**Title**
Personality and Scheduling in Online Courses



**Author**
Cung, Bianca

**Publication Date**
2018

UNIVERSITY OF CALIFORNIA,
IRVINE

# Personality and Scheduling in Online Courses

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Education

by

Bianca Cung

Dissertation Committee:
Professor Mark Warschauer, Chair
Assistant Professor Di Xu
Assistant Professor Rachel Baker

2018

# DEDICATION

To

my mother and sisters

who have supported me

every step of the way

# TABLE OF CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

# LIST OF APPENDICES

## ACKNOWLEDGMENTS

I am grateful for the endless amount of support that I received throughout this journey. The road to a doctorate's was not easy, and friends, family, and mentors all played integral roles along the way. Below are some select individuals who deserve special recognition.

First and foremost, I would like to thank my mom and sisters for their endless support. Throughout my life, I have made many mistakes—some minor and some more than minor. My mom has always been there to help me through them, and to help me learn from them. In addition, she has made many sacrifices to help me get to where I am today, and I am extremely grateful for all that she has done. My sisters have been my source of motivation and encouragement. Though hundreds of miles away, I know that I can always count on them when I need them the most.

I would also like to thank Dr. Mark Warschauer, who has guided me through graduate school for the very first day I stepped foot onto UC Irvine. Being the first in my family to attend university, I truly did not know what to expect of graduate school. Dr. Warschauer helped me get started on my first project, and he helped me with my first academic paper and presentation as a graduate student. I had a lot of academic and career questions for Dr. Warschauer all throughout my time as a graduate student, and when he did not know the answer, he would put me in contact with someone who did know the answer. I am extremely thankful to have Dr. Warschauer as an academic advisor

I would additionally like to thank Dr. Di Xu for her mentorship and guidance even before she came to Irvine. Her expertise in quantitative research methods have greatly encouraged me to improve my own research. In addition, I truly appreciate her enthusiasm for academic research and in teaching students. She is open to new research ideas and she is very happy to include others in her projects. I am very grateful for all of the opportunities that I had to work with Dr. Xu.

I am also thankful to Dr. Rachel Baker, who has also helped me improve my academic research. She has encouraged me to think more deeply about my own research. Her questions like "why?" and "how?" have taught me to go beyond answering the basic question at hand. In doing so, she has also helped me better communicate my research findings. Dr. Baker has served as an excellent role model for my academic pursuits.

In addition, I am extremely grateful to all of the members of the IVLE group. I especially would like to thank Fernando Rodriquez, Qiujie Li, Jihyun Park, Mariela Rivas, Peter McPartlan, Renzhe Yu, Kameryn Denaro, and Christian Fischer. Not only have the different members of the IVLE group engaged me in meaningful discussions and thought about my research, they have also offered me valuable feedback and guidance in my studies. I am truly appreciative of the variety of perspectives that IVLE had because of everyone's diverse perspectives.

Other graduate students who have been extremely integral to my learning are Grace Lin, Youngmin Park, Soobin Yim, Veronica Newhart, Tamara Tate, Connie Kang, and Ying Xu. As a young researcher, the learning curve was steep. Their kindness, encouragement, and feedback have made each step easier to manage. Their work has served as an inspiration to my own. I am truly happy to have met them along the way.

Lastly, I would like to thank Daisuke Furukawa, who grew to be a bigger part of my life than I imagined. He has been there for me through what I thought was the most challenging, and even impossible, times. His patience and forgiving nature has helped me through a multitude of obstacles. I am truly grateful to have his unconditional love and support.

## Curriculum Vitae

**Bianca Cung**
Ph.D. Candidate, School of Education E-mail: bcung@uci.edu
University of California, Irvine
3200 Education, Irvine, CA 92697

**Research Interests:**
Data Mining, Applied Statistics, Media and Technology in Education, Higher Education, Online Learning, Adaptive Tutorial Instruction, Social Networking, Student Behavior, STEM Education, Second and Heritage Language Learning, International Education

**Education:**
University of California, Irvine
- **Ph.D. Student**, September 2013 – December 2018
- **Master of Arts**, December 2016
  - Major: Education, emphasis in Language, Literacy, and Technology
  - Advisors: Mark Warschauer and Di Xu

University of California, Los Angeles
- **Master of Science**, June 2013
  - Major: Statistics
  - Advisor: Robert L. Gould
- **Bachelor of Science**, June 2013
  - Double Major: Statistics and Mathematics with Specialization in Computing
- **Bachelor of Arts**, June 2013
  - Major: East Asian Studies

**Honors and Awards:**
- National Science Foundation – EAPSI Fellowship, 2017
- Best Graduate Student Poster Award – UCI Data Science Symposium, 2015
- UCLA Statistics Departmental Scholar, 2013
- Research in Industrial Projects for Students (Los Angeles) Awardee, 2012
- Research in Industrial Projects for Students (Hong Kong) Awardee, 2011

**Academic Projects:**
09/2015 – 09/2018 **Graduate Student Researcher**, University of California, Irvine
School of Education
*Online Collaborative Problem Solving in Remedial College Mathematics*
*Investigating Virtual Learning Environments*
(PI: Mark Warschauer)

06/2017 – 08/2017 **EAPSI Research Fellow**, Kyoto University
Center for the Promotion of Excellence in Higher Education
*Investigating the Convergence between Engagement Patterns, User Perceptions, and Course Outcomes in Science MOOCs*
(Advisor: Toru Iiyoshi)

07/2014 – 09/2014 **Graduate Student Researcher**, University of California, Irvine
School of Education
*Formative Assessment in Mathematics*
(PI: Jamal Abedi)

08/2013 – 06/2014 **Graduate Student Researcher**, University of California, Irvine
School of Education and Distance Learning Center
*Distance Learning Courses*
(PI: Mark Warschauer)

06/2012 – 08/2012 **Student Researcher**, Institute for Pure and Applied Mathematics
Research in Industrial Projects for Students, Los Angeles
*Spectral Clustering Approximation Algorithms*
Project funded by NSF and sponsored by IBM

06/2011 – 08/2011 **Student Researcher**, Institute for Pure and Applied Mathematics
Research in Industrial Projects for Students, Hong Kong
*Risk Aggregate Modeling for Life Insurance*
Project funded by NSF and sponsored by MetLife

**Publications:**

**Cung, B.**, Xu, D., Eichhorn, S., Warschauer, M. (in press). Getting academically underprepared students ready through college developmental education: Does the course delivery format matter? *American Journal of Distance Education*.

Baker, R., Evans, B., Li, Q., **Cung, B.** (2018). Does Inducing Students to Schedule Lecture Watching in Online Classes Improve Their Academic Performance? An Experimental Analysis of a Time Management Intervention. *Research in Higher Education*, 1-32.

**Cung, B.**, Xu, D., Eichhorn, S. (2018). Increasing Interpersonal Interactions in an Online Course: Does Increased Instructor Email Activity and Voluntary Meeting Time in a Physical Classroom Facilitate Student Learning?. *Online Learning*, 22(3), 193-215.

**Cung, B.**, Jin, T., Ramirez, J., Thompson, A., Boutsidis, C., & Needell, D. (2012). Spectral clustering: An empirical study of approximation algorithms and its application to the attrition problem. *SIAM Undergraduate Research Online*, 5, 283-303.

**Works in Progress:**

**Cung, B.** (submitted to a refereed journal). Engagement and activity in a MOOC.

Sadehvandi, N., **Cung, B.**, Iiyoshi, T. (submitted to a refereed journal). Prize as a group achievement: Incentivizing fair peer grading in MOOC.

Torres, J., **Cung, B.** (submitted to a refereed journal). Comparing advanced heritage and L2 learners' task performance across face-to-face and ichat contexts.

Xu, D., Glick, D., Rodriguez, F., **Cung, B.**, Li, Q., Warschauer, M. (submitted to a refereed journal). Does blended instruction enhance English language learning in developing countries? Evidence from Mexico.

**Presentations:**

**Cung, B.**, Torres, J. (2018, May). *Comparing Advanced HL-HL and HL-L2 Learners' Writing Task Performance across FTF and SCMC modes*. Paper presented at Digital Learning in the Humanities and Beyond Symposium, Irvine, CA.

Rivas, M., **Cung, B.** (2018, May). *Multitasking During Video Watching and Student Performance*. Poster presented at Association for Psychological Science Annual Convention, San Francisco, CA.

**Cung, B.**, Li, Q., Evans, B., Baker, R. (2018, March). *Improving Academic Outcomes through Time Management: A Randomized Control Trial of a Scheduling Intervention to Improve Performance in Online Postsecondary Courses*. Paper presented at Association for Education Finance and Policy Annual Conference, Portland, OR.

**Cung, B.**, Torres, J. (2018, February). *Comparing Advanced HL-HL and HL-L2 Learners' Writing Task Performance across FTF and SCMC modes*. Paper presented at NHLRC's Third International Conference on Heritage/Community Languages, University of California, Los Angeles.

**Cung, B.**, Rivas, M., Xu, D., Warschauer, M. (2017, May). *Attitude Changes Working in Groups or Alone in Online Math*. Poster presented at Association for Psychological Science Annual Convention, Boston, MA.

Torres, J., **Cung, B.** (2017, April). *Comparing Advanced Heritage and L2 Learners' Task Performance Across Face-to-Face and Ichat Contexts.* Paper presented at International Conference of Task-Based Language Teaching, University of Barcelona, Spain.

**Cung, B.**, Torres, J. (2017, March). *Comparing Advanced Heritage and L2 Learners' Task Performance Across Face-to-Face and Ichat Contexts.* Poster presented at American Association for Applied Linguistics Conference, Portland, OR.

Baker, R., Evans, B., Li, Q., **Cung, B.** (2017, March). *Nudging Students to Better Academic Outcomes: A randomized control trial of a scheduling intervention to improve persistence and performance in online postsecondary courses*. Paper presented at Association for Education Finance and Policy Annual Conference, Washington, D.C.

**Cung, B.**, Thrall, K., Torres, J. (2017, February). *Personality traits and HL-L2 task performance across Face-to-Face and CMC contexts*. Talk presented at National Symposium on Spanish as a Heritage Language, University of California, Irvine.

Lin, G., **Cung, B.**. Wu, M. (2016, October). *Interactive Tales: World Languages*. Demo presented at Society for Research in Child Development: Technology and Media in Children's Development, University of California, Irvine.

Lin, G., **Cung, B.** (2016, October). *Interactive Tales*. Demo presented at the Digital Media and Learning Conference, University of California, Irvine.

**Cung, B.** (2016, May). *How Does Number of Deadlines and Course Format Affect Time Online?* Poster presented at Data Science Initiative: Fostering Literacy and Learning with Text and Data Mining, University of California, Irvine.

**Cung, B.** (2016, April). *Learning in an Online Course: How Study Habits Can Affect Course Performance.* Paper presented at Association of Graduate Students Symposium, University of California, Irvine.

**Cung, B.** (2015, May). *Chance to Meet: Does Offering Face-to-Face Time Help an Online Math Course?* Poster presented at Data Science Initiative: An Afternoon of Data Science, University of California, Irvine.

**Cung, B.** (2015, May). *Chance to Meet: Does Offering Face-to-Face Time Help an Online Math Course?* Poster presented at Data Science Initiative: Text and Data Mining for Interactive Online Learning Symposium, University of California, Irvine.
* Received "Best Graduate Student Poster" Award

**Cung, B.** (2014, January). *Student Participation and Drop-Off: A Case Study of a Preparatory Biology MOOC*. Paper presented at High Impact Practices: Interdisciplinary Collaborations and Creative Connections, University of California, San Diego.

Warschauer, M., Jiang, S., Williams, A., **Cung, B.**, Domina, T., Smyth, P., O'Dowd, D. (2013, December). *Peer Assessment and Academic Achievement in a Gateway MOOC*. Paper presented at MOOC Research Initiative, University of Texas, Arlington.

**Cung, B.** (2012). *Risk aggregation*. Poster presented at the annual MAA and JMS Joint Mathematics Meetings, Boston, MA.

**Teaching Experience:**

09/2015 – 06/2018 **Assistant Instructor**, University of California, Irvine
Anteater Recreation Center
"Kendo"

10/2016 – 12/2016 **Teaching Assistant**, University of California, Irvine
Teaching and Learning Research Center
"R for Data Analysis"

06/2015 – 07/2018 **Math Lab Instructor**, The Harker School
Taught at the Middle School Summer Institute over multiple summers
"Geometry Lab" and "Algebra II Lab"

10/2014 – 06/2015 **Teaching Assistant**, University of California, Irvine
School of Education
"Multicultural Education in K-12 Schools" and
"Theories of Development and Learning Applied to Education"

06/2013 – 08/2016 **Teaching Assistant and ELI Program Coordinator**, The Harker School
Taught at the English Language Institute and assisted with administrative aspects over multiple summers

**Leadership:**

09/2015 – 06/2018 **Council Member**, University of California, Irvine Association of Graduate Students

09/2013 – 06/2015 **Lab Manager**, UCI School of Education Digital Learning Lab

**Technical and Language Skills:**

Statistical Software (R, STATA, MATLAB)
Programming (Python, Java, C++, C#)
Languages (English, Cantonese, Mandarin, Teochew)

## ABSTRACT OF THE DISSERTATION


Personality and Scheduling in Online Courses

By

Bianca Cung

Doctor of Philosophy in Education

University of California, Irvine, 2018

Professor Mark Warschauer, Chair

Online education is becoming increasingly integral to postsecondary education. However, student performance in online courses do not parallel performance in in-person courses. This dissertation presents a series of studies investigating three undergraduate online courses at a large public four-year university.

The first study examines how online instruction affects different personality types, as measured by student responses to items on the Big Five Inventory. Using historical course grades for all students enrolled in two online public health courses and employing student fixed effects in a linear regression models, I estimate student performance in online courses relative to their own performance in in-person courses. I include interaction effects with personality to examine whether online courses differentially impact academic performance for different personality types. Findings suggest that highly conscientious students and students with high openness to experience are differentially impacted by online instruction.

The second and third study examine a scheduling intervention implemented in three different courses. Using a randomized control design, students in the courses were assigned either to a treatment group or control group. Treatment students were asked to schedule their

coursework on a weekly basis for the full duration of each course. One of the two studies incentivized the act of studying as scheduled while the other simply incentivized the act of scheduling without regard to whether or not the student followed his or her schedule. Using course clickstream data, gradebook data, institutional data, and survey data, a series of linear regression models are built to estimate the intent-to-treat effect of the scheduling intervention on students' study times and course grades. Mixed results are found across the two studies, though they suggest neutral to positive results for the scheduling intervention.

The last study examines students' subsequent course performance and choice of major following the scheduling intervention in the second of the two studies. Using subsequent course grades and course performance averages for both the given and past terms, regression models are built to estimate the effect of the scheduling intervention. Findings suggest long-term positive impacts of the intervention on academic performance.

# CHAPTER 1: INTRODUCTION

Online education is taking on an increasingly prominent role in higher education institutions. About one in four students take at least one online course, and half of those students take exclusively online courses (NCES, 2015). Moreover, an increasing proportion of midsize to large institutions, those that enroll over 3,000 students, report that online education is critical to their long-term strategy (Allen, Seaman, Poulin, & Straut, 2016). Over 70% of academic leaders believe that online courses are as good as or better than face-to-face courses (Allen, Seaman, Poulin, & Straut, 2016). However, with online courses taking a prominent role in higher education, research focuses will need to shift from whether online is as good as traditional instruction to how we can make online instruction better-suited for learners.

Reasons for adopting online instruction can vary for many institutions. Some common reasons include the need to free up classroom space, accommodate a growing campus population, address a broader range of student needs, meet budget constraints, improve current teaching methods, and facilitate on-time graduation. In particular, the ability for online instruction to overcome classroom space and meeting time constraints is advantageous for both institutions and students. By freeing up classroom space during peak lecture times, online instruction allows campuses to offer courses that would otherwise not be available. Instructors can also reach a broader range of students, including those with busy work schedules or those with longer commutes. For students, online instruction opens up opportunities for registration in high-demand courses that would otherwise have limited seats in-person. It also decreases opportunities for schedule conflicts since many online courses can be taught asynchronously.

Widespread sentiment that online courses are as good as face-to-face courses by academic leaders (Allen, Seaman, Poulin, & Straut, 2016) suggests that online courses has the potential alleviate many challenges faced by educational institutions. Unfortunately, online education in its current state has many challenges, including high dropout rates (Carr, 2000; Dutton, Dutton, & Perry, 2001; Lee & Choi, 2010; Levy, 2007; Tello, 2007), reduced engagement (Conrad, 2002; Lyons, 2004; Singh & Pan, 2004), and poor student performance (Jaggars & Xu, 2010; Taylor, 2003; Xu & Jaggars, 2011). If these problems persist, online courses may turn out to be more of a problem than a solution.

Online education, as it is now, helps some students more than others. Empirical studies have found individual characteristics to be predictors of success in online courses. Xu and Jaggars (2014), for example, found that males, younger students, Black students, and students with lower grade point averages (GPAs) tended to have the strongest declines in online courses compared to in-person courses. Figlio, Rush, and Yin (2013) found that students with lower levels of academic preparation also saw greater declines in the online instructional mode. Other studies have looked at less permanent student characteristics. Highly motivated students with self-regulated learning and time management skills (Bambara, Harbour, & Davies, 2009; Deal III, 2002; Liu, Gomez, Khan & Yen, 2007) tend to do better in online courses compared to their peers. In addition, some studies have suggested that personality traits, particularly conscientousness, also predict academic performance in online courses (Al-Dujaily, Kim, & Ryu, 2013; Alkış & Temizel, 2018; Boghikian-Whitby & Mortagy, 2016).

In order to improve online education, studies need to further investigate the different individual characteristics that are advantaged or disadvantaged through online learning, compared to in-person learning. In addition, researchers and practitioners need to then act on the new discoveries and uncover interventions or new teaching methods to help key groups. The effect of the intervention should be evaluated both during and after the intervention period, in case there are any withdrawal effects. In the following series of studies, presented in three chapters, I investigate these different aspects of online education.

In Chapter 2 of this dissertation, I investigate the role of personality in online courses. In particular, I examine whether personality predicts academic performance in online courses, and whether different personality types are equally advantaged (or disadvantaged) in online courses compared to in-person courses. I measure student personality in five dimensions through a widely used and validated and instrument, the Big Five Inventory (John & Srivastava, 1999). Using student responses to the instrument and historical course data for students enrolled in two undergraduate online courses, I employ linear regression to estimate the effect of personality characteristics on academic performance. Taking advantage of the panel nature of historical course data, I use student fixed effects in my linear regression models to effectively compare students' academic performance in online courses to their own performance in in-person courses. To uncover any differential effects of personality, I examine the interaction between online courses and each personality measure. Results suggest that personality characteristics do predict performance in online courses and that students with different personality characteristics do not see equal gains or losses in online instruction.

Chapter 3 then investigates a scheduling intervention in undergraduate online courses, presented as two studies. Both studies use random assignment to determine treatment and control groups for the investigation. The first of the two studies examines the scheduling intervention implemented in a five-week online physics course taught during the summer. Treatment students in the first study were asked at the end of each week to schedule when they would watch each of the upcoming five weekly lecture videos for the course. They were then awarded a nominal amount of extra credit for following their scheduled times. The second of the two studies examines the scheduling intervention implemented in two similarly-formatted eleven-week online public health courses taught during fall of the regular school year. Treatment students in the second study were asked at the beginning of the week to schedule when they would work on their online coursework. They were awarded a nominal amount of extra credit for scheduling. In both studies, control students received a theoretically inert weekly survey in place of the scheduling prompts. Results for the two studies show mixed, though generally neutral to positive, effects for students' study habits and academic performance.

In Chapter 4, I follow up on second study of Chapter 3 by investigating student grades in the subsequent term and their choice of major by the end of the school year. Using institutional data on students' subsequent course grades and choice of major, I examine whether the original scheduling intervention has effects on academic performance beyond the treatment period. In particular, I use linear regression models to compare treatment and control students' grades earned while controlling for demographic background characteristics. I also compare treatment and control students' grades relative to classmates, as well as their selection of courses by difficulty. Finally, I look at whether treatment and control students differ in their choice of major by the end of the school year. Findings suggest promising results for the scheduling intervention as treatment students earned higher grades and tended to pick harder courses. While both treatment and control groups saw an increase in the number of public health majors, there was a steeper increase for treatment students.

Online education is taking on an increasingly important role in higher education. However, students are not benefitting equally from online courses. Similarly, universities may also not be helping the students who need the most support. Thus, to improve on the current status on online education, studies need to investigate how various groups of students are differentially affected by online education and whether an intervention can effectively address the needs of students in online courses. This dissertation thus aims to gain a better understanding of how different groups of students are impacted by the change in instructional modality and how a known issue with online courses, time management, can be addressed through a low-cost intervention. It achieves these goals by investigating the differential impacts of online education according to individual student characteristics (as defined by five personality dimensions) and by implementing a scheduling intervention in online undergraduate courses through a randomized control trial.

# CHAPTER 2: THE DIFFERENTIAL IMPACT OF ONLINE COURSES ACROSS PERSONALITIES

Online education has been rapidly expanding in higher education. Well over 5 million students enroll in at least one online course each year (Allen, Seaman, Poulin, & Straut, 2016). In addition, over 70% of academic leaders believe that online courses are as good as or better than face-to-face courses, and many institutions have adopted online education as part of their long-term strategy (Allen, Seaman, Poulin, & Straut, 2016). However, a large body of research looking in to the effects of online instruction on academic performance has found mixed results (e.g. Bernard et al., 2004; Lack, 2013; Means, Toyama, Murphy, Bakia, & Jones, 2010; Russell, 1999). These meta-analyses highlighted the complex nature of determining the effect of online instruction compared to in-person instruction.

Student characteristics, such as age, sex, motivation, and self-regulation skills, are some major factors that prior studies considered when determining the effect of online instruction. Studies aiming to obtain causal estimates of online courses have found that students from disadvantaged backgrounds and students with lower levels of preparation see a greater disadvantage in academic performance when enrolled in online courses (Figlio, Rush, & Yin, 2013; Xu & Jaggars, 2013). Other studies have found that only highly motivated students with high levels of self-regulation and time management skills perform well in online courses (Bambara, Harbour, & Davies, 2009; Deal III, 2002; Liu, Gomez, Khan & Yen, 2007).

Another set of individual characteristics that potentially affects learning gains in an online course is personality—the qualities that define a person's distinctive character. It has long been established that personality and learning are closely related (Eysenck, 1978; Messick, 1984). Empirical studies have shown that personality predicts college performance (Noftle &

Robins, 2007; Poropat, 2009; Rosander, Bäckström, & Stenberg, 2011; Stajkovic, Bandura, Locke, Lee, & Sergent, 2018; Vedel, 2014; Wolfe & Johnson, 1995). Studies in online contexts have suggested that personality predicts students' preference for online courses (Boghikian-Whitby & Mortagy, 2016; Harrington & Loffredo, 2010), satisfaction with online courses (Cohen & Baruth, 2017; Daughenbaugh, Ensminger, Frederick, & Surry, 2002), and performance in online courses (Alkış & Temizel, 2018; Boghikian-Whitby & Mortagy, 2016). However, findings thus far that try to examine the relationship between personality and performance are correlational studies or face selection bias and compare unlike groups.

The current study investigates personality as a predictor of online course performance compared to in-person course performance. It aims to make a causal inference by using individual fixed effects to compare students' grades in online courses with their own grades in in-person courses. This paper starts by first estimating the overall gap between online and in-person courses. This paper then investigates whether the gap varies according to different personality factors, as indicated by responses to personality questions.

## Prior Literature and Theory

### Personality Measures

Though personality can be broadly defined as individual differences enduring behavior, attitudes, emotions, and cognition, researchers have yet to agree on a single definition of personality. As a result, a number of models have been proposed to capture broad idea of personality. Two of the most common models used in studies of learning are the five-factor, "Big Five," personality model (Costa & McCrae, 1992; McCrae & Costa, 1987) and the Myers-Briggs Type Indicator (MBTI; Myers, McCaulley, Quenk, & Hammer, 1998).

**Big Five.** The Big Five model was derived from natural language words used by individuals to describe themselves and others (Costa & McCrae, 1992; John, Angleitner, & Ostendorf, 1988; McCrae & Costa, 1987). Other personality models have been based off of questionnaires with items measuring discrete constructs; often, the constructs were determined because they were important for a specific practical application (Goldberg, 1971). In contrast, the Big Five model grew from extracting thousands of English trait terms (Allport & Odbert, 1936), grouping the terms (Cattell, 1946), and then obtaining ratings of the groups for factorization (Norman, 1963; Tupes & Christal, 1992). McCrae and Costa (1987) used extraversion, agreeableness, conscientiousness, neuroticism, and openness to experience to describe the five dimensions of the Big Five. A brief description of the dimensions are provided below:

- Extraversion – the extent of an individual's outgoingness, assertiveness, and sociability
- Agreeableness – the extent of an individual's cooperativeness, trustworthiness, and willingness to help others
- Conscientiousness – the extent of an individual's self-discipline, organization, achievement-orientation, and competence
- Neuroticism – the extent of an individual's feelings of security, emotional stability, and self-control
- Openness (to experience) – the extent of an individual's willingness for new experiences, intellectual curiosity, creativity, and independent-mindedness

The Big Five personality traits have been found to be fairly stable across cultures (McCrae & Terracciano, 2005). It is important to note, however, that despite the consistent divides into each of these dimensions, some cultures do not see all five dimensions (Gurven, von

Rueden, Massenkoff, Kaplan, & Lero Vie, 2013; Saucier & Goldberg, 2001). One explanation is that the dimension might not exist in the culture at all. Similarly, some cultures may have other dimensions not captured in the Big Five (Cheung & Leung, 1998).

