label. For PPID 181, the research team coded receiving two
samples on Class 1. However, it is indicated on the label that samples were received on “Class 2 or 3”. Again, I do not have record of receiving sample from PPID 181 for Class 2 or 3. Also, the research team coded receiving two samples from PPID 350 for Class 1. However, I have two samples from PPID 350 that are from Class 4, and I have no record of receiving sample from PPID 350 for that class. For these four samples, I also followed the information that the research team coded as the labeling process took place based on the information that the research team coded. Also, the information written on the saliva questionnaires aligned with the information the research team coded.

Lastly, there were two samples for which the time of day participants submitted samples were noted differently on the coded information and the label. The research team noted that PPID 403 submitted samples for the morning session of Class 2. However, on the labels it is noted that PPID 403 submitted samples for the afternoon session of Class 2. On the saliva questionnaire for PPID 403 for Class 2, it was noted that the saliva questionnaire was submitted for the afternoon session. For these samples, I followed the information on the label, as it also aligned with the information on the saliva questionnaire. Our findings do not change in significance and direction after excluding participants who had an information discrepancy between the label and the information coded after class.
Appendix C

Calculating Intra and Inter-Assay Coefficient

Intra-assay coefficient represents degrees to which duplicate results differ from each other. A lower intra-assay coefficient value represents lower variability within each sample. First, to calculate intra-assay coefficient, the coefficient of variation (\% CV) for each duplicate sample is calculated. This is calculated by computing the standard deviation of results 1 and 2, dividing that by the duplicate mean, and then multiplying by 100. Lastly, the average of individual \% CV is calculated; this is the intra-assay coefficient.

In cases with larger samples, samples have to be assayed on multiple assay plates. Each plate is assayed with saliva controls with both high and low known values. The inter-assay coefficients represent plate-to-plate variation, and it is calculated based on variability of control values in each plate. A lower inter-assay coefficient value represents stability in assay results across plates. To calculate the inter-assay coefficient, the mean, standard deviation, and \% CV (SD of plate means ÷ mean of plate means x 100) of the high and low controls are calculated separately. Then, the average of the high and low \% CV is the inter-assay CV.
## Appendix D

**Predicting Averaged Adjusted Exam Scores**

Table 11. Regression Coefficients Predicting Averaged Adjusted Exam Scores Step 1-3

<table>
<thead>
<tr>
<th>Variable</th>
<th>Step 1</th>
<th>Step 2</th>
<th>Step 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>b</td>
<td>se(b)</td>
<td>CI(95%)</td>
</tr>
<tr>
<td>Intercept</td>
<td>72.46***</td>
<td>0.79</td>
<td>70.90</td>
</tr>
<tr>
<td>Cortisol level&lt;sup&gt;1&lt;/sup&gt;</td>
<td>-0.11</td>
<td>3.31</td>
<td>-6.62</td>
</tr>
<tr>
<td>Race/ethnicity&lt;sup&gt;2&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender&lt;sup&gt;3&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SES&lt;sup&gt;4&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prior performance</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time since awakening&lt;sup&gt;5&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medication&lt;sup&gt;6&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brush Teeth&lt;sup&gt;7&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Missing prior performance&lt;sup&gt;8&lt;/sup&gt;</td>
<td>2.19</td>
<td>3.30</td>
<td>-4.31</td>
</tr>
</tbody>
</table>

N = 271

Note. \( R^2 \) change between steps significant at \( p < .001 \). \(^1\) Log-transformed cortisol level averaged across four classes and mean centered. \(^2\) Non-stereotyped racial group = -1, Negatively stereotyped racial group = 1. \(^3\) Male = -1, Female = 1. \(^4\) An average of standardized SES variables (parent’s income, parent’s education level, social status). \(^5\) Time since awakening averaged across four classes (in hours). \(^6\) Whether students’ took medication 24 hours before saliva collection averaged across four classes. \(^7\) Whether students brushed teeth an hour before the saliva collection averaged across four classes. \(^8\) Participants who had prior performance data = 0, Participants who did not have prior performance data due to not consenting to release their transcript = 1.

\( * p \leq .05. ** p \leq .01. *** p \leq .001. \)
References