In addition, personality traits emerging from the Big Five have shown to be fairly consistent, with small changes happening gradually over time rather than in response to adverse events (Cobb-Clark & Shurer, 2012). Age-related changes from childhood to adulthood generally tend towards increases in agreeableness and conscientiousness and decreases in the other dimensions (McCrae, 2002). Pronounced changes tended to be in late childhood or adolescence, when gender differences also emerge (Soto, John, Gosling, & Potter, 2010).

**MBTI.** While widely criticized in education, the MBTI is another commonly used model for measuring personality. Like most other personality models, the MBTI was developed with discrete constructs in mind for a specific practical application (Goldberg, 1971). In particular, it was derived from the clinical work of Swiss psychiatrist Carl Jung who postulated that humans experienced the world in four psychological functions (Huber, Kaufmann, & Steinmann, 2017). Isabel Myers and her mother Katharine Briggs developed an inventory and tested it on medical students and college students, and over time the inventory was used for various purposes (Mccaulley, 1990). The MBTI, at present, has several different forms of varying question lengths. It yields 16 personality types, described by four main dichotomies (Myers, McCaulley, Quenk, & Hammer, 1998):

- extraversion/introversion – where an individual gets his or her energy
- sensing/intuition – how an individual gathers information
- thinking/feeling – how an individual makes decisions
- judging/perceiving – how an individual organizes his or her world

**Overlapping Measures.** Between the Big Five personality model and the MBTI, most studies of personality in the computer-based learning context use the MBTI (Tlili, Essalmi, Jemni, Kinshuk, & Chen, 2016) though other studies in the postsecondary context use the Big Five model (Noftle & Robins, 2007; Poropat, 2009; O'Connor & Paunonen, 2007; Vedel, 2014). Both models of personality are similar and have overlaps across dimensions. Extraversion in the Big Five model, for example, is strongly related to extraversion in the MBTI. Agreeableness is related to the thinking/feeling dichotomy, conscientiousness is related to the judging/perceiving dichotomy, and openness is related to the sensing/intuition dichotomy (Furnham, 1996). The MBTI, however, does not cover the emotional stability dimension found in the Big Five.

To capture a broader range of personality descriptors, and include the emotional stability dimension, the current study uses the Big Five model to measure personality. This study uses the 44 items from the Big Five Inventory (John & Srivastava, 1999) due to its high reliability and fewer question items compared to the NEO Personality Inventory Revised and NEO Five Factor Inventory (Gosling, Rentfrow, & Swann Jr., 2003).

### Theory of Personality in Learning

This investigation is based on psychology research on personality and job performance, and personality and job satisfaction. A long line of research has investigated the predictability of personality measures of an individual's job satisfaction and performance (for reviews, see Barrick & Mount, 1991; Barrick, Mount, & Judge, 2001; Judge & Bono, 2001; Judge, Heller, & Mount, 2002; Judge, Rodell, Klinger, & Simon, 2013; Kristof-Brown, Zimmerman, & Johnson, 2005; Tett, Jackson, Rothstein, 1991). Likewise, the development of the MBTI came, in part, from Myers and Briggs' observations of a mismatch between the jobs that people took and their personality types (Mccaulley, 1990).

Meta-analyses have generally found some personality traits that suggest high job performance (Tett, Jackson, & Rothstein, 1991). However, the relationship does not hold true across all job types. Barrick and Mount (1991) found that certain types of personality were valid predictors depending on the job type. Extraversion, for example, was an important predictor for jobs involving social interaction. Additionally, Kristof-Brown, Zimmerman, and Johnson (2005) found that person-job fit was a strong positive predictor of job satisfaction and organizational commitment. Taken together, studies have shown the importance of personality on job success.

In the academic setting, studies have investigated personality's role in academic success. Feldman, Smart, and Ethington (1999), for example, found a relationship between students' personality and college major choice, and that students saw increases in their dominant skills when they were in congruent fields of study. Studies have also looked at personality as predictors of grades (for reviews, see Noftle & Robins, 2007; Poropat, 2009; O'Connor & Paunonen, 2007; Vedel, 2014) and learning preferences (e.g. Busato, Prins, Elshout, & Hamaker, 1998; Jessee, O'Neill, & Dosch, 2006; Kamal & Radhakrishnan, 2018). These studies suggest that personality also play a role in the learning context, though it is unclear whether the role is the same for both online and in-person settings.

**Personality and Academic Performance**

Following the concept of person-job success, a wide range of studies have also examined the relationship between personality and academic performance. Most studies conducted have been observational and were only able to generate correlations, at best. Nevertheless, the multitude of studies have enabled researchers to identify trends and inconsistencies in a series of meta-analyses and reviews (Noftle & Robins, 2007; Poropat, 2009; O'Connor & Paunonen, 2007; Vedel, 2014).

Of the Big Five factors, conscientiousness has had the most consistent association with college performance. Students who were more conscientious tended to have higher GPAs (Chamorro-Premuzic & Furnham, 2003a, 2003b; Conard, 2006; Furnham, Chamorro-Premuzic, & McDougall, 2003; Hakimi, Hejazi, & Lavasani, 2011; Phillips, Abraham, & Bond., 2003; Stajkovic, Bandura, Locke, Lee, & Sergent, 2018; Wolfe & Johnson, 1995). Other studies have shown that the positive relation also holds for other measures of academic performance, such as midterm grades (Busato, Prins, Elshout, & Hanmaker, 2000; Furnham & Chamorro-Premuzic, 2004; Hair & Hampson, 2006), final course grades (Conard, 2006; Dollinger & Orf, 1991; Lounsbury et al., 2003; Paunonen & Ashton, 2001), and essay grades (Hair & Hampson, 2006). Two meta-analyses have found an overall correlation of 0.24 between conscientiousness and academic performance (O'Connor & Paunonen, 2007; Vedel, 2014), suggesting that conscientiousness is a strong predictor of academic performance.

Other personality measures have yielded less straightforward results. Meta-analyses by O'Connor and Paunonen (2007) and Vedel (2014) have both found a weak positive correlation between openness and academic performance, a weak positive correlation between agreeableness and academic performance, no correlation between neuroticism and academic performance, and no correlation between extraversion and academic performance. However, individual studies across these four personality measures have shown positive, non-significant, and negative results.

Some studies have found context-specific cases when the other four personality measures are significantly correlated with academic performance outcomes. A review by Noftle and Robins (2007), for example, found that openness was a significant positive predictor of SAT verbal scores. Additionally, Rosander, Bäckström, and Stenberg (2011) found a positive relation between academic performance on practical topics (e.g. art, music, home/consumer economics, crafts) and agreeableness, and between academic performance in languages (English, Swedish, foreign languages) and neuroticism. Rothstein, Paunonen, Rush, and King (1994) found a positive relation between extraversion and MBA students' classroom participation grades. Thus, while each of these other personality measures do not strongly correlate with overall college grades, they are predictive of narrower measures of academic performance.

**Personality in Online Spaces**

Given that personality predicts overall academic success, it could be argued that personality's predictability should be similar in in-person and online learning. However, many researchers have noted key differences across instructional mediums, including the isolated nature of distance instruction compared to in-person instruction. As a result, learning advantages may be exacerbated or dampened in online environments. Terrell (2005), for example, noted that extraverts may be at a disadvantage in online learning environments due to the asynchronous nature of communication. Similarly, Watjarakul (2016) noted that students high in neuroticism may be disadvantaged due to stress from learning in an unfamiliar environment.

Research is starting to uncover general trends of the role of personality in online learning spaces. Studies have looked into personality as a predictor of preference for online learning (Harrington & Loffredo, 2010; León, Morales, & Vértiz, 2017; Watjatrakul, 2016), perceptions of online learning (Keller & Karau, 2013; Tlili, Essalmi, Jemni, Kinshuk, & Chen, 2016), satisfaction in online courses (Cohen & Baruth, 2017; Daughenbaugh, Ensminger, Frederick, & Surry, 2002; Shih, Chen, Chen, & Wey, 2013), and performance in online education (Al-Dujaily, Kim, & Ryu, 2013; Alkış & Temizel, 2018; Boghikian-Whitby & Mortagy, 2016). Findings have confirmed the general predictability of certain traits. For example, Alkış & Temizel (2018) found that conscientiousness was a positive predictor of academic performance. Shih, Chen, Chen, and Wey (2013) found that extraversion and conscientiousness positively predicted online learning satisfaction in English online courses. Cohen and Baruth (2017) similarly found that extraversion and conscientiousness positively predicted satisfaction in teacher education online courses, though extraversion was not significant and openness was instead a significant positive predictor.

However, while general trends suggest that certain personality types predict success in online courses, it is unclear whether students would have had higher gains in the in-person context or whether online instruction afforded a boost in learning. Boghikian-Whitby & Mortagy (2016) did find differential learning advantages for different personality types in online and in-person courses. However, their study was observational and compared unlike groups. This study aims improve on existing studies of differential learning gains by personality by comparing students to themselves in online and in-person courses.

## Context

### Context Description

This study obtained student-reported personality data from two online, undergraduate public health courses taught during fall 2017 at a four-year public university in the western United States. The courses were delivered over an 11-week time span with the first ten weeks dedicated to instruction and the eleventh week dedicated to the final exam. The courses, which will be referred to as PH1 and PH2, were required entry-level courses for students intending to complete a public health major. The courses were also open to students in other majors. Both courses were taught entirely online by the same instructor. Exams were administered online through an online proctoring service, so the course never met in person.

A randomly selected group of students in the course received question items from the Big Five Inventory (John & Srivastava, 1999). Other students in the course received a different set of question related to a separate study (see Chapter 3). Students in the course were asked 8 out of 44 items from the Big Five Inventory on a weekly basis. Question items were administered to student in the form of a weekly survey made available online to students at the beginning of the week Students were asked each question item twice, once during the first half of the term and once again five to six weeks later. Though studies have indicated that personality was a stable measure (McCrae & Terracciano, 2005), I checked the correlation between the first response and the second to confirm stability in the online course. Students were not required to complete any or all items, though they were awarded a nominal amount of extra credit as an incentive to complete each week's set of questions.

**Participants**

The two public health courses had an open enrollment period when students were allowed to freely add or drop courses without any transcript record or penalty fees. Thereafter, students were assessed tuition fees for their selected courses. A total of 228 students were enrolled in PH1 and 96 students were enrolled in PH2. However, due to random assignment of students into treatment and control groups for a different study, only 162 students across both courses (117 students from PH1 and 45 students from PH2) received questions from the Big Five Inventory (John & Srivastava, 1999). Of the 162 students who received personality questions, personality measures were obtained for 157 participants. The analysis sample thus includes all 157 students with personality measures.

## Data

### Demographics

The institution provided administrative data on the demographic background of all students enrolled in PH1 and PH2. Table 1 presents demographic summary statistics of students in the analytic sample. Students were on average 20 years of age. Most students in the sample were female (74.5% compared to 25.5% males). Less than half of students in the sample were underrepresented minorities (33.8%) and from low-income households (39.5%). About half of the students were first-generation college students. 36.3% of students were from households that do not speak any English at home.

### Personality Measures

Table 1 reports the composite scores for each of the five personality dimensions. Composite scores were calculated as the average of all answered items belonging to each personality dimension. Students in the course were, on average, moderate in extraversion and neuroticism. They were moderately high in conscientiousness and openness, and high in agreeableness.

Correlations between composite scores from the first half of the course and the second half of the course ranged between 0.70 (neuroticism) and 0.87 (extraversion). The high test-retest reliability suggested that students were stable in their personality measures throughout the course of the term.

Between 55.0% of students and 88.8% of students completed the personality questions each week. The lowest rate was found in week 2 and the highest rate was found in week 7. Overall, slightly more students submitted responses to the personality items in the later weeks of the course. As a result, more students had all personality composite scores available in the second half of the course. Since personality scores were stable over the duration of the term, scores in the second half of the term were also used in this study. For students who had both sets of composite scores (from the first half of the course and from the second half of the course), composite scores were averaged between the two halves. For students who had only one set of composite scores, the available set of composite scores was used.

**Academic Performance**

**Public Health Scores.** Course gradebook data was obtained and used to estimate the impact of online instruction for different personality types. Both PH1 and PH2 had one midterm exam during the fifth week of the course and one final exam during the eleventh week of the course. Each exam was worth 60 points each. On average, students in the analysis sample earned 49 points (82.2%) on the midterm and 44 points (73.1%) on the final. The exams each accounted for 25% of the final course score. The remaining 50% came from other course assignments, including weekly short written assignments (9%), discussion forum posts (9%), a peer-reviewed presentation (6%, due week 6), research paper (24%, due week 9), and weekly quizzes in all weeks except 5, 6, and 9 (2%). Students' average final course score was 82.14% after accounting for extra credit awarded. On average, the course awarded grades between "B-" and "B."

**All Course Grades.** Grades for all courses up until fall 2017 were obtained for students in the sample. Across the 157 students in the sample, 3,142 course outcomes were included in this analysis, of which 2,727 were from in-person courses and 415 were from online courses (including public health). In addition, 360 of the grades were pass or no pass grades and 4 were withdrawals. Figure 2.1 shows the distribution of all grades in this study. About half of the grades in the study were "B" or higher. Non-passing grades ("C-" or below, including no-passes) accounted for 14.6% of grades in the current data set.

In the analyses, letter grades were converted to a 4.0-scale using the same university conversion rules. Both “A+” and “A” equated to four points. All other letter grades followed the usual rules of the 4.0-scale (e.g. “B” equated to three points). In some courses, students were allowed to choose a pass or no pass grading option. This allowed students to take courses without the course score impacting the student’s GPA. The university’s rule for earning a “pass” is that students had to earn a minimum of a “C” grade. As such, “pass” and “no pass” were converted to the lowest possible point value, two and zero respectively. Withdrawals were converted to zero points. Grade points were averaged for each student. The sample’s average grade was about a “B-” (Table 1). The overall average grade was slightly lower than the average public health final course score.

## Method of Analysis

The method of analysis used in this paper is similar to that of Xu and Jaggars (2014), who used student-level fixed effects to address individual-level unobserved variable bias in their comparison of in-person and online courses. To start, I investigate how personality predicts student academic performance in the online instructional mode. I first look at scores from the two online public health courses. I use the following ordinary least squares model:

$$Y_i = a + \beta_w personality_{w,i} + \gamma \Theta_i \tag{1}$$

In Equation (1), $personality_w$ are the five personality dimensions of key interest in predicting the outcome variable Y, which stands for the midterm score, final exam score, and final course score. To account for basic individual differences, student-level covariates ($\theta_i$) are also included in the model. Student-level covariates include seven demographic background items.

A similar model is adopted to understand the association between different personality dimensions and academic performance in students’ in-person courses. However, since course difficulty and grading vary across campus, course fixed effects are included in the regression model to account for biases resulting for variation in course difficulty. The model with grade outcome Z thus looks like:

$$Z_{i,c} = a + \beta_w personality_{w,i} + \gamma\theta_i + \rho_c \tag{2}$$

By accounting for course differences and including the course fixed effects term $\rho_c$, differences in scores between students are thereafter due to differences in other characteristics beyond course-level differences, such as individual-level characteristics.

Finally, I investigate whether academic performance is differentially influenced between online and in-person instruction for different personality types. To do this, I include interaction terms between an online course indicator and each personality measure. Since there may be unobserved characteristics that underlie a student's propensity to enroll in online courses, I also include student-level and department-level fixed effects to account for overall differences across students:

$$Z_{i,c,d} = a + \varphi online_{i,c,d} + \beta_w personality_{w,i} \times online_{i,c,d} + \sigma_{i,d} \tag{3}$$

The department-level fixed effects are included to account for different departmental grading criteria or difficulty. Personality as a main effect is accounted for in the variable $\sigma_{i,d}$ since it is an individual level characteristic. Averaging Equation 3 over courses in the same department gives the following equation:

$$\bar{Z}_{i,d} = a + \varphi \overline{online}_{i,d} + \beta_w personality_{w,i} \times \overline{online}_{i,d} + \sigma_{i,d} \tag{4}$$

Subtracting Equation 4 from Equation 3 yields the within-department course demeaned data:

$$\ddot{Z}_{i,d} = a + \varphi onl\ddot{i}ne_{i,d} + \beta_w personality_{w,i} \times onl\ddot{i}ne_{i,d} \tag{5}$$

Importantly, the term $\sigma_{i,d}$ disappears, indicating that student-level unobserved bias has been eliminated. The model thus compares courses within the same department taken by the same student. The online coefficient $\varphi$ thus represents the performance gap between online and in-person courses within-student, and the coefficients $\beta_w$ represents the differential online performance gap for students whose personality deviates from the average.

The above equations, however, do not account for time-varying changes. For example, students may take harder courses as they advance in grade level and thus have lower scores in later terms. Since the public health courses were the last grading period available, score for in-person courses would be biased upwards. Similarly, students may become more experienced with college courses as time goes on, which then biases scores for in-person courses downwards. In addition upper division courses may also award higher (or lower) grades to students on average and bias students' scores for in-person and online formats. As a robustness check accounting for time-varying changes, I conducted the same set of analyses only on courses taken in the fall 2017, the same time as PH1 and PH2 in this study.

## Results

### Personality in the Online Course

In order to explore the how academic performance for different personality types are impacted by online course, I examined the overall course score for PH1 and PH2. I also examined the individual assignments that made up the overall course score. Linear regression results for the final course score and major (non-weekly) course assignments are presented in Table 2. A course fixed effect, whether or not the course was PH2 (as opposed to PH1), is included in the regression models.

In line with other studies of academic performance, conscientiousness is a very strong positive predictor of scores across all assignments and the final course grade. It is the only significant predictor for the composite weekly course scores. The standard deviation of overall course score is 17.077 percentage points. All else held equal, students who are higher on the conscientiousness dimension by one standard deviation (0.705 points) have an advantage by almost two letter grades.

Extraversion, agreeableness, and neuroticism are other significant predictors of the final course grade (Table 2). The directionality of these variables for major assignments is consistent in that higher levels of extraversion and neuroticism both positively predict assignment score while higher levels of agreeableness negatively predicts assignment score. Between the midterm and final exam, significance differs for each dimension. Extraversion is significant for the midterm, but not the final. Agreeableness is significant for the final, but not the midterm. These coefficients, however, are similar across midterm and final exam models, so it is possible that the lack of significance is due to the sample size. However, the coefficient for neuroticism is 0.175 for the midterm and close to zero (0.006) for the final. It is unclear why this discrepancy is present. Models predicting weekly quiz scores do not show any declining trends in the effect of neuroticism (Figure 2.2).

Openness is not a significant predictor across all scores. Point estimates, however are consistently negative for major assignments (Table 2). Coefficients are as large as -0.103 (midterm). No consistent trends were found for the weekly quiz scores (Figure 2.3).

**Personality in In-Person Courses**

The next analyses explore the same personality measures as predictors of in-person course performance, as measured by grade. Since students in the sample took a variety of courses, and since courses varied in difficulty, course fixed effects are included to effectively compare students within the same course. Some of the other courses that students took were also taught online. Results in Table 3 show three panels: all courses, in-person courses only, and online courses only. I exclude PH1 and PH2 courses from results in Table 3 since detailed scores are also provided in Table 2. Results with PH1 and PH2 included can be found in appendix Table A2.1.

Across all three panels, conscientiousness is again a strong significant predictor of grade (Table 3). The coefficient for the online courses (0.230) are twice as large as that for the in-person courses (0.098). The coefficient is even greater for PH1 and PH2 (0.432, Table 2).

Extraversion and agreeableness are both significant predictors of academic performance when considering all courses together (Table 3). However, most of the predictability can be explained from the online courses. Extraversion is a significant predictor for online courses, but not in-person courses (0.175 versus 0.020). Agreeableness is not a significant predictor for either online or in-person courses when separated, though this is possibly due to sample size. The coefficient for agreeableness in online courses is larger than the full sample model (column 1) and twice as large as the in-person model. Results for extraversion and agreeableness for other online courses are consistent with PH1 and PH2 findings in Table 2.

Neuroticism and openness are not significant predictors of academic performance (Table 3). Neuroticism coefficients are negative, but close to zero, across in-person and online courses. This finding contrasts that of PH1 and PH2 in Table 2, where neuroticism is found to be positive. The coefficient for openness is close to zero for in-person courses and negative for online courses.

Across all personality dimensions, only conscientiousness is a significant predictor in in-person courses. All other coefficients are close to zero, with agreeableness having the largest coefficient magnitude (-0.032). For online courses, extraversion and conscientiousness are both significant, and all other personality dimensions are slightly further from zero. Neuroticism has the smallest coefficient magnitude (-0.051), and it is the only personality dimension that does not align with findings for PH1 and PH2 in Table 2. The discrepancy between in-person and online courses suggests the possibility of a differential effect of personality across the two instructional mediums.

**Within Student Differential Effects**

To obtain the differential effect of personality on online courses, I include interaction terms between each personality dimension and whether or not a given course is taught in the online format. I include student and department fixed effects so as to compare students to themselves. Department fixed effects is used in place of course fixed effects since students do not usually take the same course twice and in different formats. Since courses within the same department tend to grade similarly, the department fixed effects addresses the issue of varying course difficulty. Results between course fixed effects and department fixed effects are similar (see appendix Table A2.1 for department fixed effects), with at most a 0.013 difference in coefficients.

Students in the sample earn higher grades in their online courses compared to their in-person courses (Table 4). This finding is surprising given that prior literature about performance in online courses compared to in-person suggested otherwise. However, the finding does align with the summary statistics (Table 1) which shows that students earned lower grades on average ("B-") compared to their course score in PH1 and PH2 (between "B-" and "B").

Since personality is a student-level variable, the main effects are accounted for in the student fixed effects. The interactions of conscientiousness with online instruction and openness with online instruction are both significant (Table 4). This suggests that personality does have a differential impact on instructional format. Namely students high on conscientiousness perform even better in online courses relative to their less conscientious peers. That also means that students low on conscientiousness perform even worse in online courses relative to their highly conscientious peers. The reverse is true for openness. Students who have a high degree of openness tend to perform worse in online courses relative to their peers. This finding aligns with the different coefficient signs across instructional mediums for openness found in Table 3 (close to zero but positive for in-person courses and negative for online courses).

To address concerns about changing course difficulty over time (e.g. students take harder courses later or students become more experienced in college coursework later), the second column of Table 4 reflects the same regression model for courses in Fall 2017 only. Conscientiousness is significant and similar in size to the first column. Openness decreased in magnitude and is no longer significant. In contrast, neuroticism grew from a coefficient close to zero (0.003) in the all courses model to a relatively large positive coefficient (0.122).

The decrease in openness can be explained by a possible time-varying effect (Figures 2.4 and 2.5). Given that the all-courses model includes courses *before* Fall 2017, grades in earlier courses reflect that of students with less college experience and other time-varying factors. To confirm this, Figure 2.4 shows the same openness and online interaction coefficient separated by course year. There is a decreasing trend as students are more advanced in college grade. The coefficient decreases from -0.057 for freshmen to -0.231 for seniors.

Figure 2.5 looks at coefficients from courses in Fall 2017 and each term in the prior academic year for students in their second year or above. Freshmen did not take courses in the 2016-17 academic year. There is also a decreasing trend from Fall to Spring in the 2016-17 academic year. Figure 2.5 shows a slight increase in trend from Spring 2017 to Fall 2017. It is possible that the increase is by chance, or that the increase is due to the summer break in between. The increase in Fall 2017, however, does not reach the level of what it was in Winter 2017.

Figures 2.4 and 2.5 taken together, as well as results from Table 4, suggest that openness has a negative differential impact in online courses and that it is more negative for students who have more college experience. Although Table 4 shows a large change in coefficient magnitude for neuroticism, Figures 2.4 and 2.5 do not show any time-varying trends for neuroticism.

## Discussion

This study looked at the impact of personality in online courses and whether personality had a differential effect on online versus in-person learning. It found that conscientiousness was an important predictor of academic performance and that conscientiousness did differentially influence online academic performance. Other personality dimensions were also important in the online format. Extraversion and agreeableness were important predictors in the online format but not in the in-person format when comparing between students. However, no differential effects were found when comparing students' in-person scores to their own online scores, suggesting an unobserved student-level characteristic. In contrast, openness was not significant when comparing across students in different formats, but it was significant when looking at differential impacts.

Contrary to the findings of Xu and Jaggars (2014), this study finds that students performed better in their online courses. There are multiple possible explanations for this. The first is that this study is conducted at a four-year public research university. As such, the students at the university are, on average, more academically prepared. A second reason is that this study picked specifically students who were enrolled in PH1 and PH2, which means both that the sample of students in this study is unique and that the two public health courses might have just been easier. Since PH1 and PH2 were entry-level courses, it is likely that course difficulty was not fully captured even when using department fixed effects. The two public health courses awarded, on average, a grade between “B-” and “B.” Average course grades in other courses were closer to “B-”.

The finding that conscientiousness is an important predictor of academic performance is not surprising as many prior studies have also found conscientiousness to be a strong and consistent predictor of academic performance. This study has confirmed that conscientiousness not only is a strong predictor of academic performance, but also that students who are less conscientious tend to do even worse online relative to their highly conscientious peers. Since conscientious encompasses terms like *reliable*, *hardworking*, *well-organized*, and *self-disciplined* (McCrae & Costa, 1987) it does not come as a surprise that students who are conscientious do even better in online courses.

Neuroticism is one of the less consistent personality dimensions. It is a positive predictor for PH1 and PH2. However, it is a negative predictor for all other courses. The jump from a nearly zero interaction coefficient when considering all courses together to a relatively large positive (though not significant) coefficient when considering only Fall 2017 courses is likely due to PH1 and PH2’s overrepresentation in Fall 2017. Neuroticism is a significant positive predictor for specifically the PH1 and PH2 courses, but not all courses in general.

Openness yields an unexpected finding given that prior meta-analyses found a weak positive association between openness and academic performance (O’Connor & Paunonen, 2007; Vedel, 2014). The weak positive association is only true for in-person courses in this study. For online courses, this study finds slightly stronger negative associations. The coefficient for openness is significant after controlling for student-level fixed effects. However, further analyses also show a time-varying factor. Over time, the differential effect of openness in online courses becomes more negative. Since this study shows a positive estimate for online courses, more college experience and a greater degree of openness will mitigate the positive estimate.

The time-varying effect of openness brings about another important issue—that this study assumes that personality is a fixed measure. However, preliminary test-retest work only showed that personality was stable, but not fixed. Thus, it is also possible that during students' time at the university, openness changed thereby explaining the time-varying pattern. Unfortunately, the current data set does not have personality information for students prior to Fall 2017 to examine this aspect. Given prior studies on the stability of personality (e.g. Cobb-Clark & Shurer, 2012), it is unlikely for personality to drastically change during the courses included in this study (four to five years at most for seniors and above).

Even though personality is difficult to change, it is still useful for schools to understand how personality impacts academic performance in different instructional mediums. Schools can leverage their understanding of the strengths and weaknesses of each personality type when designing online courses. For example, the finding that conscientiousness is important for academic performance, and even more so in online contexts, indicates a need to help less conscientious students in their studies. Knowing that they are less self-disciplined and disorganized, instructors may want to think about how information is presented and if the information is clear. Interventions, such as a scheduling intervention, may also help less conscientious students.

Although course performance was obtained for students in multiple departments across campus, the personality instrument was only administered to students who were enrolled in lower division public health courses. As a result, this study may not generalize to the whole population as there is almost certainly an underlying reason why students chose to enroll either of the two public health courses. In order for results to be generalizable school-wide, a study would need to randomly select enough students who take both online and in-person courses throughout campus and obtain their personality scores. Even then, however, those results may not generalize to other campuses. Nevertheless, this study does provide a better understanding regarding how personality applies to students in similar departments, and it also provides an initial step to how other unlike departments can investigate the role of personality or other individual differences across various instructional mediums.

# CHAPTER 3: IMPROVING ACADEMIC PERFORMANCE IN ONLINE COURSES: RANDOMIZED CONTROL TRIALS OF A SCHEDULING INTERVENTION

Online education is taking on an increasingly prominent role at postsecondary institutions. About 28% of postsecondary students take at least one online course, and about 14% of postsecondary students take only online courses (NCES, 2015). Both the number and proportion of college students enrolling in online courses have increased over the past few years as both traditional and for-profit universities are offering a greater share of credit bearing courses online. Given that nearly 70% of higher education academic leaders say that online education is part of their long-term strategy, online courses will continue to play an important role in higher education.

Online courses are an attractive means of instruction at colleges for a multitude of reasons, including its potential to expand enrollment counts with minimal investments in infrastructure and to slow the rapidly growing costs of delivering higher education (Deming, Goldin, Katz, & Yuchtman, 2015). Further, incorporating online courses and programs opens tertiary education opportunities to a broader population, particularly students unable to attend traditional colleges. In turn, increasing access to tertiary education also expands human capital development.

Unfortunately, online higher education also raises a number of concerns pertaining to persistence (Carr, 2000; Dutton, Dutton, & Perry, 2001; Lee & Choi, 2010; Levy, 2007; Tello, 2007), engagement (Conrad, 2002; Lyons, 2004; Singh & Pan, 2004), and learning outcomes (Jaggars & Xu, 2010; Taylor, 2003; Xu & Jaggars, 2011). Prior studies have documented alarmingly low persistence rates in online courses (Cochran, Campbell, Baker, & Leeds, 2014;

Evans, Baker, & Dee, 2016; Leeds, Campbell, Baker, Ali, & Brawley; 2013). While online course features play a role in students' persistence (Evans, Baker, & Dee, 2016), persistence is also affected by individual student characteristics, such as motivation, interaction in the course, proclivity towards self-regulated learning, and time management skills (Cochran et al., 2014; Evans, Baker, & Dee, 2016; Hart, 2012; Rostaminezhad, Mozayani, Norozi, & Iziy, 2013). Thus, the students who are likely to succeed in online courses are those with better study skills.

Although the risk factors in online courses are similar to those in traditional face-to-face courses, low study skills may interact in especially pernicious ways in the online setting. The lack of face-to-face connection with instructors and classmates, for example, inhibits student accountability to stay on track. It further detracts from the benefits of peer effects. The often asynchronous setting of online courses creates delays in communication. Additionally, the lack of designated course times also pose challenges to consistent scheduling as students may prioritize other urgent tasks or preferable activities.

Research is beginning to look into simple, low-cost, and scalable interventions aimed at improving online course outcomes. One such intervention is a scheduling device that facilitates student time management by prompting students to explicitly think about and state when they would work on a given online course. Prior studies used the scheduling device on a weekly basis. Baker, Evans, and Dee (2016) first examined the scheduling device on a science-based massive open online course (MOOC), but found negative effects on the number of lectures watched, course grade, and whether students earned a course completion certificate or not. Baker, Evans, Li, and Cung (2018) replicated the study in a five-week for-credit undergraduate course and found diminishing effects on weekly quiz scores with significant positive effects in the first week of the course and significant negative effects in the last week of the course.

Two studies are presented in this chapter. Study 3.1 examines the same online physics course as that examined by Baker et al. (2018) taught in a subsequent year. Study 3.2 examines two public health online courses taught during the regular school year. The physics and public health courses were taught at the same four-year selective public university. The two studies aim to provide empirical evidence for the efficacy of a scheduling intervention on undergraduate online courses. Both studies are guided by the same primary research question: what is the effect of encouraging students to schedule their coursework on academic performance in an online, for-credit postsecondary course?

The current work extends prior investigations of the scheduling device across two different studies (Studies 3.1 and 3.2) in four different ways. First, I monitored whether students followed up with their stated study time. Baker and colleagues (2018) awarded extra credit to students who scheduled their coursework regardless of whether students followed their schedule or not. In Study 3.1, students were not awarded extra credit unless they were active in the course within one hour of their stated time. Students were thus incentivized to schedule and work as scheduled.

Second, I adapted the wording of the device to prompt students to schedule the online coursework overall. Baker et al. (2016) and Baker et al. (2018) asked students to schedule when they would watch specific videos in the course. While Study 3.1 adopts the same wording (asking students to schedule each lecture video), Study 3.2 asks students to schedule when they would work on the online coursework. Since not all online courses have the same number of lecture videos, if any lecture videos at all, the prior device is not necessarily applicable to all online courses. By introducing slight variation in the wording, Study 3.2 examines a more flexible version of the scheduling device. Thus, the scheduling device was adapted to better fit the fact that the course did not have the same number of videos each week. In the modified scheduling device, students were given seven time slots to fill. They were not expected to fill out all seven slots.

Third, Study 3.2 is the first time that the scheduling device is implemented in a for-credit course taught during the regular school year. Prior studies have looked at the scheduling device in either a non-credit MOOC course or a for-credit summer course. Although online summer courses are popular among universities, structural differences between summer courses and regular-term courses warrant a need to obtain empirical evidence for both. Summer courses, for example, are typically shorter duration, but also more intensive in terms of the weekly work load. Students typically do not take as many courses at the same time in the summer, and enrolling in summer is not required to be a degree-seeking student at the university. Thus, those who enroll in the summer either need recovery credit, are motivated to get ahead, or have a different unique reason for doing so. Students' work load potentially varies in the summer depending on the number of courses they choose to enroll in. Students during the regular school year typically are degree-seeking students of the university with a minimum course load requirement. Thus, during the regular school year, students are tasked with juggling responsibilities across multiple courses.

Finally, both Studies 3.1 and 3.2 prolong the administration of the scheduling device. The study by Baker et al. (2018) raised concerns about withdrawal effects after seeing decreasing effect sizes. This study eliminates the possibility of withdrawal effects in the given course by extending the scheduling prompt to the full duration of the course. Students were asked every week when they would work on the online courses. In Study 3.1, students were prompted to schedule every week for the full five-week course duration. In Study 3.2, students were prompted to schedule every week for the full eleven-week course duration.

# Prior Literature and Theory

## Academic Performance in Online Higher Education

The growing role of online education at postsecondary institutions has generated a lot of interest in the effects of online courses on student academic performance. A long line of meta-analyses have attempted to uncover the effect of online instruction on academic performance (e.g. Bernard et al., 2004; Lack, 2013; Means, Toyama, Murphy, Bakia, & Jones, 2010). These meta-analyses have made some progress in determining factors that contribute to different findings across empirical studies, but they also highlight the complex nature of determining the effect of online instruction compared to in-person instruction. Lack (2013) focused on undergraduate for-credit courses and found mixed results. One major reason for inconsistent results was that many studies comparing online to in-person instruction were observational and faced selection bias.

A number of studies aiming to obtain causal estimates of the effects of online courses have made improvements to earlier observational studies. Xu and Jaggers (2013) compared community college students' performance in online courses to their own performance in in-person courses. They found that students generally performed worse in online mediums. In one randomized control trial, Bowen, Chingos, Lack, and Nygren (2014) randomly assigned students to either a hybrid or an in-person statistics course. They found no significant difference in learning outcomes between the two mediums. In a different randomized control trial, Figlio, Rush, and Yin (2013) found that there was a significant difference between students who were randomly assigned to an online economics course with compared to students who were assigned to the in-person version of the same course. In particular, male, Hispanic, and lower achieving students performed worse in the online economics course.

While research in online education still yields largely mixed results, the negative findings for community college students (Xu & Jaggars, 2013) and for lower achieving students (Figlio et al., 2013) suggest an important concern for underprepared students that other studies do not account for (Frantzen, 2014; Xu & Jaggars, 2011). This is especially true given that most postsecondary leaders agreed with the statement that "students need more discipline to succeed in an online course than in a face-to-face course" (Allen & Seaman, 2006). Given that universities are increasingly turning to online courses, it is important to explore interventions targeting students who lack the discipline needed to succeed in an online course.

**Time Management**

Time management is an important skill for academic success. Prior research has demonstrated the importance of time management skills on course performance in both online and in-person college courses. Time management and study hours have been found to be leading predictors of academic performance (Beattie, Laliberté, Michaud-Leclerc, & Oreopoulos, 2017; Brint & Cantwell 2010; Stinebrickner & Stinebrickner, 2004, 2008). Several studies have found a negative correlation between performance in online courses and procrastination (Elvers, Polzella, & Graetz, 2003; Michinov, Brunot, Le Bohec, Juhel, & Delaval, 2011). Unfortunately, many students spend a low amount of time studying (Babcock & Marks, 2011; Beattie et al., 2017) while a high proportion engage in procrastination (Steel, 2007). Hartwig and Dunlosky (2012) found a positive correlation between regular studying, rather than cramming before a deadline, and college GPA. Macan, Shahani, Dipboye, and Phillips (1990) similarly found a positive correlation between scores on a robust time management scale and a number of other variables: college GPA, self-perceptions of performance, and general satisfaction with life.

Additionally, students with better time management skills demonstrated higher cognitive test scores and greater efficiency with study time (van Den Hurk, 2006).

In the context of online education, many students and instructors agree that online courses require greater time management skills, in addition to personal responsibility and motivation, compared to in-person courses (Allen & Seaman, 2006; Bork & Rucks-Ahidiana, 2013). Many online courses tend to be asynchronous and typically have little to no predetermined class meeting times. Thus, students are tasked with individually setting aside time to learn course materials.

Implementing an intervention to help students with their time management could benefit students' academic learning. One study showed that students credited their success in an online course to the act of developing a time management strategy (Roper, 2007). There are several reasons why proper time management would be positively correlated with academic outcomes. First, time management could induce students to study at optimal times and reduce the chance that students will work during off hours. Studies have shown that students who worked during their preferred time of day and students who started classes later in the morning tended to have better academic performance (Goldstein et al., 2007; Carrell, Maghakian, & West, 2011). Second, planning could prevent students from waiting until a deadline to complete work. Cramming and procrastination have been found to be negatively correlated to academic performance (Elvers et al., 2003; Michinov et al., 2011). Third, time management could reduce levels of stress and anxiety, which in turn could improve academic performance (Misra & McKean, 2000).

### Scheduling Intervention

One critical difference between a traditional in-person course and an asynchronous online course is that the latter lacks scheduled class meeting times. Although online courses typically have regular deadlines for submitting assignments, students are ultimately in charge of when and how regularly they would like to work on course materials. Therefore, students can choose to work on new lecture materials at regularly scheduled times throughout the week (e.g. Mondays and Wednesday at 1 pm) or they can choose to study work on new lecture materials a few minutes before a homework deadline.

The goal of the scheduling intervention in this study is not to force students to study at set times, as that would detract from the flexibility of an online course. Rather, the intervention aims to preserve the flexibility of online courses while simultaneously encouraging structure and timeline. By having students think about the near future and commit scheduled days and times for their online coursework, students should be more likely to hold themselves to that schedule. I hypothesize that scheduling will help students' academic performance. I further hypothesize that incentivizing students to study as scheduled will have an even greater effect on their study behavior and academic performance.

The scheduling intervention used in this study is similar to a precommitment device, which has been found to be effective in various contexts such as employee effort (Kaur, Kremer, & Mullainathan, 2015), smoking (Giné, Karlan, & Zinman, 2010), and savings (Ashraf, Karlan, & Yin, 2006). A precommitment device works by binding a person's future behavior to reduce the risk of other preferences taking over when the future comes. People's preferences are constantly changing. As a result, people are often engaged in a different activity from what was previously planned because a different set of preferences took precedence (Frederick, Loewenstein, & O'donoghue, 2002). For example, a student may intend to work on an online class every evening. However, the student may end up working only on the evening of the deadline because other more attractive activities arose each night. A precommitment device would help students commit to and formalize a study schedule with the larger goal of observing overall coursework commitment.

In this chapter, there are two levels of consequences for missing scheduled study times. In Study 3.1, students lose out on an extra credit opportunity for missing scheduled study times. In Study 3.2, students are not penalized for missing scheduled study times, but may instead face self-scolding. Although the consequences for not studying as scheduled may be considered light, the act of explicitly planning out and stating a study schedule may be sufficient to see an improvement in student performance. The intervention in this study further addresses prior calls for interventions targeting the development of time management strategies in online courses (Nawrot & Doucet, 2014; Song, Singleton, Hill, & Koh, 2004).

Several studies that look at a time management interventions in online higher education exist. However, most of these studies were conducted in the context of MOOCs (Baker, Evans, & Dee, 2016; Kizilcec, R. F., Pérez-Sanagustín, M., & Maldonado, 2016; Patterson, 2014). Findings in the context of MOOCs have been mixed. Only one has examined the scheduling device in the context of a credit-bearing online higher education course (Baker et al. 2018). However, the study was conducted in the summer term when the course duration was shorter (though more intensive) and when the student population differed from the regular school year. The study only implemented the scheduling device in the first two weeks of the five-week course, and it found positive effects only in the first week followed by declining to negative effects in the later weeks.

This study improves on prior literature by examining the scheduling device in two contexts: a summer course with the device administered for the full course duration and a course taught during the regular school-year with the device administered for the full course duration. The study makes one further improvement by incentivizing the device and only awarding a nominal amount of extra credit for students who study as scheduled in the summer course.

## Study 3.1

### Context

**Course.** This study was conducted in an online, undergraduate physics course taught over the summer at a selective, public four-year university. The course was five weeks in duration and conferred full credit as if enrollees had taken the same ten-week course during the academic year. The course was a required lower division course for students in the major. Students taking the course were required to take calculus concurrently if they had not taken calculus prior to the start of the term. The course was taught entirely online and met only for a final exam.

Students had frequent deadlines to meet. Each week included five lecture videos, six quizzes, daily homework assignments, and weekly "challenge problem sets." Each lecture video was approximately 40 to 50 minutes in duration. Each lecture video had an associated quiz, making up five of the six quizzes administered in a week. Students received credit for watching the lecture videos and taking the lecture video quizzes if they completed the quizzes by Friday at midnight each week. The sixth weekly quiz accounted for a larger proportion of students' grade, and it was only available on Sundays. For clarity, I will refer to the lecture video quizzes as "video quizzes" and the Sunday quiz as the "weekly quiz." Daily problem sets were due every day from Monday to Saturday in addition to one or two challenge problem sets due each week. In addition to the weekly tasks, the course also had a final exam on the last day of the course.

The lecture videos, quizzes, homework assignments, problem sets, and final exam all counted towards students' grades.

**Participants**. A total of 150 students enrolled in the course, but due to several students enrolling late, only 143 students were randomly assigned into treatment (N=71) and control (N=72) groups on the first day of the course. Two students from the control group and three students from the treatment group dropped out of the course and thus were excluded from the analysis. Another seven students were also excluded because they had been randomly selected from the control group for another time related study. Among the 131 students in the potential sample, 113 of them completed the pre-course survey and 105 had full demographic data available. A large proportion of students were missing demographic data because they were not regular degree-seeking students of the university. Students outside of the university were allowed to enroll in courses during the summer. In total, 97 students had both pre-course survey and demographic data available. However, one student was excluded as the student never logged in to the online course. The analysis sample included the 96 students who had both pre-course survey and full demographic data available.

## Data

**Demographics.** The institution provided administrative data on demographic data for all students enrolled in the course who were also enrolled as degree-seeking students in the institution. I use the data to check for balance across the treatment and control groups, and I also include demographic variables in the treatment effect regressions.

We use demographic background to check for balance across treatment and assignment groups. Table 1 presents demographic summary statistics for students in the analysis sample, control group, and treatment group. Students in the analysis sample were on average 20.2 years old (SD 0.85). A little more than half of the students were female (58.3% compared to 41.7%). Underrepresented minorities made up approximately 7.3% of the students. Most students has just completed their sophomore year (74.0%). Recent freshmen made up 15.6% of the course, juniors made up 9.4% of the course, and seniors made up 1% of the course. Incoming freshmen were excluded from this study as demographic information was not readily available for this group. About 30.2% of the students came from low income households while 35.4% were first generation college students (neither parents attended college). Most students were living off-campus but nearby for the summer (68.8%). Others were commuting from their parents' home (22.9%) or commuting from more than one hour away (8.3%).

Individual t-tests were conducted between the control and treatment groups to assess balance across randomization. Only age appeared to be statistically significant with the treatment group being older by approximately half a year. This difference corresponds with the slightly higher proportion of juniors and seniors in the course. There were nine juniors or above in the treatment group and two juniors or above in the control group. This difference does raise a concern for balance across the two groups. Given the small proportion of students, I proceeded with the analyses. However, I also checked the consistency of results with the juniors and seniors excluded from the sample.

**Time Management and Self-Regulation Skills.** A pre-course survey was administered to students to capture students' self-perceptions of their own time management and self-regulation skills. The pre-course survey adapted questions from a widely used and validated measure of students' self-reported self-regulation and self-management skills (Pintrich, 1991). In addition, it also included a set of widely used online readiness assessment questions (Williams, n.d.). Students responded to questions on a Likert scale from 1 (strongly disagree) to 5 (strongly agree). The pre-course survey also included one question about students' time commitment to the course (measured in hours). I include the items in the treatment effect regressions to increase the estimate precision.

This study distinguishes between students' actual time management abilities and their perceived time management abilities as the two are not necessarily correlated (i.e. Dunning–Kruger effects, Kruger & Dunning, 1999). The data in this study suggests very little to no correlation between survey responses and actual click activity. However, those who report low planning skills tended to wait until the last minute to start of coursework more. They were also less likely to study as scheduled.

Pre-survey responses are used as additional checks of balance across treatment and control groups. Table 1 presents summary statistics of pre-course survey responses for both control and treatment groups. Students expected to spend about 11 hours on the course each week. The estimate was slightly higher for control students, though the difference was not statistically significant ($p = 0.374$). Overall, students reported high levels of self-regulation abilities, with most variables averaging close to four (agree). Three self-regulation items significantly differed between the treatment and control group. The control group indicated higher advanced planning skills and higher persistence through dull or unfavorable work. While it is possible for significance to occur by random chance due to multiple testing, I note this difference between the treatment and control groups. I include the pre-course survey items as controls in this study's models. Since students were randomly assigned, including pre-course survey items should yield more precise estimates without drastically changing the effect size.

**Course Performance.** I observe achievement outcomes (weekly quiz scores, daily homework scores, final exam score, and final course grade) as well as video click stream data from the course management platform. This study's analyses focus on weekly quiz scores as opposed to daily homework scores for two reasons. First, the weekly quizzes are most closely aligned with the content presented in the lectures. Second, the intervention is aimed to affect weekly, not daily, scheduling, so any effects should be most present in weekly assignments (Koch & Nafziger, 2017). I also include final exam and final course grade in the main analyses as these measure the student's overall retention of course materials and overall performance in the course.

Table 2 presents summary statistics of students' scores on their weekly quizzes, final exam score, and their overall course grade. Students averaged between four to five points out of six on each of the weekly quizzes. They also earned about 49 points out of 75 (65.0%) on the final exam. Students' overall course reflected their weekly quiz scores, final exam score, completion of the video quizzes, and completion of homework assignments. On average, students earned a "B-" in the course.

**Course Engagement.** I include clickstream data from the course management platform to assess student engagement, including when students viewed videos and completed lecture video quizzes during the week. These behavioral measures allow me to examine potential mechanisms affected by the treatment that could explain effects of the treatment on student academic performance. I specifically investigate students' procrastination and cramming behavior, as measured by how much earlier (time remaining, on average across the five weekly videos) and how much time fell in between videos (the standard deviation across the five weekly videos) before the Friday midnight deadline. For clarity, I will refer to these two measures as the time remaining and spacing, where higher values meant that the student started earlier or had more time in between videos.

## Experimental Design

The analysis sample included 48 treatment students and 48 control students. On the Monday of the first week of the course and Fridays every week, all students received an e-mail with a survey link (which was separate from the pre-course survey) from the course instructor. Treatment students received an e-mail with a link to an online scheduling survey asking them to schedule the day of week and time that they planned to watch each of the five lecture videos for that week (see Figure 3.1 for a screenshot of the scheduling survey). Control students received an e-mail with a link to a theoretically inert survey containing eight items from the Big Five Inventory (see Chapter 2). Both treatment and control groups received the same number of e-mail contacts from the instructor.

Treatment students were asked to schedule an exact day and time of day for each video. Students were given five time frame options: (a) morning between midnight and noon, (b) early afternoon from noon to 2 pm, (c) late afternoon from 2 pm to 5 pm, (d) evening from 5 pm to 8 pm, and (e) night from 8 pm to midnight. Although an uneven amount of hours were allocated to each time frame, this division was determined by results from a prior study (see Baker et al., 2018) and aimed to have a roughly even amount of schedules for each block. That is while the morning session had the most hours allocated (twelve hours), about the same number of videos were scheduled before noon as between noon and 2 pm.

Treatment and control students were offered a nominal amount of extra credit for completing the weekly surveys. Treatment students were given the additional task of needing to watch videos at the time and day scheduled in order to receive credit. Out of five possible extra credit points each week, treatment students were awarded one point for each video watched within their scheduled time. Students were therefore “monitored” for following up on their scheduled times. Control students only needed to complete the weekly survey in order to receive the maximum possible extra credit points for any given week.

Except for the first week, treatment and control weekly survey uptake was high. In the first week of the course, 32 out of 48 (66.7%) treatment students completed the treatment survey while 34 out of 48 (71.0%) control students completed the control survey. In the following weeks, weekly survey uptake ranged between 83.0% and 92.0% for both the treatment and control groups. Only about four to eight students in each group missed the weekly survey each week. Both groups had approximately similar rates of uptake.

**Method of Analysis**

We use linear regression to estimate the effect of treatment assignment on students' course time as well as several performance outcomes: weekly quiz, final exam, and overall course grade. While this study is specifically interested in students' academic performance, I investigate course time as a potential mechanism for treatment students' differential academic outcomes.

I also measured when students submitted video quizzes by looking at clickstream timestamps. I focus on video quiz submission times rather than video viewing time because of the ambiguous nature of video clicks (i.e. whether the first click was intentionally to watch lecture videos, exploratory, or accidental). I define spacing as the standard deviation of time when students submitted their video quiz. A higher "spacing" value means that students had more time in between each video quiz submission. I define procrastination as the average time remaining before the Friday midnight deadline for each of the five video quizzes. A higher "time remaining" value means that students started the video quiz earlier in the week. Students who did not complete the video quizzes on time or at all were considered to have submitted the video quiz right at the deadline.

We standardize all outcome variables in the linear regression models. I also include 18 student-level covariates, $\theta_i$. Covariates include eleven pre-course survey items and seven demographic background items. The regression equation for the intent-to-treat effect (ITT) for student *i* is as follows:

$$Y_i = a + \beta treatment_i + \gamma\theta_i \tag{1}$$

We estimate the ITT in the models rather than the treatment-on-the-treated (TOT), which gives the effect of actually completing the scheduling survey, since there are multiple ways to define treatment uptake in this experiment (i.e. completing any one of the weekly surveys, completing the majority of the weekly surveys, completing all of the weekly surveys). The ITT estimate yields the most conservative estimate, where treatment uptake could be defined as receiving the treatment survey.

## Results

**Watching as Scheduled.** Among all videos that were scheduled, approximately half were watched as scheduled (Figure 3.2). A video is considered to be "watched as scheduled" if students clicked to watch a specific video during the time that they indicated on the survey (e.g. Monday in the early afternoon). The actual proportion varied greatly for each of the 25 videos. The third video during week 1 had the highest percentage of treatment students who watched as scheduled (68.2%), while the second video during week 5 had the lowest percentage of treatment students who watched as scheduled (29.1%). Between 3.5% and 23.8% additional students watched each video before scheduled. Figure 3.2 shows the proportions of scheduled videos that were watched as scheduled. While the proportion watched as scheduled for each video within a given week varied slightly, students tended to watch videos as scheduled more in the last two weeks of the course.

**Treatment Effect on Course Time.** The distribution of days and times when treatment students watch course videos is different from the distribution of days and times when control students watch course videos (Figure 3.3). The KS-test, which tests for the difference between two distributions, suggests a statistically significant difference between the two groups' video watching times ($p<0.001$). Control students tend to have a single peak in their first-time video-watching distribution on Friday mornings, whereas treatment students tend to have a bimodal distribution with peaks on Monday evenings and Friday mornings. Incidentally, Monday evening was a popular time to schedule videos. Although students did not necessarily watch videos at the time scheduled (KS-test $p < 0.001$), the bimodal nature of treatment students' video-watching time suggest some influence from scheduling.

We examine the treatment effect on spacing and time remaining before the video quiz deadline on Friday midnight (Table 3). With the exception of the first week of the course, I find overall positive effects on spacing and completing work early. Estimates are significant for the later weeks of the course, with effect sizes as high as 0.587 (week 4) for working in advance and 0.707 (week 4) for spacing. This equated to treatment students working approximately 20 hours earlier than control students in week 5 and treatment students spacing work out by 12 standard deviation hours more than control students in week 4.

The higher effect size estimates of time remaining and spacing in weeks 4 and 5 are consistent with the high proportion of videos watched as scheduled (Figure 3.2). Likewise, the first two weeks saw both a low proportion of videos watched as scheduled and a nonsignificant treatment effect. One possible explanation for a lower proportion of videos watched as scheduled and nonsignificant treatment effect in the first few weeks could be that students needed a better understanding of the course structure and of their own schedule for the term. Figure 3.4 shows the contrast between scheduled and actual video-watching times in weeks 1 and 5. Week 1 tended to have more videos scheduled in the morning, night, and late afternoon. However, a higher proportion of videos were actually watched at night, as opposed to other times during the day. Across both treatment and control groups, not as many students actually watched videos in the late afternoon. Week 5 tended to have most videos scheduled at night, which falls in line with when students have been watching lecture videos throughout the term.

**Treatment Effect on Academic Performance.** Seeing an effect on times that students work, I turn to examine whether the treatment similarly impacted academic performance. I use three different measures of course performance: weekly quiz scores, final exam score, and overall course grade (Table 4). In contrast to the findings of Baker and colleagues (2018), the current study does not find any significant effects in the first week's quiz scores. The study also does not find significant effects for the remaining four weeks or any trends in effect size for the treatment group relative to the control group.

In addition, non-significant effect sizes were found for treatment students' final exam (0.046) and overall course score (0.197). Null findings suggest that if there were any kind of effect of the treatment on final exam or course score, it was very small and could not be detected given the current sample. Thus, while the scheduling treatment with monitoring was effective for starting coursework early and spacing out coursework, it did not change academic performance as expected.

Further investigation of students who watched videos as scheduled (as opposed to not scheduling or not watching as scheduled) shows that watching as scheduled tended to yield higher, but not significant, quiz and final exam scores (Table 4). For each weekly quiz score, I ran a regression model that used the number of videos for the given week that were watched as scheduled. For final exam and course grade, I ran a regression that used the overall number of videos watched as scheduled. Findings were significant only for the final course grade, which had a small effect size of 0.047. Effect size estimates are mostly positive for other academic performance scores, with the exception of week 3. Positive findings are expected and consistent with prior literature suggesting a correlation between time management skills and academic performance (Beattie, Laliberté, Michaud-Leclerc, & Oreopoulos, 2017; Brint & Cantwell 2010; Stinebrickner & Stinebrickner, 2004, 2008). However, a lack of significant effect on academic performance despite changes in starting early and spacing suggests that cramming and procrastination does not directly impact academic performance.

## Study 3.2

### Context

**Course.** This second study was conducted in two online, undergraduate public health courses taught during the academic year at the same four-year public institution as Study 3.1. Like other courses at the institution, the courses were ten weeks in duration with an additional week for final exams. The courses were required lower division courses for students in the public health major. One course, which I will call PH1, was a pre-requisite for the other, PH2. PH1 did not have any pre-requisites. Both courses were taught by the same instructor and had the same assignment structure despite different content. The courses were taught entirely online and did not meet in-person. Exams were administered online, scheduled by the student, through an online proctoring service.

Each week, new course materials were released on Monday morning at 8 am. Students had three different deadlines (Wednesday, Friday, and Sunday night) to meet for several different assignments, including viewing course videos, writing two discussion forum posts (worth 9% of the final grade), and completing a short response assignment (9%). Students additionally had to complete seven online quizzes (2% total), a peer-reviewed presentation (6%), research paper (24%), midterm exam (25%), and final exam (25%). The instructor estimated that students should spend about ten hours each week for course assignments. The number of course videos varied each week but required approximately 2 hours of students' time. It included both lecture videos from the instructor as well as documentary videos from other sources. Students were not graded for watching the lecture videos but were expected to do so given that the content of the videos related to the discussion forum and short response assignment prompts, as well as course quizzes and exams.

The two discussion forum posts required of students each week were due on Wednesday and Sunday night. Students were assigned to discussion forum groups. Each student had to make an original discussion forum post in response to a prompt by Wednesday night. Students then had to respond to a group member's post by Sunday night. The instructor provided guidelines and a rubric for completing the discussion forum posts on the course syllabus.

The short response assignment had a similar guideline and rubric also stated on the course syllabus. Students had one short response assignment to complete each week. The short response assignment was due on Sunday night at the same time as the second discussion forum post.

In addition to written assignments, students in both public health courses had to take a 30-minute quiz almost every week. On weeks that did not have a quiz, students either had to complete either a midterm exam (week 5), peer-reviewed presentation (week 6), or research paper (week 8). The peer-reviewed presentation was a 5-minute recorded PowerPoint lecture uploaded to a designated course page. The quizzes and midterm were available all day on Friday and had to be submitted by Friday night. The peer-reviewed presentation and research paper also had the same Friday night deadline. Since the presentation was peer-reviewed, students additionally had to submit an evaluation by Sunday night of the same week. The final exam was available on Monday all day during the institution's finals exam week (week 11).

**Participants**. A total of 401 students were randomly assigned into treatment and control groups either on the first day of the term or on a rolling basis during the university's add/drop period. By the end of the second week of the term, the university's add/drop deadline, a total of 228 students were enrolled in PH1 and 96 students were enrolled in PH2. The remaining 77 students, of those randomly assigned, dropped the course before the add/drop deadline. At the university, it is common for students to switch courses in the initial two weeks of the term. The treatment group included 162 students: 111 students from PH1 and 51 students from PH2. The control group included 162 students: 117 students from PH1 and 45 students from PH2. Among the 324 students in the sample, 235 of them completed all pre-course survey items related to self-regulation and self-management. Given the large reduction in sample size, both analyses with the full student sample (N=324) and analyses with the reduced student sample are reported. In the full student sample analyses, only demographic characteristics are included covariates. In the reduced student sample analyses, both demographic characteristics and self-reported study skills are included as covariates.

## Data

**Demographics.** As in Study 3.1, I use a similar institutional data set to capture students' demographic background and prior academic achievement. I check for balance of demographic and academic background variables across the treatment and control groups to ensure successful randomization. I also include demographic and prior academic achievement as covariates in the treatment effect regressions.

Table 5 presents summary statistics of demographic and prior academic achievement for students in the full sample, control group, and treatment group. Students in the full sample were on average 20.20 years old (SD 1.92). Almost three-quarters of the students were female (74.1% compared to 25.9%). Approximately one-third of the student were underrepresented minorities. Most students enrolled in the course are in their sophomore year (53.1%), followed by freshmen (20.4%), juniors (16.7%), and seniors or above (9.9%). About 40.1% of students came from low-income households and 38.3% came from households that do not speak any English. About half of students in the course were first-generation college students.

Individual t-tests were conducted between the control and treatment groups to assess balance across randomization. None of the variables appeared to be statistically significant, which provides confidence that randomization produced equal treatment and control groups.

**Time Management and Self-Regulation Skills.** The pre-course survey administered in PH1 and PH2 was very similar to the pre-course survey administered in Study 3.1. I used the same self-reported self-regulation and self-management items (Pintrich, 1991; Williams, n.d.) to test for balance across treatment and control groups. I also included one question about students' time commitment (measured in hours) and a set of questions from a widely used and validated measure of students' self-reported self-regulation and self-management skills (Pintrich, 1991;

Williams, n.d.). Students responded to self-regulation questions on a Likert scale from 1 (strongly disagree) to 5 (strongly agree). Self-regulation items are used as an additional balance test across treatment and control groups. I also include the items in the treatment effect regressions to increase the estimate precision.

Table 5 presents summary statistics of pre-course survey responses for both control and treatment groups. On average, students expected to receive a between an “A” and “A-” in the course. Most students expected to receive an “A” while only five students expected to receive a “B-.” None of the students expected a “C+” or below. Students expected to spend approximately 9 hours on the course per week. Overall, students reported high levels of self-regulation abilities. Their Likert scale scores averaged close to four (agree). None of the pre-survey items significantly differed between the treatment and control group, giving added confidence that randomization between treatment and control groups successfully yielded equivalent treatment and control groups.

**Course Performance.** I observe students’ overall course grade, as well as the different scored components that make up the course grade. Graded components include students’ weekly discussion forum and short response assignments, quizzes, midterm and final exams, paper, and presentation. I focus the analyses on students’ overall course grade, though I also examine the various assignment scores that constitute the overall course grade as the intervention may have impacted weekly tasks or overall retention of course content.

Table 6 presents summary statistics describing students’ scores on their weekly quizzes, midterm exam, final exam, research paper, course presentation, and their overall course grade. I also include summary statistics for students’ overall homework and discussion forum grade, both of which were assigned weekly. Almost all students received full credit on the weekly quizzes.

On average, students earned a “B” (50 points out of 60) on the midterm exam and a “C” (45 points out of 60) on the final exam. Students also earned grades of “B” and “A-” on the research paper and course presentation, respectively. On average, students earned a “B” in both their weekly written homework assignments and discussion forum participation, as well as in their overall course grade.

**Experimental Design**

A total of 246 students enrolled in PH1 on the first day of the term were randomly assigned into treatment and control groups. Another 103 students enrolled in PH2 on the first day of the term were also randomly assigned. I additionally randomly assigned students who enrolled late, on a rolling basis. In total, 46 students enrolled in PH1 and 6 students enrolled in PH2 at a later point during the university’s add/drop period, the first two weeks of the course.

All students were given the same time frame to complete either the treatment weekly survey or control weekly survey, as assigned. Weekly surveys were made available on Monday morning and closed on Wednesday evening, which also aligned with when weekly course materials were made available and the first assignment deadline of the week, respectively. Links to the weekly surveys were posted on the course webpage. Treatment students had access to links that led them to an online scheduling survey asking them to schedule the day of week and time that they planned to work on their public health course. Unlike Study 3.1, which had students choose from a general time of day, this survey asked students to pick start and end times in half-hour increments. The scheduling survey had space to schedule up to seven study sessions. Control students had access to links that led them a theoretically inert survey with questions about their personality. Both treatment and control students were offered a nominal amount of extra credit for completing the weekly surveys, in addition to the pre-course survey. Extra credit information, including links and deadlines, was posted at the top of the course webpage alongside the general course information section. Students were able to access the survey links on the main course page.

Treatment and control weekly survey uptake was moderately high over the eleven week duration of the course (Figure 3.5). Uptake percentage ranged between 55% and 89%. Survey uptake saw an increasing trend over the first half of the school term. It peaked at 89% for the control group and 85% for the treatment group during the sixth and seventh weeks of the course, which also corresponds to the weeks following the midterm exam. Uptake between the treatment and control groups were similar across all eleven weeks.

**Method of Analysis**

I use a similar linear regression model as Study 3.1 to estimate the effect of treatment assignment on students' overall course grade, as well as the different graded components (e.g. final exam grade, quiz scores, presentation score) that made up the overall course grade. I also examined individual assignment scores. However, for reporting, I condense each weekly assignment type into one overall score as assignments were similar each week. I therefore use one overall summed score for the ten short response assignments. I also use one overall summed score for the ten discussion forum scores.

As in Study 3.1, I standardize all outcome variables in the linear regression models. I include the same 18 student-level covariates, $\theta_i$, as those found in Study 3.1. In addition, I add one additional course-level covariate to account for any differences between PH1 and PH2. The regression equation for the intent-to-treat effect (ITT) is the same as Equation (1) found in Study 3.1.

## Results

**Studying as Scheduled.** Treatment students tended to schedule study sessions in the afternoon or early evening. However, they often did not work on the course until the late evening (Figure 3.7). Furthermore students tended to schedule study sessions uniformly on weekdays. However, their click activity suggest that they were more active in the evenings on Wednesdays, Fridays, and Sundays when the course had an assignment due.

A very low proportion of students actually studied as scheduled. With the exception of week 11, between 13.4% and 17.2% of scheduled study sessions were met each week. Students met their scheduled study session if they made any single click within their stated start and end time. On average, students scheduled a little over four study sessions. Very few students worked on the online course as scheduled, and it was unlikely for students to have met all scheduled study sessions. In Week 11 46% of scheduled sessions were met, most likely because students only needed to worry about the final exam that week, and the final exam was administered early in the week. On average, students scheduled two study sessions during week 11. They met approximately half of the scheduled sessions.

**Treatment Effect on Course Time.** I examine whether the treatment had any effect on when students were active in the course. Table 7 reports the ITT estimates on the earliest time of week that students made their initial click, the spread of students' clicks (as measured by the standard deviation of clicks each week), and the earliest time of day that students made their initial click. Though not presented, I also check results for the latest time of week, latest time of day, average time of week, and average time of day and found consistent results. In general, the treatment did not have an effect on how early and how spread out students worked on the course throughout the week. Only the spread was significant during week 5, though this finding is likely due to chance as all other weeks do not show consistent positive or negative trends.

The treatment also did not have a clear significant effect of the time of day that students worked. However, there is consistent positive estimate for all but two weeks (weeks 3 and 4). The estimate for week 2 (0.201) is marginally significant. The positive findings are mirrored in Figure 3.6 where there is a slightly higher proportion of treatment students active earlier in the day on Monday, Tuesday, Thursday, and Saturday. These days also happen to be days that assignments were not due. For the most part, however, treatment students' study hours were similar to that of the control groups even though they scheduled to work on the course earlier in the day. Late evening, or night, was a popular time to work on the course for both treatment and control students. This was particularly true for days when assignments were due.

Overall, Figure 3.6 shows a high degree of overlap between treatment and control students, which starkly contrasts Figure 3.3 from Study 3.1. Like the scheduling device found in Baker and colleagues' (2018) study, it is likely students are minimally affected by the intervention without any incentive to follow-up as scheduled. Another likely explanation for the lack of an effect is that the course was highly structured and students had assignments due almost every other day. Control and treatment students were therefore both incentivized to work early and throughout the week. Another possible explanation could be due to the fact that students were not prompted to schedule until the start of the week. Thus, students did not have an initial start session to encourage starting on coursework early.

**Treatment Effect on Academic Performance.** I next investigate if the treatment had any effect on students' academic performance. I find a positive effect of the treatment on overall course grade (0.199, Table 8). This remains significant after controlling for pre-survey covariates for the reduced sample (M2). The effect size (0.199) suggests that the treatment boosted student scores by slightly more than one-fifth of a grade point (grade SD = 1.057). Although insufficient to boost all students' course grades, the treatment helps bring students who are near the cut-off up to the next letter grade.

Since the overall course grade is made up of different assignments and course requirements, I investigate which component of the overall course grade benefitted from the treatment. Table 8 additionally reports ITT results for the assignment components that make up the course grade: weekly quizzes, overall assignment, overall discussion, research paper, presentation, midterm exam, and final exam.

Given that the treatment was administered on a weekly basis, one hypothesis might be that the weekly assignments would be impacted by the treatment. However, a significant positive effect is only found for the week 2 and week 4 quizzes. All other weeks, except week 10, are positive. This consistent trend suggests promising results for the treatment on weekly quizzes. However, if there is a positive effect on weekly scores, it is too small to detect with the current sample. The treatment also does not yield statistically significant results for the combined quizzes. Thus, the quizzes alone are insufficient in explaining the positive effects on quiz scores. Similar results are found for the weekly homework and weekly discussion participation.

Another hypothesis might be that the larger and theoretically more time-consuming one-time assignments would be impacted by the treatment. On paper and presentation assignments, the treatment has generally zero to positive effects. The treatment does not appear to have any impact on students' presentation assignment, which demanded a lot of time on students' part to create a slide presentation with audio-recording. The treatment does seem to have a small, but non-significant, positive effect on students' research paper performance. The effect size 0.065

suggests that if the treatment did give students an advantage in their research paper grade, it would increase paper scores by about 1.2 percentage points.

The largest effects, however, are seen for the midterm and final exams. The most conservative effect size estimates (0.141 for midterm and 0.230 for final exam) equate to the treatment increasing midterm and final exam scores by approximately 2.5% and 4.4% respectively. These estimates remain significant even after controlling for pre-survey covariates in the reduced sample (M2). While short-term weekly measures (e.g. quizzes, homework) do not appear to be significantly impacted by the treatment in this study, significant positive effects on the final exam suggests long-term impacts of continuous treatment administration.

**Other Mechanisms.** In earlier results, I find a generally positive effect of the treatment on some academic performance measures, but no impact on study times. I consider two other possible mechanisms through which the treatment can impact academic performance. The first is changes in students' self-perception. The second is differences in students' progress-monitoring.

Given that students' perceived time management abilities are not necessarily correlated with their actual time management abilities (Kruger & Dunning, 1999), the treatment may guide students to better re-evaluate their study skills. To evaluate changes in student self-perceptions, I refer to a post-course survey that included questions similar to the pre-course survey. Not all students who completed the pre-course survey also completed the post-course survey. Thus, the ITT estimates in Table 9 reflect only students with valid post-course survey responses.

Treatment students' post-course survey responses marginally differed from control students' post-course survey responses on some measures. In particular, students in the treatment group were less likely to agree that they kept a record of their assignments. This runs counterintuitive to the scheduling device, which aimed to help students with their self-regulation.

Though not significant, they were also less likely to agree with other measures of self-regulation. One possible explanation is that the treatment prompted students to become more cognizant of their self-regulation abilities. Correlational estimates suggest that this was not the case. The correlation between students' post-course rating of keeping a record of assignments and their final exam scores, for example, was only 0.034 for the treatment group compared to 0.070 for the control group.

This study additionally investigates if the treatment induced students to be more conscious of their learning progress. That is, while treatment students may not have changed the times that they studied, they may have differed from control students in their focus while studying. I look at evidence in the clickstream data, focusing particularly on how often students check their grades and the discussion forums. For grade-checking, effect sizes each week ranged from 0.058 to 0.270. For discussion forums-checking, effect sizes each week ranged from -0.108 to 0.251. For overall grade-checking and discussion forums-checking throughout the whole school term, the treatment had effect sizes of 0.167 and 0.114 respectively. I find generally positive estimates for how often students check their grades and the discussion forums. The estimates are significant only for checking grades in week 4 for checking the discussion forums in week 3. It is not clear why these weeks might be outstanding. However, these estimates should be interpreted with caution as they suggest that treatment students checked their grades or the discussion forums by a fraction of a click more than control students each week. Across the whole term, the treatment induced students to check their grades three to four more times.

## Discussion

The two studies presented in this chapter found very different, but also promising, results for a scheduling intervention designed to help students' time management and academic performance in online courses. Contrary to the findings of Baker and colleagues (2018), neither of the two studies found a positive effect on academic performance in the first week's assignments. Instead, Study 3.1 found a positive effect on when students studied, but no effect on their academic performance. Study 3.2 found a positive effect on students' final exam score and course grade, but not on their study times. Both studies were similar in having students schedule on a weekly basis. However, Study 3.1 had an added incentive for students to work as scheduled whereas Study 3.2 only required that students plan a schedule in order to receive extra credit.

Study 3.1 found that the scheduling intervention requiring students to work as scheduled did impact when students watched course lecture videos even though students met only about half of their scheduled sessions. Control students tended to watch course videos towards the end of the week, closer to the weekly deadline. Treatment students displayed a similar characteristic, but to a lesser extent. They scheduled to watch videos towards the beginning of the week and ended up distributing the five weekly videos towards both the beginning and end of the week.

Study 3.2 found that prompting students to schedule without incentivizing students to follow their schedules yielded a very small proportion of students who studied as scheduled. Treatment students' study habits aligned closely with that of control students despite intentions to study earlier in the day and throughout the week. Even with similar study habits, however, treatment students consistently performed the same as or better than control students on course assignments. In particular, treatment students outperformed control students on the final exam.

Taken together, the findings across these two studies suggest that time management and academic performance are not directly related. Study 3.1 did not observe a direct improvement in academic performance following earlier starts and greater spacing. Similarly, Study 3.2 did not observe changes in study activity leading to improved academic performance. Thus, while prior studies have found a negative correlation between procrastination and academic performance (Elvers, Polzella, & Graetz, 2003; Michinov, Brunot, Le Bohec, Juhel, & Delaval, 2011), future research needs to further investigate potential mechanisms that connect study timing and academic performance.

One such mechanism worth exploring is self-monitoring. I did see evidence that the scheduling device served as a means for students to reflect on their studying. Treatment students in Study 3.1 tended to schedule videos at night (as opposed to the afternoon) in the later weeks of the course. Scheduling at night more closely matched their actual study hours throughout the day. Though small and not significant, treatment students in Study 3.2 also checked their grades a few more times throughout the quarter compared to control students.

It is important to note that findings in this chapter may have also been influenced by differences across the courses in the two studies. The courses differed in terms of course duration and structure. Study 3.2 saw a positive effect on academic performance on the final exam, which was taken in the eleventh week of the course. The course in Study 3.1 was only five weeks in duration. It is possible that Study 3.1 could have had an unobserved positive effect on long term outcomes. However, negative effects in the last three quizzes and on the final exam of the summer physics course suggest that this was not likely to be the case.

The two public health courses in Study 3.2 also had more deadlines on the course learning platform than the physics course in Study 3.1. Any differences in study times for students in Study 3.2 may not have been captured because students had to log onto the course platform throughout the week to submit assignments. The physics course did have homework due every day. However, any activity related to the homework assignments could not be captured as homework was submitted on a different platform. It is possible, though unlikely, that treatment and control students in Study 3.1 were actively studying on different platforms at the same time.

A third major difference that between the two studies that may have influenced findings was that the student population greatly differed. As Baker and colleagues (2018) found, the scheduling device did not consistently have the same effect across different types of students. Given the differences in student populations, it is reasonable to expect that the two studies did not yield the exact same results. Study 3.1 was conducted in a summer course while Study 3.2 was conducted in two courses taught in the regular school year. Students taking summer courses differ from students enrolled during the regular school year. Since summer is not a required term, students who do enroll in the summer have very particular reasons for doing so (e.g. to graduate on time, to get ahead, schedule conflict in the upcoming school year).

Other differences across the two studies are observed in the demographic background characteristics. All courses were majority female. However, three-quarters of the student population in the public health courses were female compared to only one-half of the student population in the physics course. Less than 10% of students in the physics course were underrepresented minorities compared to over 30% of students in the public health courses. Most students in the physics course were not first-generation college students while most students in the public health courses were first-generation college students.

The two studies presented in this chapter have one major shortcoming. Student activity was only captured through their clickstream on the course learning management platform. Thus, study activity could not be captured for students who downloaded materials and studied offline or students who watched the course lecture videos with a friend. Similarly, students may have clicked on course webpages without much regard to the content. In the case of Study 3.1, it is entirely possible that a student clicked on a course video but did not watch the content. While clickstream data in an online course serves as an excellent proxy for student study habits, especially because students need to log online to access virtually all of the course materials, it does not fully capture how and when students study.

The findings of these replication studies, nevertheless, suggest promising potential for a scheduling intervention to improve student learning and academic performance in online courses. However, there are many ways to implement a scheduling intervention. The two studies presented here reflect the findings of two variations of a scheduling intervention. Given the promising results, further investigation is needed to uncover what intervention components (e.g. treatment duration, monitored versus unmonitored) are most effective, as well as what types of courses and students benefit most from the scheduling intervention.

# CHAPTER 4: SUBSEQUENT EFFECTS FOLLOWING A RANDOMIZED CONTROL TRIAL OF A SCHEDULING INTERVENTION

In Baker, Evans, Li, and Cung's (2018) recent study of a scheduling intervention, researchers found an initial positive effect on students' weekly quiz score. However, that positive effect declined and became negative following the withdrawal of the scheduling intervention. Chapter 3 of this dissertation replicated and modified the experiment done in Baker and colleagues' study by extending the scheduling intervention for the full duration of a 5-week and an 11-week course. Like the experiment conducted in Baker and colleagues' study, Study 3.2 did not incentivize students to follow-up on their scheduled study times. Students were awarded extra credit simply for declaring a schedule at the start of the week.

Seeing a decreasing trend in Baker and colleagues' (2018) study raises concerns for student performance following the withdrawal of a weekly scheduling intervention. In this chapter, I examine the long term effects following regular administration of a scheduling intervention. In particular, I investigate the same two public health courses found in Study 3.2. I focus my investigation on the following research question: does regularly prompting students to schedule have an effect on their subsequent academic performance, subsequent selection of courses, or choice of major?

## Rationale and Theory

### Post-Incentive Effects

Chapter 3 discusses prior literature on time management and academic performance, and how a scheduling intervention can help time management. The scheduling intervention used in Chapter 3 can be viewed as an incentive program, where students earn are rewarded (in this case, with extra credit) for completing an action.

The economics literature has looked at the withdrawal of incentive programs extensively in various contexts. Most behavioral economics studies have found lingering results shortly after the end of the incentive period, but not long after (Acland & Levy, 2015; Barte & Wendel-Vos, 2017; Charness & Gneezy, 2009; Just & Price, 2013; Royer, Stehr, & Syndnor, 2015). Just and Price (2013), for example, found that a short incentive program targeting school food choice had effects two weeks after the incentive period ended, but not four weeks after. In general, researchers have indicated that the persistence of lingering effects depends largely on the strength of the incentive (Acland & Levy, 2015; Charness & Gneezy, 2009), context of the incentive (Belot, James, & Nolen, 2013; Royer, Stehr, & Syndnor, 2015), and duration of the incentive program (Loewenstein, Price, & Volpp, 2016).

Researchers in psychology, however, have expressed a major concern for potential negative effects following incentive programs. Opponents of incentive programs argue that incentives may crowd out intrinsic motivation and that outcomes may be worse after the program than before (Frey & Jegen, 2001). One meta-analysis of 128 studies did find that rewards of various types (e.g. tangible and verbal) undermined motivation (Deci, Koestner, & Ryan, 1999). However, more recent, and smaller-scale, meta-analyses have found positive long-term effects following incentives (Giles, Roballino, McColl, Sniehotta, & Adams, 2014; Maki, Burns, Ha, & Rothman, 2016).

Festré and Garrouste (2014) note the controversial issue between economics and psychology and suggest fundamental differences in theoretical tradition across the two fields. In empirical studies, the authors argued that psychologists tended to look at schools or voluntary situations while economists tended to look at non-voluntary work organizations.

**Study Skills and Long-Term College Performance**

Lack of academic preparation has shown to be a major barrier to college success and retention (e.g. Adelman, 2006). There are many studies of interventions aiming to increase student academic performance, college retention, and choice of college major. Most studies are conducted in the context of remedial courses and summer bridge programs (e.g. Bir & Myrick, 2015; Slade, Eatmon, Staley, & Dixon, 2015; Wathington, Pretlow, Barnett, 2016), which predominantly focus on content-related materials in preparation for upcoming courses. Summer bridge programs, however, have drawn on the importance of study skills (for meta-analysis of study skills, see Robbins et al., 2004) and often include some form of nonacademic skills training (e.g. Wathington, Pretlow, Barnett, 2016; Zhang & Smith, 2011). During the regular school year, student success courses have been found to have a positive effect on college persistence and grades (Hoops, Yu, Burridge, & Wolters, 2015; Zeidenberg, Jenkins, & Calcagno, 2007).

Given the success of academic term-long programs targeting students' nonacademic skills on college success, an intervention administered on a weekly basis like those found in Chapter 3 should have impacts on students' long term college success. The current study is unique from prior interventions in that the students did not enroll in a course or program designed specifically to improve long-term academic performance. This is an important distinction because schools are not always able to offer study skills programs or coerce students to enroll in them. However, the intervention used in the current study is potentially just as effective as summer bridge programs or student success courses because students similarly received regular soft skills training (time management) administered continuously over several weeks. The findings of this study provide evidence for regularly-administered light-touch interventions that could affect students' long-term college success.

## Context

This study examines the same population of students who enrolled in the online public health courses found in Study 3.2. It follows students into their subsequent term at the university and assesses differences in treatment and control students' course-taking, course grades, and college major.

In total, 324 students were enrolled in one of the two public health courses by the end of the add/drop period. Half of the students were randomly assigned into the treatment group (111 from PH1 and 51 from PH2) and half of the students were randomly assigned into the control group (117 from PH1 and 45 from PH2). Of the 324 students who enrolled in either PH1 or PH2, 309 students continued to take courses in the next term. Eight students from the control group did not take courses in the subsequent term and seven students from the treatment did not take courses in the subsequent term. The analysis sample in this study consists of 309 students who took courses in the subsequent term. Among the 309 students, 220 had taken the pre-course survey in public health. Analyses with pre-course survey covariates are also included for the reduced sample of 220 students.

### Data

**Demographic Background and Self-Regulation Skills.** Less than 5% of the original 324 public health enrollees were excluded from this sample. Given the small proportion, the student demographic and reported self-regulation skills did not drastically change from Chapter 3. About the same number of students were excluded from both treatment and control groups. T-tests suggest that both groups were still balanced. Appendix Table 4.1 gives summary statistics of the treatment and control groups in this analytic sample, as well as students who did not take courses in the following term.

One outstanding difference between students who continued and students who did not was the age. Students who did not continue in the subsequent term were, on average, over one year older than students who did continue. This finding is in line with how 8 out of 15 of the students who did not take courses in the subsequent term were in their senior year.

**Subsequent Course Outcomes.** This study examined students' subsequent course outcomes as provided through institutional data. This analysis focuses on subsequent course grades, but also takes into account whether students enroll in courses that are historically easier or harder. Table 4.1 presents summary statistics of students' GPA on a 4.0 scale in the subsequent quarter, as well as historical grade averages of courses that students enrolled in.

On average, students earned between a "B-" and a "B" in their courses. They performed about 0.175 grade points worse than peers in the same course. Courses that students chose to enroll in have historically awarded "B" grades on average.

**College Major.** In addition to subsequent course performance, this study is interested in whether the treatment had any impact on students' academic trajectory as indicated by their choice of major. Since PH1 and PH2 are lower division entry-level courses, students are potentially formulating their first impression of the department or other similar departments. Their impression of the course could influence their impression of the major and other similar majors. There were 45 unique majors among students in the sample. The most common two majors were biological sciences with 107 students and public health with 33 students at the start of the study. Dummy variables were generated to indicate whether students were declared as a biological science major, public health major, or other STEM major. Student choice of major is assessed at the end of the school year to see if the treatment had an impact.

Figure 3.4.1 presents the proportion of students enrolled in specific majors at the beginning of the school year (when the treatment was implemented) and by the end of the school year. At the start of the school year, 34.6% of students were biological science majors, 10.7% of students were public health majors, and 22.7% were some other STEM major. The number of biological sciences majors decreased slightly between both treatment and control groups. However, the number of public health majors sharply increased. By the end of the school year, there were almost as many public health majors as biological sciences majors. The increase appears to be larger for treatment students. Overall, 32% of students changed majors over the course of the year.

## Method of Analysis

This study uses a similar linear regression model as that found in Chapter 3. The goal of the models is to estimate the effect of the treatment assignment on students' subsequent course grades and their choice of college major. I take into account inflated grades if the treatment induced students to enroll in easier or harder courses by computing the difference between the course average score and a given student's awarded score. I will call this score the differential grade. A regression model of students' differential grade ($D_i$) with 18 student-level covariates, $\theta_i$, and one additional course-level fixed effect ($\theta_{19}$, indicative of whether the student took PH1 or PH2) looks like the following:

$$D_i = a + \beta Treatment_i + \gamma\theta_i + \varepsilon_i \quad (1)$$

where

$$D_i = \frac{\sum_n u_n \left(A_{0,n,i} - A_{0,n,\bar{x}}\right)}{u_{total}} \quad (2)$$

In the model, *u* is the number of units that each course (*n*) is worth. *A* is the grade that students receive in a given course.

In addition, I also look to historical course scores from the prior year and two years earlier to see if students may have based their course selection on grades that were awarded in prior years. Since students take multiple courses in a given term, I computed a course load difficulty score for each student using the average grade awarded for each course in prior years. For the difficulty score based on historical data from one year prior, the calculation is as follows:

$$S_{-1,i} = \frac{\sum_n u_n \, A_{-1,n,\bar{x}}}{u_{total}} \qquad (3)$$

Note that in equation (2), average scores from the current year are used, and in equation (3), average scores from the past year are used. The calculation for scores two years prior is the same as that in equation (3), except that it would use average scores from two years prior.

## Results

### Subsequent Academic Performance

To assess whether the treatment had an effect on students' subsequent course performance, I estimated the ITT effects on subsequent GPA. Results suggest that the treatment had a significant positive effect on students' subsequent grades (Table 4.2). The effect becomes non-significant when including the grade that students expected to earn in PH1 or PH2 (Model 3), though the point estimate remains positive. This suggests that students' grade expectations, and self-perceptions, did not fully explain grades in the subsequent term. As effect size of 0.142 as estimated in Model 3 suggests that treatment increased students' GPA by one-tenth of a grade point, which would also equate to half of a grade point in one course (assuming that a student takes three to four courses each term).

In consideration of the possibility that the treatment induced students to take courses that awarded higher grades, I examine the ITT on each students' average difference in score from their classmates. Models 4-6 in Table 4.2 reflect the ITT estimate and suggest that treatment students' grades relative to classmates were better than control students' grade relative to classmates. In Table 4.1, both treatment and control groups had negative values in their grades compared to classmates, suggesting that both groups tended to earn lower grades. The positive estimates in Models 4-6 in Table 4.2 suggest that despite lower grades compared to classmates overall, the gap between treatment students and their classmates was not as drastic as that for control students. The estimate remains significant, marginally, even after a reduction in sample size and controlling for pre-course survey responses.

**Subsequent Course Taking**

As another check to see if the treatment induced students to take easier or harder courses, I turn to grades that have been awarded historically in the same courses. Resources such as ratemyprofessor.com or classmates who have taken the course in the past might have informed students' enrollment decisions. Table 4.3 gives the ITT estimate of the average grade awarded in the prior year and 2 years before. The models show a positive effect for age and higher grade levels, suggesting that older and more senior students tend to take easier courses, or courses that give higher grades on average. The models also show a negative effect for the treatment, suggesting that the treatment induced students to take harder courses.

There is a small negative effect, however, on the number of courses that students took. The estimate is not significant, though it does reflect an adjustment in student schedules to counterbalance the harder course load. It is also possible that the time management aspect of the treatment helped students better judge their time in the subsequent term. Given that the estimate is very small, it is difficult to draw any conclusions on the treatment's impact on the number of courses that students take.

**Heterogeneous Effects on Grades**

Results presented in Tables 4.2 and 4.3 suggest that the treatment had a positive effect on student grades in the subsequent term. Not only did treatment students earn higher grades than control students in the subsequent term, they also enrolled in courses that awarded lower grades on average. Other covariates in the models suggest that males, underrepresented minorities, and freshmen tend to perform worse in the subsequent term. I examine these subgroups, as well as students who expected higher or lower grades in public health, for any heterogeneous effects.

Table 4.4 provides estimates of the treatment effect on subgroups of students. There is a large effect of the treatment on males and freshmen students in their subsequent course grades. The effect sizes are 0.538 and 0.634, respectively. This equates to the treatment helping males and freshmen earn more than half a standard deviation in GPAs, or half a letter grade in all courses. The effect size grows to 0.677 and 0.856 for differential grades from classmates. For females and more senior students, the effect size is much smaller and no longer significant. The point estimates remain positive, which suggest promising results but on a smaller scale.

Positive effects are also present for both groups of students who initial reported high and low levels of planning work in advance and scheduling assignments (Table 4.4). Heterogeneous point estimates suggest an even higher treatment effect on students who report low levels of planning and low levels of scheduling. For students with low levels of planning, the effect is not significant though the magnitude of the effect (0.459 for differential grades) is greater than that for students with high levels of planning (0.253), which is moderately significant for differential grades. Since only 39 students in the survey sample were neutral or disagreed with the planning in advanced survey item, the model is unable to detect significance in the 0.459 effect size. For students who reported low levels of scheduling, the treatment had a moderately significant effect on subsequent grades (0.657 for GPA and 0.630 for differential grades). The effect size increases when controlling for other survey items. These estimates greatly contrast the estimates for students with high levels of scheduling (0.095 for GPA and 0.238 for differential grades). This suggests that the scheduling intervention encouraged students who do not regularly schedule coursework to continue doing so in their other courses.

Given especially positive results for males, freshmen, and students with low self-regulation, one might hypothesize that the treatment is effective for students who traditionally struggle in online courses (Xu & Jaggars, 2013). However, this is not the case for underrepresented minorities, possibly due to differences in the URM population. While the treatment has a strong positive effect overall (0.253 for GPA and 0.337 for differential grade), the effect is less than half that size for underrepresented minorities (0.047 and 0.121, respectively). Non-underrepresented minorities benefit from the treatment more in the subsequent term. The treatment has an effect size of 0.358 for their subsequent course GPA and 0.447 for their subsequent differential grade. Although not a terrible result, this finding is slightly disappointing as underrepresented minorities are among one of the top groups to struggle in college.

Another marker of students who struggle in college are those who expect lower grades. Since an overwhelming majority of students in PH1 and PH2 indicated that they expected to earn an “A” in the course, I examined if the treatment is more or less effective for those with lower self-expectations. Results in Table 4 .4 suggest that the treatment is almost equally effective for the two groups. It appears to be slightly higher for students who expected below an “A” though combining the models in Table 4.4 and including an interaction term to represent differential effects of the treatment by grade expectation suggests that there is no significant difference between students who expected an “A” and students who did not expect an “A” in PH1/PH2 ($p$ = 0.676).

**College Major Choice**

Figure 4.1 shows that both treatment and control groups saw an increase in the number of public health majors by the end of the school year. This makes sense given that both PH1 and PH2 are entry-level courses to the major, though students can take the course as an elective or to fulfill general education requirements. However, Figure 4.1 also shows a sharper increase in public health majors for the treatment group compared to the control group.

Table 4.5 finds students are equally likely to be public health majors (or not) at the start of the school year. By the end of the school year, weeks after administration of the treatment, treatment students are more likely to be in a public health major even after student-level characteristics are taken into account. In other words, there is a positive effect of the treatment on whether or not students will eventually become public health majors by the end of the school year. Some of the effect can be explained by students’ pre-course survey responses. Namely, students who expected to spend more hours on the course are significantly more likely to be a public health major by the end of the school year. Yet a positive point estimate suggests that the scheduling device is a promising tool for helping students maintain a positive view of the subject matter. Students are more likely to change into the public health major if they received the treatment.

Virtually no effect was found for whether students declared themselves as biological science major or other STEM major was found. Effects sizes ranged between -0.052 and 0.027. For better or worse, an effect was instead found for whether or not students changed majors. Most of that effect was attributed to students who changed into the public health major as both neither models predicting changes to any other STEM major nor models predicting changes out of a STEM major were significant.

## Discussion

The study found a positive effect from a weekly-administered scheduling intervention on students' grades in the subsequent term. Not only did students in the treatment group earn higher GPAs, but they also picked harder courses as measured by the historical average grade awarded. Both groups of students earned lower grades on average compared to classmates in the subsequent term, but the treatment students had a smaller gap in grades from classmates compared to how control students performed relative to classmates. In addition, students were more likely to change in to the public health major. The treatment was more effective for males, freshmen, non-underrepresented minorities, and students who report lower levels of planning and scheduling.

These findings are promising in that the study has identified a low-cost intervention that greatly helps students' long term academic success. The intervention is particularly helpful for two groups of students that many universities are concerned about: students with low levels of planning and scheduling and freshmen who are new to navigating college.

Although the original intent of the scheduling device was to help improve students' academic performance in the online course by developing their time management skills, the current study found evidence the intervention also impacted students' subsequent academic performance, subsequent selection of courses, and choice of major. This falls in line with empirical studies in economics and psychology, which have demonstrated treatment effects of incentive programs even after the end of the incentive period. In this study, students are incentivized with extra credit to plan out online coursework on a weekly basis.

Withdrawal effects were suspected in a prior study of the scheduling device (Baker, Evans, Li, & Cung, 2018). However, the prior study found declining effects with eventually negative estimates, while the current study found positive effects following withdrawal which follows the overall positive trend found in Study 3.2. Indeed, the outcome measures of this study compared to other studies of the scheduling device are different. This study is concerned with students' overall subsequent academic performance while other studies have focused solely on current course outcomes in the specific course that administered the scheduling device. However, it is still important to consider other reasons why academic performance is positively impacted by the scheduling device after withdrawal in this study but not in Baker and colleagues' study.

There are several possible explanations for the observed difference in the treatment effect following withdrawal. One explanation is that the population of students across the two studies are different, which may also explain findings in Chapter 3. Baker, Evans, Li, and Cung (2018) looked at students in an online physics course taught during the summer. Not only do students who take physics courses greatly differ from students who take public health courses, but students who choose to enroll in summer courses also greatly differ from the general population of all students during the regular school year. From the investigation of heterogeneous effects, the current study found that males, freshmen, and non-underrepresented minorities see a greater treatment effect. Seeing heterogeneous effects, it is likely that there is a key difference in student populations that generated the different outcomes.

Furthermore, the treatment duration may have been a factor. As mentioned, researchers have suggested that the lingering effects of an intervention depends largely on the duration of the intervention program (Loewenstein, Price, & Volpp, 2016). The scheduling device in this study was administered for the full 11-week duration of the online course, compared to the initial two weeks of a five-week course. With a short treatment duration in Baker and colleagues' (2018) study, students did not have enough time to establish a sense of regular scheduling. As a result, quickly withdrawing the treatment left a negative effect on students' academic performance by the end of the course while administering the treatment over an extended period of time maintained a neutral to positive effect on students' academic performance by the end of the course and into the subsequent course.

In addition to treatment duration, this study also might have seen a positive effect because the scheduling device was consistently administered throughout the term, as opposed to withdrawn in the middle of the term. In Baker and colleagues' study, the treatment was withdrawn only two weeks into the course. Scheduling may not have been viewed as an important enough task to complete for the full duration of a given course. In existing studies of the scheduling device, it is difficult to disentangle whether encouraging students to schedule for a longer period of time or for the full duration of the course is more important. However, this can be examined in future studies by incentivizing scheduling, without incentivizing following schedules, for the full duration a course that is taught over a fewer number of weeks (e.g. five-week summer course).

Student academic performance is important, and there are many studies that investigate ways to help students in their current courses. However, it is equally important, if not more so, to investigate any long-term effects. This study has shown that a scheduling device implemented in a fully online course does in fact have lingering effects even after the treatment period, and the lingering effects are beneficial to students' academic performance and college trajectory. However, this study only investigated academic performance in the subsequent term and major choice by the end of the school year. It is possible, and should be investigated, for the treatment to have longer term negative impacts on grades and graduation rates.

This study was conducted in a unique setting. Students were enrolled in one of two fully online public health courses, and their grades across all other courses in the subsequent term were measured. It is not clear whether the results will equally apply to other course types (e.g. face-to-face), subjects (e.g. math), student populations, or academic terms with longer or shorter course durations. It is also not clear whether the trend will hold for longer term outcomes. These are all factors that need to be investigated in future studies.

# CHAPTER 5: KEY FINDINGS AND FUTURE RESEARCH

## Key Findings

Online courses have become an extremely integral part of postsecondary education. As such, it is important to understand what types of students are affected by the change from in-person to online courses and how issues that arise can be addressed. This dissertation aimed to contribute to the existing knowledge base about online education by exploring personality types that are impacted by online instruction and investigating a scheduling intervention designed to help students' self-regulation and time management. To achieve this goal, I examined various aspects of students enrolled in three different undergraduate online courses.

In Chapter 2, I investigated personality traits according to the Big Five personality model (Costa & McCrae, 1992; McCrae & Costa, 1987). I calculated composite personality scores for students in two undergraduate online public health courses. I obtained historical course data and, using individual fixed effects, I compared how students performed in in-person courses to their own performance in online courses. By including interaction terms between personality type and an indicator for online courses, I was able to obtain an estimate for the differential impact of online education for different personality types.

Conscientiousness was found to be an important predictor of academic performance in both online and in-person courses. Further investigations of academic performance in both modalities revealed that highly conscientious students tended to perform even better in online courses. Some common terms describing students who are conscientious include hardworking, well-organized, self-disciplined, self-reliant, and punctual (McCrae & Costa, 1987). Thus, the findings are consistent with prior studies that have found that highly motivated students with self-regulated learning and time-management skills tended to perform better in online courses

(Bambara, Harbour, & Davies, 2009; Deal III, 2002; Liu, Gomez, Khan & Yen, 2007). In general, positive findings for academic performance in online courses are desirable for universities with online instruction as part of their long-term plan. However, universities also need to take caution and think through available support systems as the differential gains for less conscientious students suggests a widening gap between high performing students and low performing students in online spaces.

In addition, openness to experience was also an important predictor of academic performance. However, the differential impact of online education on students who are highly open to experience disappeared for more senior students. Prior findings for academic performance in in-person courses found a weak positive association between degree of openness to experience and academic performance in in-person courses (O'Connor & Paunonen, 2007; Vedel, 2014). This finding held true for the in-person courses examined in Chapter 2. For online courses, however, a negative association was found between degree of openness and academic performance. The negative association was stronger for more senior students.

In Chapter 3, I implemented a scheduling intervention in three undergraduate online courses. One of the courses was a five-week physics course while the other two were eleven-week public health courses. The scheduling intervention for the physics course and the two public health courses varied slightly in that students in the physics course were asked to schedule their weekly course lecture videos, and the intervention incentivized the act of watching as scheduled. Students in the two public health courses were asked to schedule when they would work on the online coursework and the intervention incentivized the act of scheduling without regard to whether or not students actually worked as scheduled. In all courses, students were prompted to schedule on a weekly basis.

The findings for the two studies presented in Chapter 3 were mixed. The first of the two studies (physics) found an effect on students' study behavior. Students in the physics course crammed less and procrastinated less on average, though no effect was found on students' academic performance. The second of the two studies (public health) found a positive effect on students' academic performance. However, the intervention did not appear to have an effect on students' study behavior. Taken together, the scheduling intervention has a neutral to positive effect on students' study behavior and students' current course academic performance.

In Chapter 4, I followed public health students into their subsequent term at the university to examine the effect of the scheduling intervention on subsequent academic performance and declared major by the end of the school year. Using historical average course scores and the average scores of peers in the subsequent term, I controlled for differences in course selection following the intervention.

On average, both treatment and control students tended to perform worse than peers in their subsequent courses. However, treatment students did not have as big of a gap in academic performance from peers as control students. Furthermore, treatment students tended to pick courses that were about the same difficulty or harder. The scheduling intervention was more effective for males, freshmen, non-underrepresented minorities, and students who report lower levels of planning and scheduling. In addition, while both treatment and control groups saw an increase in the proportion of public health majors, the treatment group saw a steeper increase in the proportion of public health majors.

Overall, the studies presented in this dissertation contribute two important findings to what is currently known about online education. First, individual characteristics, as reflected in personality descriptors, also predict differential gains or losses in online courses compared to in-person courses. By understanding and the differences that contribute to or prevent learning in online courses, academic leaders can then customize interventions to address key pitfalls. The second finding, then, pertains to an intervention targeting the lack of scheduled course time in asynchronous online courses, taking into account prior findings that students with low self-regulation and time management tend to do worse in online courses. Encouraging students to schedule coursework on a weekly basis not only has the potential to improve current academic outcomes, but it also has long-term impacts on student success.

## Future Research

The studies in this dissertation do have a number of limitations. The studies presented were conducted in lower division online undergraduate courses at a large research university in the western United States. Given heterogeneous findings within the same courses presented in this dissertation, the findings are even more likely to vary for different departments and dissimilar universities. By no means does this dissertation intend to speak for students in all postsecondary online courses broadly. Replications in different postsecondary contexts (e.g. other departments, upper division online courses, community colleges, teaching universities in a different part of the United States) are needed to confirm the external validity of current findings.

What can be taken from the studies presented is the general understanding of how individual differences can impact online learning and how interventions can impact individuals in both current and subsequent coursework. In addition, the methodology used in this dissertation can be adopted as way to investigate other online student populations and the impact of novel interventions.

Future investigations should also look into a wider time frame. Chapter 2 of this dissertation only included students' prior course enrollment, which potentially biases results in two ways. First, students take harder courses over time, so later courses may have lower scores. Second, students have more experience with college courses over time, so later courses may instead have higher scores. These two biases do not necessarily cancel each other out and disentangling their effects can be challenging. In Chapter 2, I addressed the potential biases by conducting the analysis again on only courses in the same Fall 2017 term.

In addition, future studies that are able to implement an intervention and include a wider time frame should investigate course performance beyond just the subsequent term. Chapter 4 of this dissertation included only students' grades in the subsequent term and students' choice of major by the end of the school year. However, empirical studies in economics have shown that the effects of incentive programs, in this case awarding extra credit for scheduling, fade over time. Thus, it is not clear whether treatment students will continue to choose harder courses and perform better relative to peers compared to control students beyond the subsequent term. One reason why a positive effect might only be found in the subsequent term but not even later terms (two or more terms later) is that student enrolled in their subsequent term courses during the treatment period. Thus, investigating later terms up until students graduate will better highlight the academic impact of the scheduling intervention.

## Conclusion

Online education is shifting the postsecondary education scene. However, students are not equally adjusting to the different instructional medium. As a result, many universities need to find ways to help support student learning in online courses. To achieve this task effectively, academic leaders need to gain a better understanding of the types of students that are negatively affected by online instruction. Thereafter, academic leaders can use what is known about online learners to guide well-thought interventions that aim to address key teaching and learning issues.

As this dissertation has shown, certain personality types are associated with poor performance in online courses, yet interventions do not need to be costly to have a long term positive impact. The findings of this dissertation can influence online instructors and academic leaders to better identify at-risk students and to implement a cost-effect intervention to optimize student academic performance.

# REFERENCES

Acland, D., & Levy, M. R. (2015). Naiveté, projection bias, and habit formation in gym attendance. *Management Science, 61*(1), 146-160.

Adelman, C. (2006). The toolbox revisited: Paths to degree completion from high school through college. *US Department of Education*.

Al-Dujaily, A., Kim, J., & Ryu, H. (2013). Am I extravert or introvert? Considering the personality effect toward e-learning system. *Journal of Educational Technology & Society*, *16*(3).

Alkış, N., & Temizel, T. T. (2015). The impact of individual differences on influence strategies. *Personality and Individual Differences, 87*, 147-152.

Allen, I. E., & Seaman, J. (2006). Growing by degrees: Online education in the United States, 2005. *Sloan Consortium (NJ1)*.

Allen, I. E., Seaman, J., Poulin, R., & Straut, T. T. (2016). Online report card: tracking online education in the United States. *Babson Survey Research Group and Quahog Research Group*.

Allport, G. W., & Odbert, H. S. (1936). Trait-names: A psycho-lexical study. *Psychological Monographs*, *47*(1), i-171.

Ashraf, N., Karlan, D., & Yin, W. (2006). Tying Odysseus to the mast: Evidence from a commitment savings product in the Philippines. *The Quarterly Journal of Economics, 121*(2), 635-672.

Babcock, P., & Marks, M. (2011). The falling time cost of college: Evidence from half a century of time use data. *Review of Economics and Statistics*, *93*(2), 468-478.

Baker, R., Evans, B., & Dee, T. (2016). A randomized experiment testing the efficacy of a scheduling nudge in a Massive Open Online Course (MOOC). *AERA Open*, *2*(4), 2332858416674007.

Baker, R., Evans, B., Li, Q., & Cung, B. (2018). Does inducing students to schedule lecture watching in online classes improve their academic performance? An experimental analysis of a time management intervention. *Research in Higher Education*, 1-32.

Bambara, C. S., Harbour, C. P., Davies, T. G., & Athey, S. (2009). Delicate engagement: The lived experience of community college students enrolled in high-risk online courses. *Community College Review*, *36*(3), 219-238.

Barrick, M. R., & Mount, M. K. (1991). The big five personality dimensions and job performance: a meta-analysis. *Personnel psychology*, *44*(1), 1-26.

Barrick, M. R., Mount, M. K., & Judge, T. A. (2001). Personality and performance at the beginning of the new millennium: What do we know and where do we go next?. *International Journal of Selection and assessment*, *9*(1-2), 9-30.

Barte, J. C., & Wendel-Vos, G. W. (2017). A systematic review of financial incentives for physical activity: the effects on physical activity and related outcomes. *Behavioral medicine*, *43*(2), 79-90.

Beattie, G., Laliberté, J. W. P., Michaud-Leclerc, C., & Oreopoulos, P. (2017). *What sets college thrivers and divers apart? A contrast in study habits, attitudes, and mental health*(No. w23588). National Bureau of Economic Research.

Belot, M., James, J., & Nolen, P. (2013). *Changing Eating Habits-A Field Experiment in Primary Schools* (No. 219). Edinburgh School of Economics, University of Edinburgh.

Bernard, R. M., Abrami, P. C., Lou, Y., Borokhovski, E., Wade, A., Wozney, L., ... & Huang, B. (2004). How does distance education compare with classroom instruction? A meta-analysis of the empirical literature. *Review of educational research*, *74*(3), 379-439.

Bir, B., & Myrick, M. (2015). Summer bridge's effects on college student success. *Journal of Developmental Education*, *39*(1), 22-30.

Boghikian-Whitby, S., & Mortagy, Y. (2016). Student preferences and performance in online and face-to-face classes using Myers-Briggs Indicator: A longitudinal quasi-experimental study. *Issues in Informing Science and Information Technology*, *13*, 89-109.

Bork, R. H., & Rucks-Ahidiana, Z. (2013). Role Ambiguity in Online Courses: An Analysis of Student and Instructor Expectations. Community College Research Center, Teacher's College, Columbia University.

Bowen, W. G., Chingos, M. M., Lack, K. A., & Nygren, T. I. (2014). Interactive learning online at public universities: Evidence from a six-campus randomized trial. *Journal of Policy Analysis and Management*, *33*(1), 94-111.

Brint, S., & Cantwell, A. M. (2010). Undergraduate time use and academic outcomes: Results from the University of California Undergraduate Experience Survey 2006. *Teachers College Record*, *112*(9), 2441-2470.

Busato, V. V., Prins, F. J., Elshout, J. J., & Hamaker, C. (1998). The relation between learning styles, the Big Five personality traits and achievement motivation in higher education. *Personality and Individual Differences*, *26*(1), 129-140.

Busato, V. V., Prins, F. J., Elshout, J. J., & Hamaker, C. (2000). Intellectual ability, learning style, personality, achievement motivation and academic success of psychology students in higher education. *Personality and Individual Differences*, *29*(6), 1057-1068.

Carr, S. (2000). As distance education comes of age, the challenge is keeping the students. *Chronicle of Higher Education*, 46, 39–42.

Carrell, S. E., Maghakian, T., & West, J. E. (2011). A's from Zzzz's? The causal effect of school start time on the academic achievement of adolescents. *American Economic Journal: Economic Policy*, *3*(3), 62-81.

Cattell, R. B. (1946). Description and measurement of personality.

Chamorro-Premuzic, T., & Furnham, A. (2003a). Personality predicts academic performance: Evidence from two longitudinal university samples. *Journal of Research in Personality*, *37*(4), 319-338.

Chamorro-Premuzic, T., & Furnham, A. (2003b). Personality traits and academic examination performance. *European Journal of Personality*, *17*(3), 237-250.

Charness, G., & Gneezy, U. (2009). Incentives to exercise. *Econometrica*, *77*(3), 909-931.

Cheung, F. M., & Leung, K. (1998). Indigenous personality measures: Chinese examples. *Journal of Cross-Cultural Psychology*, *29*(1), 233-248.

Cobb-Clark, D. A., & Schurer, S. (2012). The stability of big-five personality traits. *Economics Letters*, 115(1), 11-15.

Cochran, J. D., Campbell, S. M., Baker, H. M., & Leeds, E. M. (2014). The role of student characteristics in predicting retention in online courses. *Research in Higher Education*, 55(1), 27-48.

Cohen, A., & Baruth, O. (2017). Personality, learning, and satisfaction in fully online academic courses. *Computers in Human Behavior*, *72*, 1-12.

Conard, M. A. (2006). Aptitude is not enough: How personality and behavior predict academic performance. *Journal of Research in Personality*, *40*(3), 339-346.

Conrad, D. L. (2002). Engagement, excitement, anxiety, and fear: Learners' experiences of starting an online course. *The American Journal of Distance Education, 16*(4), 205-226.

Costa Jr, P. T., & McCrae, R. R. (1992). Four ways five factors are basic. *Personality and Individual Differences, 13*(6), 653-665.

Daughenbaugh, R., Ensminger, D., Frederick, L., & Surry, D. (2002). Does Personality Type Effect Online versus In-Class Course Satisfaction?.

Deal, W. F., III. (2002). Distance learning: Teaching technology online. *Technology Teacher, 61* (8), 21-27.

Deci, E. L., Koestner, R., & Ryan, R. M. (1999). The undermining effect is a reality after all—Extrinsic rewards, task interest, and self-determination: Reply to Eisenberger, Pierce, and Cameron (1999) and Lepper, Henderlong, and Gingras (1999). *Psychological Bulletin*, *125*(6), 692-700.

Deming, D. J., Goldin, C., Katz, L. F., & Yuchtman, N. (2015). Can online learning bend the higher education cost curve?. *American Economic Review*, *105*(5), 496-501.

Dollinger, S. J., & Orf, L. A. (1991). Personality and performance in “personality”: Conscientiousness and openness. *Journal of Research in Personality*, *25*(3), 276-284.

Dutton, J., Dutton, M., & Perry, J. (2001). Do online students perform as well as lecture students?. *Journal of Engineering Education*, 90(1), 131-136.

Elvers, G. C., Polzella, D. J., & Graetz, K. (2003). Procrastination in online courses: Performance and attitudinal differences. *Teaching of Psychology*, *30*(2), 159-162.

Evans, B. J., Baker, R. B., & Dee, T. S. (2016). Persistence patterns in massive open online courses (MOOCs). *The Journal of Higher Education*, 87(2), 206-242.

Eysenck, H. J. (1978). The development of personality and its relation to learning. *Critical Studies in Education*, *20*(1), 134-181.

Feldman, K. A., Smart, J. C., & Ethington, C. A. (1999). Major field and person-environment fit: Using Holland's theory to study change and stability of college students. *The Journal of Higher Education, 70*(6), 642-669.

Festré, A., & Garrouste, P. (2015). Theory and evidence in psychology and economics about motivation crowding out: A possible convergence?. *Journal of Economic Surveys*, *29*(2), 339-356.

Figlio, D., Rush, M., & Yin, L. (2013). Is it live or is it internet? Experimental estimates of the effects of online instruction on student learning. *Journal of Labor Economics*, *31*(4), 763-784.

Frantzen, D. (2014). Is technology a one-size-fits-all solution to improving student performance? A comparison of online, hybrid and face-to-face courses. *Journal of Public Affairs Education*, *20*(4), 565-578.

Frederick, S., Loewenstein, G., & O'donoghue, T. (2002). Time discounting and time preference: A critical review. *Journal of Economic Literature*, *40*(2), 351-401.

Frey, B. S., & Jegen, R. (2001). Motivation crowding theory. *Journal of Economic Surveys*, *15*(5), 589-611.

Furnham, A. (1996). The big five versus the big four: the relationship between the Myers-Briggs Type Indicator (MBTI) and NEO-PI five factor model of personality. *Personality and Individual Differences*, *21*(2), 303-307.

Furnham, A., & Chamorro-Premuzic, T. (2004). Personality and intelligence as predictors of statistics examination grades. *Personality and Individual Differences*, *37*(5), 943-955.

Furnham, A., Chamorro-Premuzic, T., & McDougall, F. (2003). Personality, cognitive ability, and beliefs about intelligence as predictors of academic performance. *Learning and Individual Differences*, *14*(1), 47-64.

Giles, E. L., Robalino, S., McColl, E., Sniehotta, F. F., & Adams, J. (2014). The effectiveness of financial incentives for health behaviour change: Systematic review and meta-analysis. *PloS one*, *9*(3), e90347.

Giné, X., Karlan, D., & Zinman, J. (2010). Put your money where your butt is: A commitment contract for smoking cessation. *American Economic Journal: Applied Economics*, *2*(4), 213-35.

Goldberg, L. R. (1971). *A Historical Survey of Personality Scales and Inventories*. Oregon Research Institute.

Goldstein, D., Hahn, C. S., Hasher, L., Wiprzycka, U. J., & Zelazo, P. D. (2007). Time of day, intellectual performance, and behavioral problems in morning versus evening type adolescents: Is there a synchrony effect?. *Personality and Individual Differences*, *42*(3), 431-440.

Gosling, S. D., Rentfrow, P. J., & Swann Jr, W. B. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality*, *37*(6), 504-528.

Gurven, M., Von Rueden, C., Massenkoff, M., Kaplan, H., & Lero Vie, M. (2013). How universal is the Big Five? Testing the five-factor model of personality variation among forager–farmers in the Bolivian Amazon. *Journal of Personality and Social Psychology*, *104*(2), 354.

Hair, P., & Hampson, S. E. (2006). The role of impulsivity in predicting maladaptive behaviour among female students. *Personality and Individual Differences*, *40*(5), 943-952.

Hakimi, S., Hejazi, E., & Lavasani, M. G. (2011). The relationships between personality traits and students' academic achievement. *Procedia-Social and Behavioral Sciences*, *29*, 836-845.

Harrington, R., & Loffredo, D. A. (2010). MBTI personality type and other factors that relate to preference for online versus face-to-face instruction. *The Internet and Higher Education*, *13*(1-2), 89-95.

Hart, C. (2012). Factors associated with student persistence in an online program of study: A review of the literature. *Journal of Interactive Online Learning*, *11*(1), 19-42.

Hartwig, M. K., & Dunlosky, J. (2012). Study strategies of college students: Are self-testing and scheduling related to achievement?. *Psychonomic Bulletin & Review*, *19*(1), 126-134.

Hoops, L. D., Yu, S. L., Burridge, A. B., & Wolters, C. A. (2015). Impact of a student success course on undergraduate academic outcomes. *Journal of College Reading and Learning*, *45*(2), 123-146.

Huber, D., Kaufmann, H., & Steinmann, M. (2014). *Bridging the Innovation Gap*. Springer Gabler.

Jaggars, S., & Xu, D. (2010). Online learning in the Virginia Community College System. Community College Research Center, Teacher's College, Columbia University.

Jessee, S. A., O'Neill, P. N., & Dosch, R. O. (2006). Matching student personality types and learning preferences to teaching methodologies. *Journal of Dental Education*, *70*(6), 644-651.

John, O. P., & Srivastava, S. (1999). The Big Five trait taxonomy: History, measurement, and theoretical perspectives. *Handbook of Personality: Theory and Research*, *2*(1999), 102-138.

John, O. P., Angleitner, A., & Ostendorf, F. (1988). The lexical approach to personality: A historical review of trait taxonomic research. *European Journal of Personality*, *2*(3), 171-203.

Judge, T. A., & Bono, J. E. (2001). Relationship of core self-evaluations traits—self-esteem, generalized self-efficacy, locus of control, and emotional stability—with job satisfaction and job performance: A meta-analysis. *Journal of Applied Psychology*, *86*(1), 80.

Judge, T. A., Heller, D., & Mount, M. K. (2002). Five-factor model of personality and job satisfaction: A meta-analysis. *Journal of Applied Psychology*, *87*(3), 530.

Judge, T. A., Rodell, J. B., Klinger, R. L., Simon, L. S., & Crawford, E. R. (2013). Hierarchical representations of the five-factor model of personality in predicting job performance: integrating three organizing frameworks with two theoretical perspectives. *Journal of Applied Psychology*, *98*(6), 875-925.

Just, D. R., & Price, J. (2013). Using incentives to encourage healthy eating in children. *Journal of Human Resources*, *48*(4), 855-872.

Kamal, A., & Radhakrishnan, S. (2018). Individual learning preferences based on personality traits in an E-learning scenario. *Education and Information Technologies*, 1-29.

Kaur, S., Kremer, M., & Mullainathan, S. (2015). Self-control at work. *Journal of Political Economy*, *123*(6), 1227-1277.

Keller, H., & Karau, S. J. (2013). The importance of personality in students' perceptions of the online learning experience. *Computers in Human Behavior*, *29*(6), 2494-2500.

Kizilcec, R. F., Pérez-Sanagustín, M., & Maldonado, J. J. (2017). Self-regulated learning strategies predict learner behavior and goal attainment in Massive Open Online Courses. *Computers & Education*, *104*, 18-33.

Koch, A. K., & Nafziger, J. (2017). *Motivational Goal Bracketing: An Experiment* (No. 10955). Institute for the Study of Labor (IZA) Discussion Paper.

Kristof-Brown, A. L., Zimmerman, R. D., & Johnson, E. C. (2005). KRISTOF-BROWN, A. (2005). Consequences of individuals' fit at work: A meta-analysis of person-job, person-organization, person-group, and person-supervisor fit. *Personnel Psychology*, *58*(2), 281-342.

Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, *77*(6), 1121.

Lack, K. A. (2013). Current status of research on online learning in postsecondary education. *Ithaka S+ R, zuletzt geprüft am*, *3*, 2013.

Lee, Y., & Choi, J. (2011). A review of online course dropout research: Implications for practice and future research. *Educational Technology Research and Development*, *59*(5), 593-618.

Leeds, E., Campbell, S., Baker, H., Ali, R., Brawley, D., & Crisp, J. (2013). The impact of student retention strategies: An empirical study. *International Journal of Management in Education*, *7*(1-2), 22-43.

León, F. R., Morales, O., & Vértiz, H. (2017). Personality traits that differentiate attendants of higher-education online courses. *Je-LKS: Journal of e-Learning and Knowledge Society*.

Levy, Y. (2007). Comparing dropouts and persistence in e-learning courses. *Computers and Education*, *48(2)*, 185–204.

Liu, S., Gomez, J., Khan, B., & Yen, C. J. (2007). Toward a learner-oriented community college online course dropout framework. *International Journal on ELearning, 6*(4), 519-542.

Loewenstein, G., Price, J., & Volpp, K. (2016). Habit formation in children: Evidence from incentives for healthy eating. *Journal of Health Economics*, *45*, 47-54.

Lounsbury, J. W., Loveland, J. M., Sundstrom, E. D., Gibson, L. W., Drost, A. W., & Hamrick, F. L. (2003). An investigation of personality traits in relation to career satisfaction. *Journal of Career Assessment*, *11*(3), 287-307.

Lyons, J. F. (2004). Teaching US history online: Problems and prospects. *The History Teacher*, *37*(4), 447-456.

Macan, T. H., Shahani, C., Dipboye, R. L., & Phillips, A. P. (1990). College students' time management: Correlations with academic performance and stress. *Journal of Educational Psychology*, *82*(4), 760-768.

Maki, A., Burns, R. J., Ha, L., & Rothman, A. J. (2016). Paying people to protect the environment: A meta-analysis of financial incentive interventions to promote proenvironmental behaviors. *Journal of Environmental Psychology*, *47*, 242-255.

McCaulley, M. H. (1990). The Myers-Briggs Type Indicator: A measure for individuals and groups. *Measurement and Evaluation in Counseling and Development*, *22*(4), 181-195.

McCrae, R. R., & Costa, P. T. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology, 52*(1), 81.

McCrae, R. R., & Terracciano, A. (2005). Universal features of personality traits from the observer's perspective: data from 50 cultures. *Journal of Personality and Social Psychology*, *88*(3), 547.

Means, B., Toyama, Y., Murphy, R., & Bakia, M. & Jones, K.(2010). Evaluation of evidence-based practices in online Learning: A meta-analysis and review of online learning studies. *Washington, D. C: US Department of Education*.

Messick, S. (1984). The nature of cognitive styles: Problems and promise in educational practice. *Educational Psychologist*, *19*(2), 59-74.

Michinov, N., Brunot, S., Le Bohec, O., Juhel, J., & Delaval, M. (2011). Procrastination, participation, and performance in online learning environments. *Computers & Education*, *56*(1), 243-252.

Misra, R., & McKean, M. (2000). College students' academic stress and its relation to their anxiety, time management, and leisure satisfaction. *American Journal of Health Studies*, *16*(1), 41-51.

Myers, I. B., McCaulley, M. H., Quenk, N. L., & Hammer, A. L. (1998). *MBTI manual: A guide to the development and use of the Myers-Briggs Type Indicator* (Vol. 3). Palo Alto, CA: Consulting Psychologists Press.

Myers, I. B., McCaulley, M. H., Quenk, N. L., & Hammer, A. L. (1998). *MBTI manual: A guide to the development and use of the Myers-Briggs Type Indicator* (Vol. 3). Palo Alto, CA: Consulting Psychologists Press.

National Center for Education Statistics. (NCES). (2015). *Digest of Education Statistics*, Table 311.15.

Nawrot, I., & Doucet, A. (2014, April). Building engagement for MOOC students: introducing support for time management on online learning platforms. In *Proceedings of the 23rd International Conference on world wide web* (pp. 1077-1082). ACM.

Noftle, E. E., & Robins, R. W. (2007). Personality predictors of academic outcomes: Big Five correlates of GPA and SAT scores. *Journal of Personality and Social Psychology*, *93*(1), 116.

Norman, W. T. (1963). Toward an adequate taxonomy of personality attributes: Replicated factor structure in peer nomination personality ratings. *The Journal of Abnormal and Social Psychology*, *66*(6), 574.

O'Connor, M. C., & Paunonen, S. V. (2007). Big Five personality predictors of post-secondary academic performance. *Personality and Individual Differences*, *43*(5), 971-990.

Patterson, R. W. (2018). Can behavioral tools improve online student outcomes? Experimental evidence from a massive open online course. *Journal of Economic Behavior & Organization*, *153*, 293-321.

Paunonen, S. V., & Ashton, M. C. (2001). Big Five predictors of academic achievement. *Journal of Research in Personality*, *35*(1), 78-90.

Phillips, P., Abraham, C., & Bond, R. (2003). Personality, cognition, and university students' examination performance. *European Journal of Personality*, *17*(6), 435-448.

Pintrich, P. R. (1991). A manual for the use of the Motivated Strategies for Learning Questionnaire (MSLQ).

Poropat, A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological Bulletin*, *135*(2), 322.

Robbins, S. B., Lauver, K., Le, H., Davis, D., Langley, R., & Carlstrom, A. (2004). Do psychosocial and study skill factors predict college outcomes? A meta-analysis. *Psychological Bulletin*, *130*(2), 261-288.

Roper, A. R. (2007). How students develop online learning skills. *Educause Quarterly*, *30*(1), 62-65.

Rosander, P., Bäckström, M., & Stenberg, G. (2011). Personality traits and general intelligence as predictors of academic performance: A structural equation modelling approach. *Learning and Individual Differences*, *21*(5), 590-596.

Rostaminezhad, M. A., Mozayani, N., Norozi, D., & Iziy, M. (2013). Factors related to e-learner dropout: Case study of IUST elearning center. *Procedia-Social and Behavioral Sciences*, *83*, 522-527.

Rothstein, M. G., Paunonen, S. V., Rush, J. C., & King, G. A. (1994). Personality and cognitive ability predictors of performance in graduate business school. *Journal of Educational Psychology*, *86*(4), 516-530.

Royer, H., Stehr, M., & Sydnor, J. (2015). Incentives, commitments, and habit formation in exercise: evidence from a field experiment with workers at a fortune-500 company. *American Economic Journal: Applied Economics*, *7*(3), 51-84.

Saucier, G., & Goldberg, L. R. (2001). Lexical studies of indigenous personality factors: Premises, products, and prospects. *Journal of Personality*, *69*(6), 847-879.

Shih, H. F., Chen, S. H. E., Chen, S. C., & Wey, S. C. (2013). The relationship among tertiary level EFL students' personality, online learning motivation and online learning satisfaction. *Procedia-Social and Behavioral Sciences*, *103*, 1152-1160.

Singh, P., and Pan, W. (2004). Online education: Lessons for administrators and instructors. *College Student Journal*, 38, 302-308.

Slade, J., Eatmon, D., Staley, K., & Dixon, K. G. (2015). Getting into the pipeline: Summer bridge as a pathway to college success. *The Journal of Negro Education*, *84*(2), 125-138.

Song, L., Singleton, E. S., Hill, J. R., & Koh, M. H. (2004). Improving online learning: Student perceptions of useful and challenging characteristics. *The Internet and Higher education*, *7*(1), 59-70.

Soto, C. J., John, O. P., Gosling, S. D., & Potter, J. (2011). Age differences in personality traits from 10 to 65: Big Five domains and facets in a large cross-sectional sample. *Journal of Personality and Social Psychology*, *100*(2), 330.

Stajkovic, A. D., Bandura, A., Locke, E. A., Lee, D., & Sergent, K. (2018). Test of three conceptual models of influence of the big five personality traits and self-efficacy on academic performance: A meta-analytic path-analysis. *Personality and Individual Differences*, *120*, 238-245.

Steel, P. (2007). The nature of procrastination: A meta-analytic and theoretical review of quintessential self-regulatory failure. *Psychological bulletin*, *133*(1), 65-94.

Stinebrickner, R., & Stinebrickner, T. R. (2004). Time-use and college outcomes. *Journal of Econometrics*, *121*(1-2), 243-269.

Stinebrickner, R., & Stinebrickner, T. R. (2008). The causal effect of studying on academic performance. *The BE Journal of Economic Analysis & Policy*, *8*(1).

Taylor, S. S. (2003). The endless class. *Community College Week*, *15*(20), 6-9.

Tello, S. F. (2007). An analysis of student persistence in online education. *International Journal of Information and Communication Technology Education*, *3*(3), 47–62.

Terrell, S. R. (2005). A longitudinal investigation of the effect of information perception and focus on attrition in online learning environments. *The Internet and Higher Education*, 8(3), 213-219.

Tett, R. P., Jackson, D. N., & Rothstein, M. (1991). Personality measures as predictors of job performance: A meta-analytic review. *Personnel Psychology*, *44*(4), 703-742.

Tlili, A., Essalmi, F., Jemni, M., & Chen, N. S. (2016). Role of personality in computer based learning. *Computers in Human Behavior*, *64*, 805-813.

Tupes, E. C., & Christal, R. E. (1992). Recurrent personality factors based on trait ratings. *Journal of Personality*, *60*(2), 225-251.

van Den Hurk, M. (2006). The relation between self-regulated strategies and individual study time, prepared participation and achievement in a problem-based curriculum. *Active Learning in Higher Education*, *7*(2), 155-169.

Vedel, A. (2014). The Big Five and tertiary academic performance: A systematic review and meta-analysis. *Personality and Individual Differences*, *71*, 66-76.

Wathington, H., Pretlow, J., & Barnett, E. (2016). A good start? The impact of Texas' developmental summer bridge program on student success. *The Journal of Higher Education*, *87*(2), 150-177.

Watjatrakul, B. (2016). Online learning adoption: effects of neuroticism, openness to experience, and perceived values. *Interactive Technology and Smart Education*, *13*(3), 229-243.

Williams, V. (n.d.). *Online Readiness Assessment*. Retrieved from https://pennstate.qualtrics.com/jfe/form/SV_7QCNUPsyH9f012B

Wolfe, R. N., & Johnson, S. D. (1995). Personality as a predictor of college performance. *Educational and Psychological Measurement*, *55*(2), 177-185.

Xu, D., & Jaggars, S. S. (2011). Online and Hybrid Course Enrollment and Performance in Washington State Community and Technical Colleges. CCRC Working Paper No. 31. Community College Research Center, Columbia University.

Xu, D., & Jaggars, S. S. (2013). The impact of online learning on students' course outcomes: Evidence from a large community and technical college system. *Economics of Education Review, 37*, 46-57.

Xu, D., & Jaggars, S. S. (2014). Performance gaps between online and face-to-face courses: Differences across types of students and academic subject areas. *The Journal of Higher Education*, *85*(5), 633-659.

Zeidenberg, M., Jenkins, D., & Calcagno, J. C. (2007). Do Student Success Courses Actually Help Community College Students Succeed? CCRC Brief. Number 36. Community College Research Center, Columbia University.

Zhang, P., & Smith, W. L. (2011). From high school to college: The transition experiences of Black and White students. *Journal of Black Studies*, *42*(5), 828-845.

# APPENDIX A: FIGURES

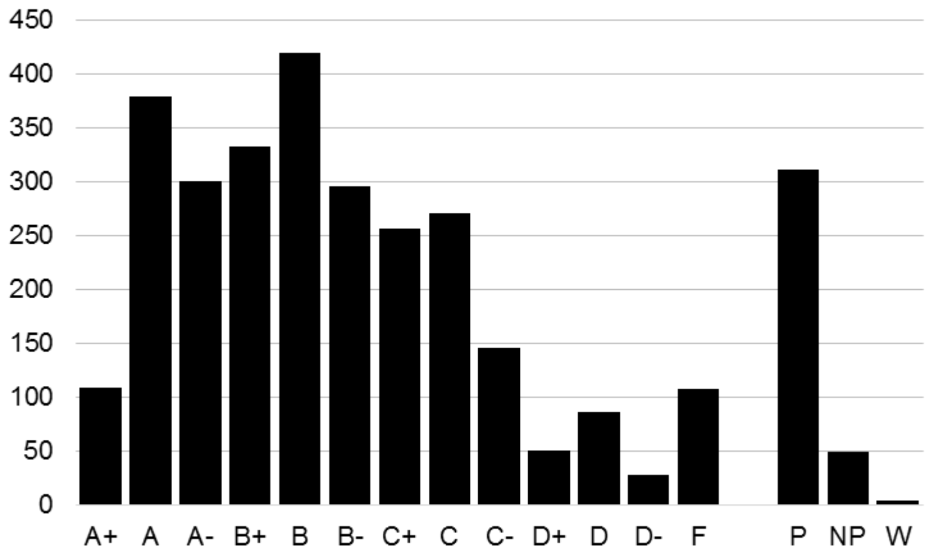



Figure 2.1 Grade Distribution (Count) for All Courses

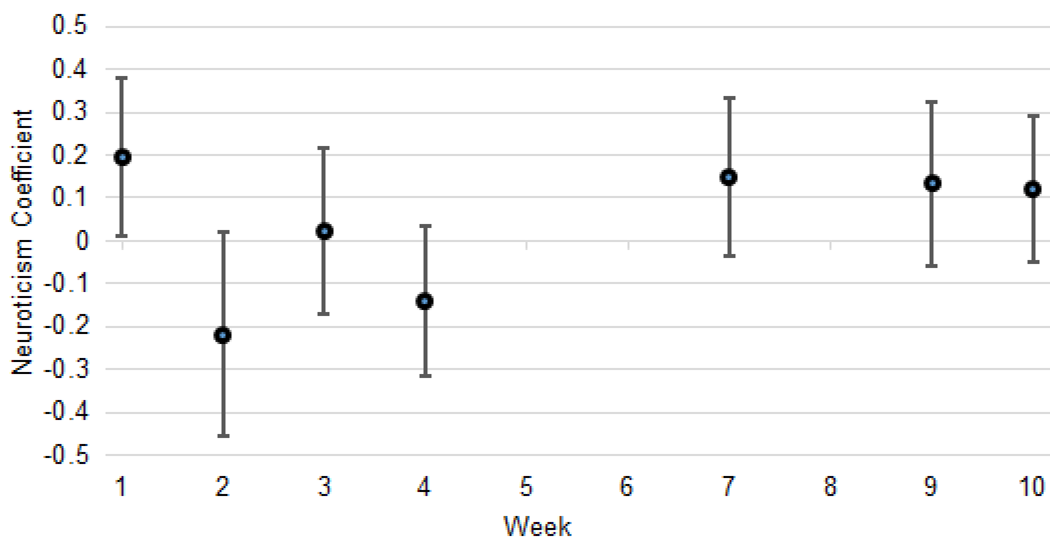



Figure 2.2 Weekly Quiz Neuroticism Coefficient with 95% Confidence Intervals

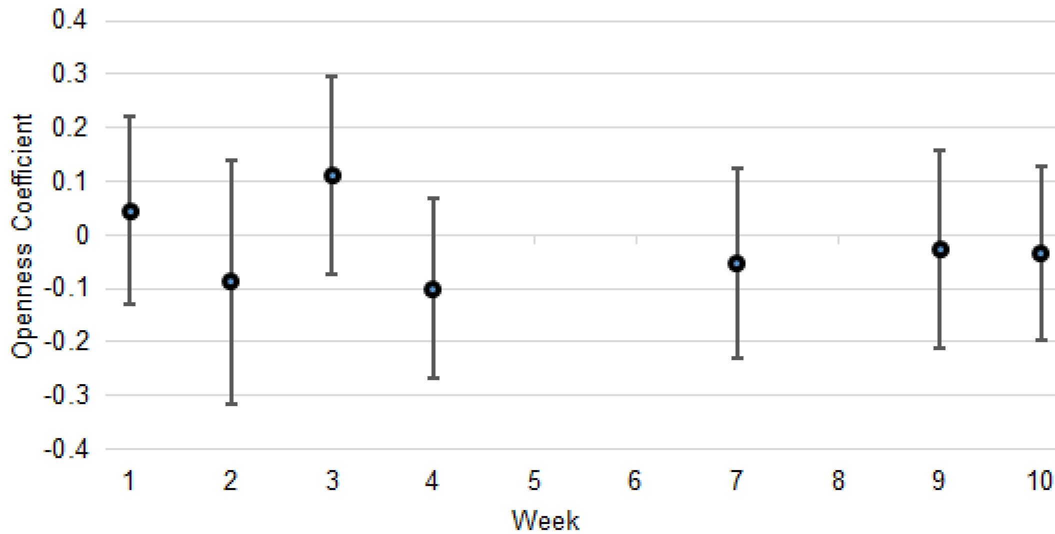



Figure 2.3 Weekly Quiz Openness Coefficient with 95% Confidence Intervals

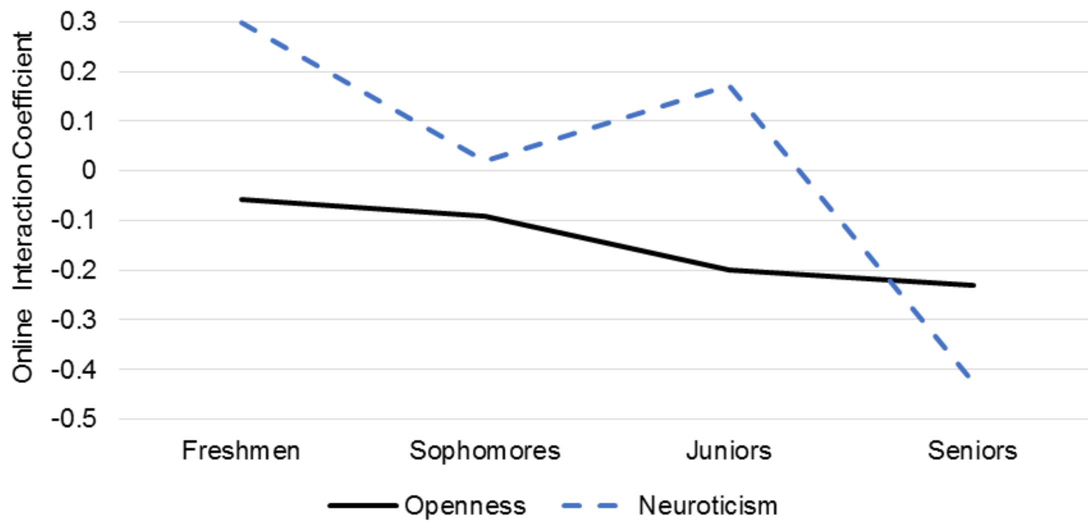



Figure 2.4 Personality-Online Interaction Coefficient by Academic Grade

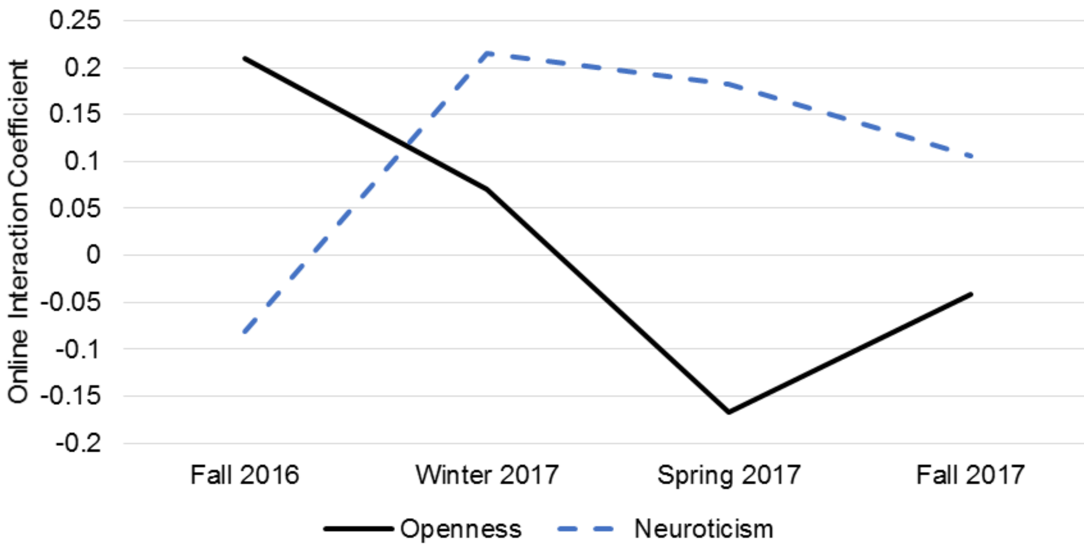


Figure 2.5 Personality-Online Interaction Coefficient by Academic Term

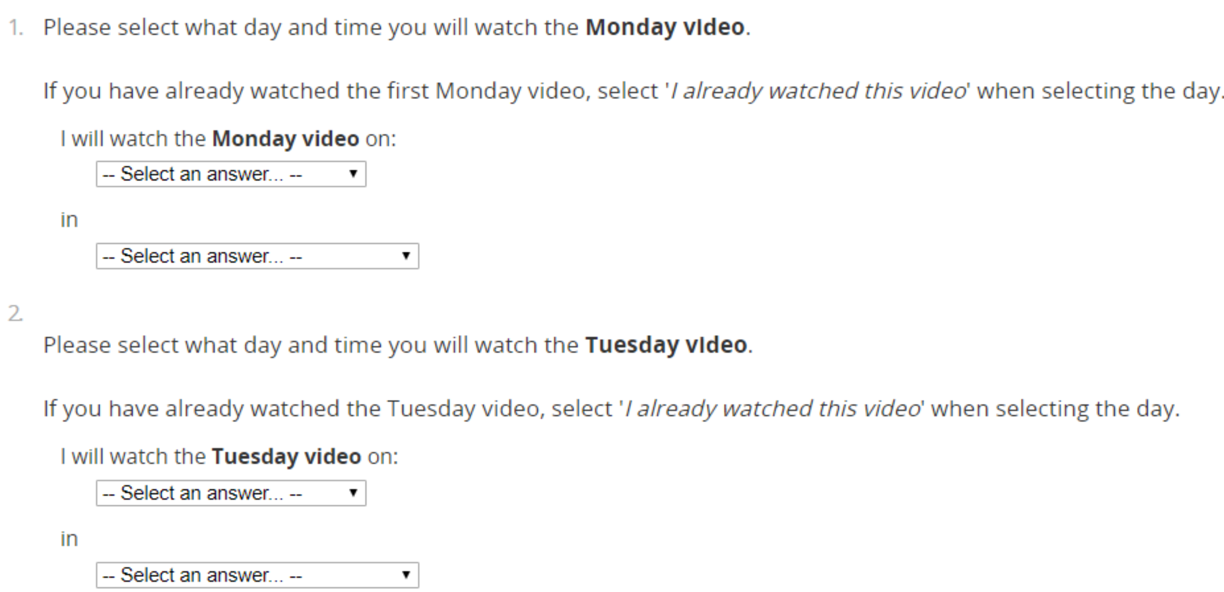


Figure 3.1. Screenshot of Physics Course Scheduling Device

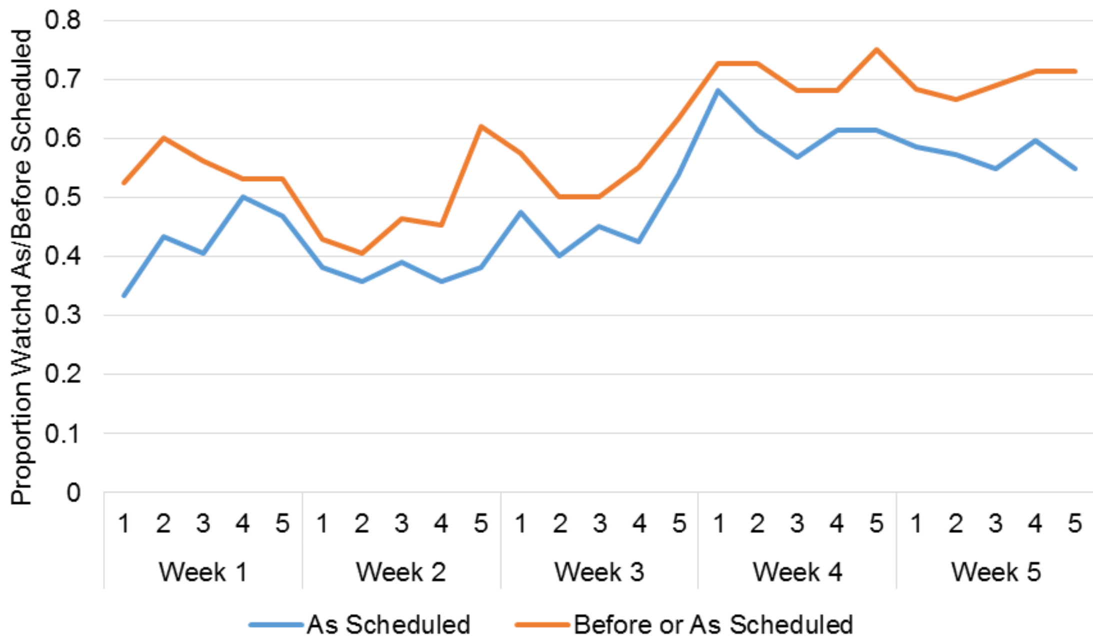



Figure 3.2. Physics Proportion of Students who Watched as Scheduled

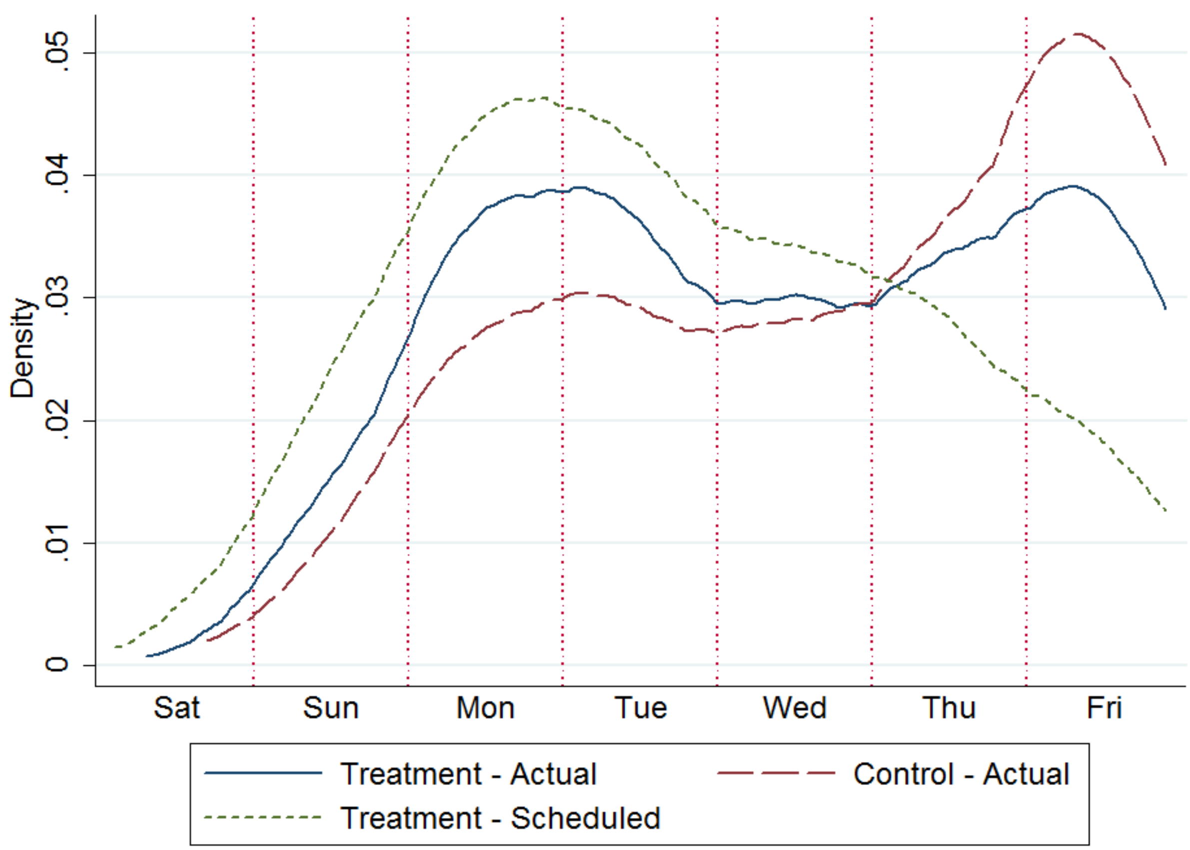



Figure 3.3. Physics Distribution of Actual and Scheduled Video Times Across Week

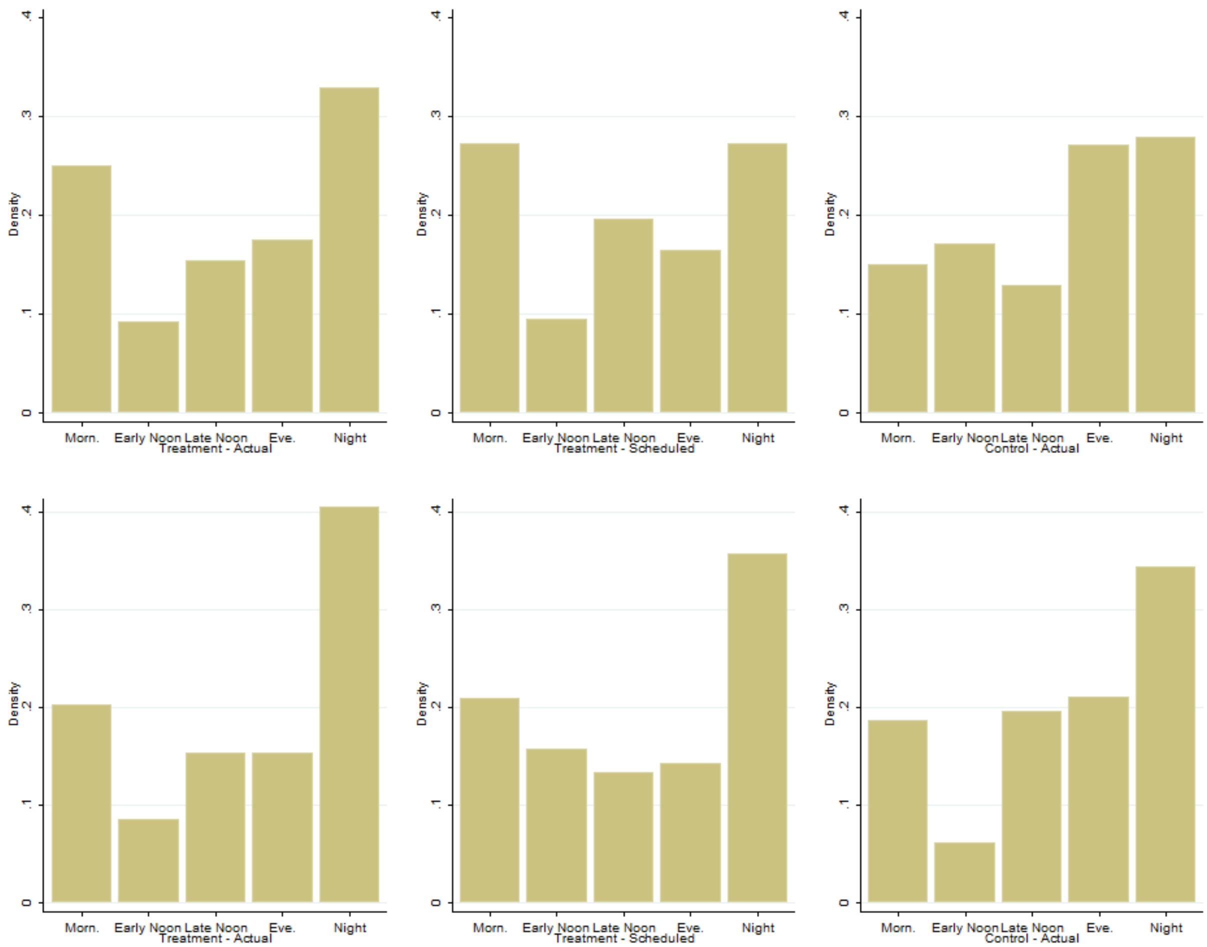
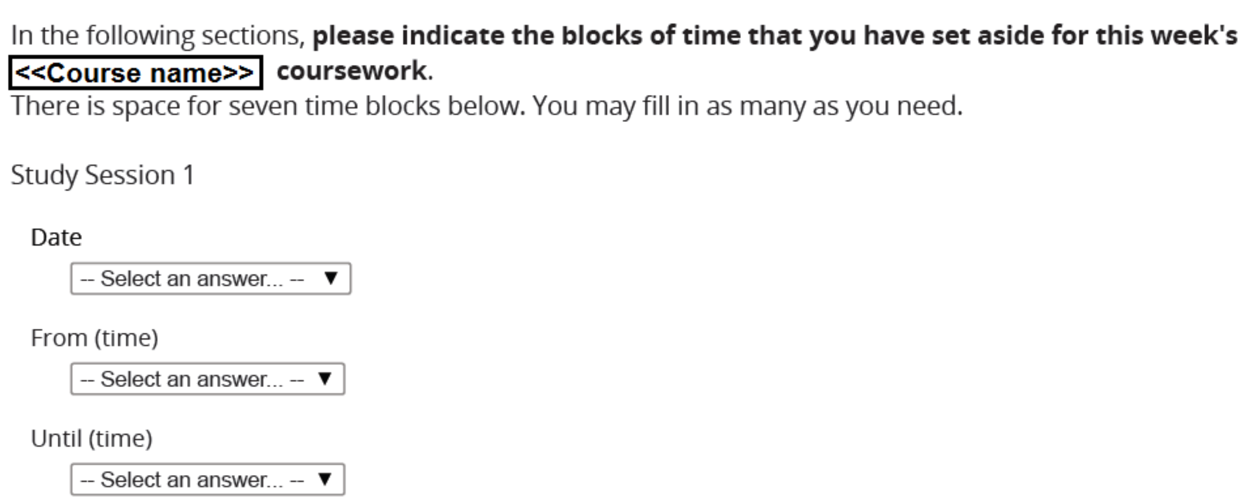



Figure 3.4. Physics Distribution of Treatment Actual (left), Treatment Scheduled (middle), and Control Actual (right) Daily Video Watching Times in Week 1 (top) and Week 5 (bottom)


In the following sections, **please indicate the blocks of time that you have set aside for this week's <<Course name>> coursework.**
There is space for seven time blocks below. You may fill in as many as you need.

Study Session 1

Date

-- Select an answer... --

From (time)

-- Select an answer... --

Until (time)

-- Select an answer... --

Figure 3.5 Screenshot of Public Health Course Scheduling Device

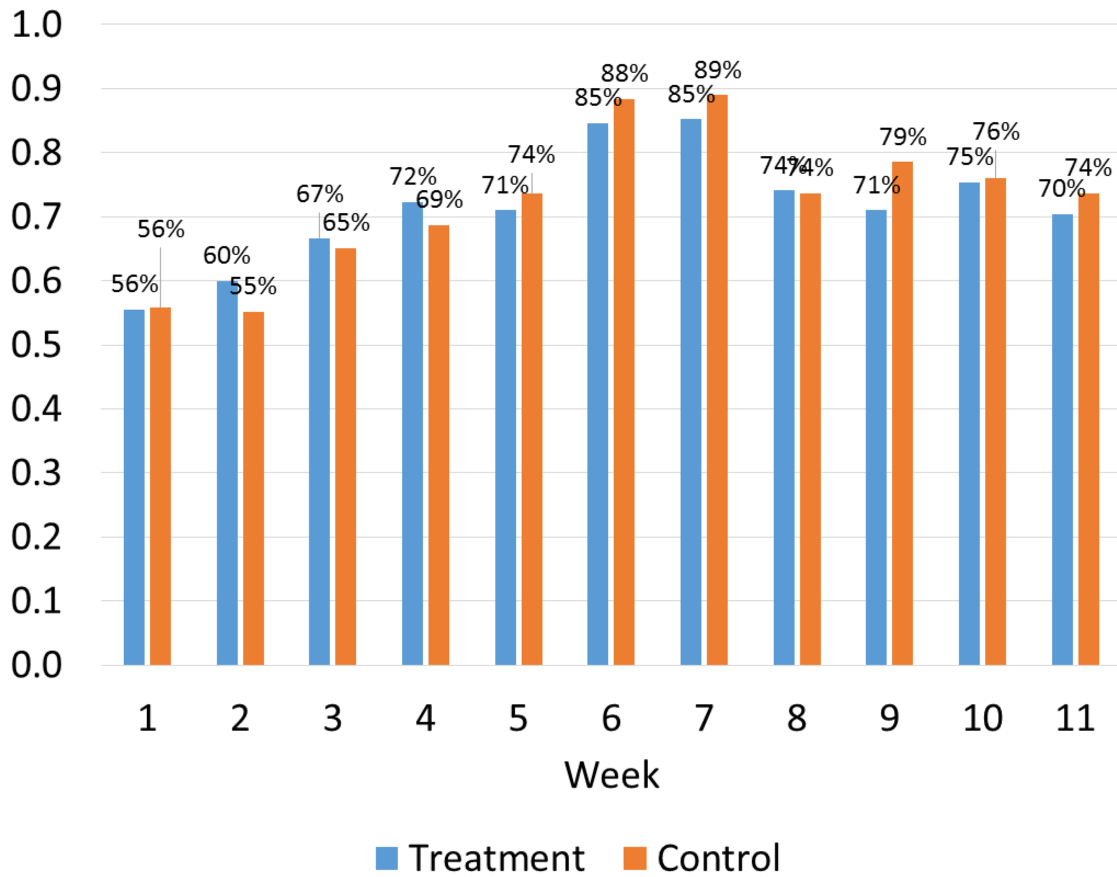



Figure 3.6. Public Health Weekly Survey Completion Rate

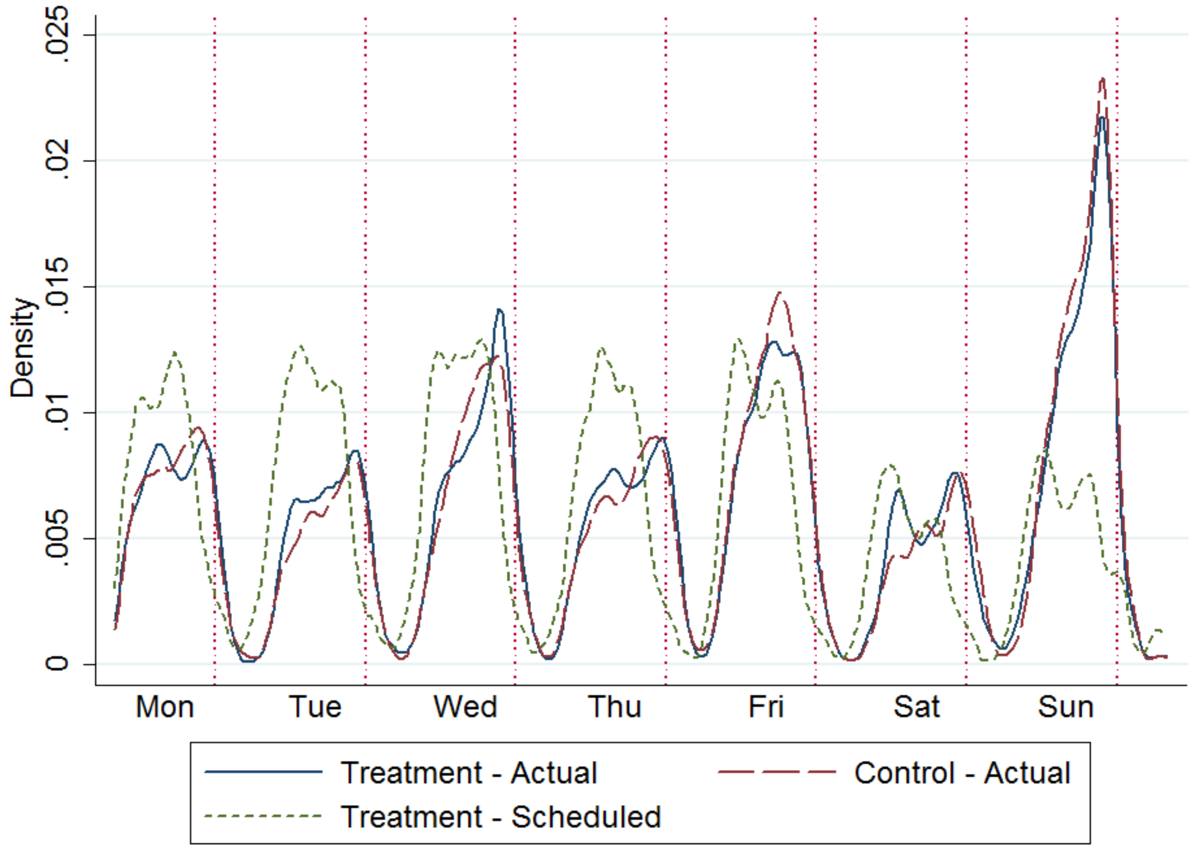



Figure 3.7. Public Health Distribution of Actual and Scheduled Watching Times

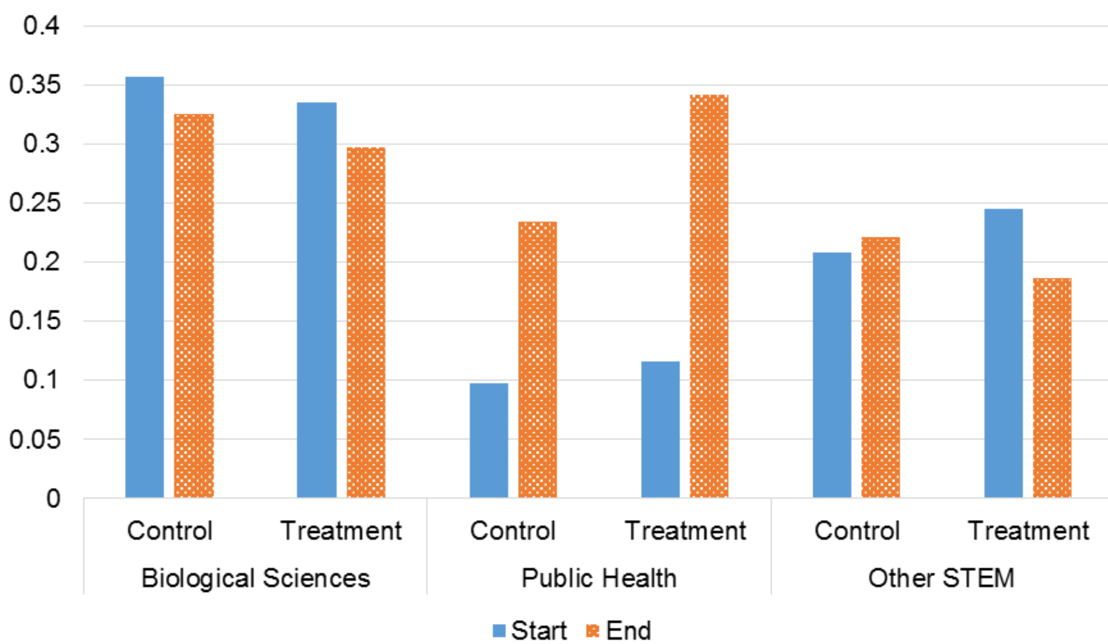



Figure 4.1. College Major at Start and End of Academic Year

# APPENDIX B: TABLES

Table 2.1. Demographic, Personality, and Outcome Summary Statistics

| | Mean | SD | N |
|---|---|---|---|
| Demographic background | | | |
| Age | 20.135 | 1.911 | 157 |
| Male | 0.255 | 0.437 | 157 |
| URM | 0.338 | 0.474 | 157 |
| Low-income | 0.395 | 0.490 | 157 |
| First-generation | 0.522 | 0.501 | 157 |
| No English at home | 0.363 | 0.482 | 157 |
| Academic year | | | |
| Freshman | 0.185 | 0.389 | 157 |
| Sophomore | 0.548 | 0.499 | 157 |
| Junior | 0.172 | 0.379 | 157 |
| Senior and above | 0.096 | 0.295 | 157 |
| Composite personality self-ratings | | | |
| Extraversion (time 1) | 3.127 | 0.840 | 144 |
| Agreeableness (time 1) | 3.942 | 0.637 | 156 |
| Conscientiousness (time 1) | 3.394 | 0.809 | 156 |
| Neuroticism (time 1) | 3.139 | 0.837 | 144 |
| Openness (time 1) | 3.414 | 0.578 | 156 |
| Extraversion (time 2) | 3.219 | 0.760 | 157 |
| Agreeableness (time 2) | 3.887 | 0.618 | 155 |
| Conscientiousness (time 2) | 3.528 | 0.667 | 155 |
| Neuroticism (time 2) | 3.133 | 0.791 | 153 |
| Openness (time 2) | 3.459 | 0.586 | 153 |
| Academic performance | | | |
| PH midterm (60 max) | 49.325 | 11.465 | 157 |
| PH final (60 max) | 43.885 | 12.037 | 157 |
| PH course score (100 max) | 82.141 | 17.077 | 157 |
| Average grade (all courses) | 2.614 | 0.571 | 157 |

All demographic variables except age are dummy variables. Personality self-ratings are on a 5-point scale with 5 being *strongly agree*. Personality scores by generated by grouping 44 items according to the Big Five Inventory (John & Srivastava, 1999). Sample size differs for composite personality self-ratings because some students did not answer the weekly personality question items.

Table 2.2. OLS of Personality on PH1 and PH2 Scores

| | Course Score | Midterm | Final | Presentation | Paper |
|---|---|---|---|---|---|
| Extraversion | 0.173* | 0.179* | 0.101 | 0.226** | 0.131 |
| | (0.080) | (0.086) | (0.078) | (0.080) | (0.085) |
| Agreeableness | -0.151* | -0.114 | -0.123+ | -0.212** | -0.110 |
| | (0.075) | (0.080) | (0.072) | (0.074) | (0.079) |
| Conscientiousness | 0.432*** | 0.412*** | 0.324*** | 0.355*** | 0.287** |
| | (0.091) | (0.097) | (0.088) | (0.090) | (0.096) |
| Neuroticism | 0.159+ | 0.175+ | 0.006 | 0.189* | 0.222* |
| | (0.089) | (0.095) | (0.086) | (0.088) | (0.094) |
| Openness | -0.059 | -0.103 | -0.073 | -0.008 | -0.044 |
| | (0.085) | (0.091) | (0.082) | (0.084) | (0.090) |

*N* = 157. Standard error in parentheses. Models include seven student-level covariates from institutional data and a fixed effect for the public health course.
+*p* < 0.10, **p* < 0.05, ***p* < 0.01, *** *p* < 0.001

Table 2.3. Personality on Grade with Course FE, Excluding PH1 and PH2

| | All | In-Person | Online Courses |
|---|---|---|---|
| Extraversion | 0.035+ | 0.020 | 0.175* |
| | (0.021) | (0.021) | (0.073) |
| Agreeableness | -0.043* | -0.032 | -0.085 |
| | (0.020) | (0.021) | (0.071) |
| Conscientiousness | 0.111*** | 0.098*** | 0.230** |
| | (0.023) | (0.024) | (0.081) |
| Neuroticism | -0.030 | -0.025 | -0.051 |
| | (0.023) | (0.024) | (0.076) |
| Openness | 0.007 | 0.021 | -0.097 |
| | (0.021) | (0.021) | (0.080) |
| *N* | 2985 | 2727 | 258 |

Standard error in parentheses. Models include seven student-level covariates from institutional data and course fixed effects. PH1 and PH2 grades are excluded.
+*p* < 0.10, **p* < 0.05, ***p* < 0.01, *** *p* < 0.001

Table 2.4. Academic Performance with Student-Level FE

| | All | Fall 2017 Only |
|---|---|---|
| Online | 0.310*** | 0.248** |
| | (0.050) | (0.083) |
| Extraversion × Online | 0.062 | -0.093 |
| | (0.055) | (0.091) |
| Agreeableness × Online | -0.043 | 0.030 |
| | (0.050) | (0.085) |
| Conscientiousness × Online | 0.234*** | 0.260* |
| | (0.060) | (0.101) |
| Neuroticism × Online | 0.003 | 0.122 |
| | (0.056) | (0.092) |
| Openness × Online | -0.150* | -0.077 |
| | (0.058) | (0.095) |
| *N* | 3142 | 625 |

Standard error in parentheses. Models include seven student-level covariates from institutional data and student fixed effects.
+$p < 0.10$, *$p < 0.05$, **$p < 0.01$, *** $p < 0.001$

Table A2.1. Personality on Grade with FE, Including PH1 and PH2

| | Course Fixed Effects (w/ PH) | | Department Fixed Effects | |
|---|---|---|---|---|
| | All | Online | All but PH | All |
| Online | | | 0.116+ | 0.115+ |
| | | | (0.067) | (0.067) |
| Extraversion | 0.044* | 0.145* | 0.032 | 0.039* |
| | (0.020) | (0.056) | (0.020) | (0.019) |
| Agreeableness | -0.047* | -0.096+ | -0.030 | -0.034+ |
| | (0.020) | (0.054) | (0.019) | (0.019) |
| Conscientiousness | 0.129*** | 0.292*** | 0.113*** | 0.130*** |
| | (0.022) | (0.063) | (0.022) | (0.022) |
| Neuroticism | -0.015 | 0.029 | -0.029 | -0.015 |
| | (0.022) | (0.061) | (0.022) | (0.021) |
| Openness | 0.004 | -0.046 | 0.011 | 0.008 |
| | (0.020) | (0.061) | (0.020) | (0.020) |
| *N* | 3142 | 415 | 2985 | 3142 |

Standard error in parentheses. Models include seven student-level covariates from institutional data.
+$p < 0.10$, *$p < 0.05$, *** $p < 0.001$

Table 3.1. Demographic and Pre-Course Survey Summary Statistics, Study 3.1

| | Full Sample (n = 96) | | Control (n = 48) | | Treatment (n = 48) | | T-Test |
|---|---|---|---|---|---|---|---|
| | Mean | SD | Mean | SD | Mean | SD | |
| *Demographic characteristics* | | | | | | | |
| Male | 0.417 | 0.496 | 0.458 | 0.504 | 0.375 | 0.489 | * |
| URM | 0.073 | 0.261 | 0.063 | 0.245 | 0.083 | 0.279 | |
| Age | 20.198 | 0.845 | 19.996 | 0.746 | 20.400 | 0.896 | * |
| Low income | 0.302 | 0.462 | 0.333 | 0.476 | 0.271 | 0.449 | |
| First generation | 0.354 | 0.481 | 0.292 | 0.459 | 0.417 | 0.498 | |
| No English at home | 0.354 | 0.481 | 0.271 | 0.449 | 0.438 | 0.501 | |
| Freshman | 0.156 | 0.365 | 0.208 | 0.410 | 0.104 | 0.309 | |
| Sophomore | 0.740 | 0.441 | 0.750 | 0.438 | 0.729 | 0.449 | |
| Junior | 0.094 | 0.293 | 0.042 | 0.202 | 0.146 | 0.357 | + |
| Senior | 0.010 | 0.102 | 0.000 | 0.000 | 0.021 | 0.144 | |
| *Pre-course survey items* | | | | | | | |
| Expected hours per week | 11.208 | 4.563 | 11.625 | 4.949 | 10.792 | 4.151 | |
| Keep record of assignments | 4.010 | 1.031 | 4.000 | 1.072 | 4.021 | 1.000 | |
| Plan work in advance | 3.927 | 0.849 | 4.083 | 0.794 | 3.771 | 0.881 | + |
| Work where there are no distractions | 4.042 | 0.939 | 4.042 | 0.967 | 4.042 | 0.922 | |
| Ignore distractions around me | 3.156 | 1.089 | 3.313 | 1.151 | 3.000 | 1.011 | |
| Willing to use email to ask questions | 3.906 | 0.963 | 3.938 | 1.019 | 3.875 | 0.914 | |
| Lazy or bored so quit (Reversed) | 2.562 | 0.880 | 2.646 | 0.956 | 2.479 | 0.799 | |
| Work hard even if do not like task | 3.875 | 0.824 | 4.063 | 0.836 | 3.688 | 0.776 | * |
| When difficult give up (Reversed) | 2.948 | 0.887 | 3.062 | 0.954 | 2.833 | 0.808 | |
| Even when dull keep working | 3.688 | 0.850 | 3.938 | 0.861 | 3.438 | 0.769 | ** |
| Off-campus 1+ hour away | 0.083 | 0.278 | 0.104 | 0.309 | 0.063 | 0.245 | |
| Off-campus nearby | 0.688 | 0.466 | 0.646 | 0.483 | 0.729 | 0.449 | |
| On-campus | 0.229 | 0.423 | 0.250 | 0.438 | 0.208 | 0.410 | |

All demographic variables except age are dummy variables. All pre-survey items except summer residence items (e.g. off-campus) and expected hours per week are measured on a Likert scale from 1 (strongly disagree) to 5 (strongly agree). Two effort regulation items have been reverse-coded. Residence items are dummy variables.
+$p < 0.10$, *$p < 0.05$, **$p < 0.01$

Table 3.2. Academic Performance Summary Statistics, Study 3.1

| | Full Sample (n = 96) | | Control (n = 48) | | Treatment (n = 48) | |
|---|---|---|---|---|---|---|
| | Mean | SD | Mean | SD | Mean | SD |
| W1 quiz (6 max) | 4.073 | 2.088 | 4.240 | 2.104 | 3.906 | 2.080 |
| W2 quiz (6 max) | 4.573 | 1.422 | 4.563 | 1.227 | 4.583 | 1.606 |
| W3 quiz (6 max) | 4.307 | 1.430 | 4.406 | 1.197 | 4.208 | 1.637 |
| W4 quiz (6 max) | 5.043 | 1.210 | 5.090 | 1.095 | 4.996 | 1.325 |
| W5 quiz (6 max) | 4.543 | 2.194 | 4.424 | 2.356 | 4.663 | 2.037 |
| Final exam (75 max) | 48.630 | 16.428 | 50.167 | 15.374 | 47.094 | 17.444 |
| Course grade (4.0-scale) | 2.761 | 0.837 | 2.818 | 0.783 | 2.704 | 0.892 |

Scores were collected at the end of the course. Course grade has been adjusted to exclude extra credit awarded from the weekly surveys.

Table 3.3. Treatment Effect Estimates on Weekly Study Times, Study 3.1

| | Time Remaining | Spacing |
|---|---|---|
| Week 1 | 0.082 | 0.083 |
| | (0.247) | (0.251) |
| Week 2 | 0.303 | 0.290 |
| | (0.256) | (0.240) |
| Week 3 | 0.377 | 0.451* |
| | (0.248) | (0.225) |
| Week 4 | 0.587* | 0.707** |
| | (0.240) | (0.212) |
| Week 5 | 0.566* | 0.591** |
| | (0.250) | (0.219) |

Each cell reports the ITT estimate from a regression of the treatment on time remaining or spacing of videos watched in a given week, measured in standard deviation units. Time remaining is defined as the average time remaining across the five videos watched in each week. Spacing is defined as the standard deviation of times that students watched videos each week. Regression models include 18 student-level covariates from institutional data and pre-course survey responses. N = 96
+$p < 0.10$, *$p < 0.05$, **$p < 0.01$

Table 3.4. Treatment Effect on Academic Outcomes, Study 3.1

| | Treatment Effect (n = 96) | # Videos as Scheduled (n = 48) |
|---|---|---|
| W1 quiz | -0.0097 | 0.071 |
| | (0.233) | (0.101) |
| W2 quiz | 0.081 | 0.093 |
| | (0.246) | (0.108) |
| W3 quiz | -0.094 | -0.002 |
| | (0.243) | (0.109) |
| W4 quiz | -0.042 | 0.211 |
| | (0.255) | (0.130) |
| W5 quiz | 0.027 | 0.053 |
| | (0.240) | (0.091) |
| Final exam | 0.046 | 0.028 |
| | (0.219) | (0.021) |
| Course grade | 0.197 | 0.047* |
| | (0.217) | (0.020) |

Each cell reports the standardized ITT estimate from a regression of the treatment (using the full sample) or number of videos watched as scheduled (using only the treatment sample) on the academic outcome indicated by each row. Regression models include 18 student-level covariates from institutional data and pre-course survey responses.
*$p < 0.05$

Table 3.5. Demographic and Pre-Course Survey Summary Statistics, Study 3.2

| | Full | | Control | | Treatment | |
|---|---|---|---|---|---|---|
| | Mean | SD | Mean | SD | Mean | SD |
| *Demographic characteristics* | | | | | | |
| Male | 0.253 | 0.435 | 0.259 | 0.440 | 0.247 | 0.433 |
| URM | 0.318 | 0.466 | 0.333 | 0.473 | 0.302 | 0.461 |
| Age | 20.147 | 1.891 | 20.197 | 1.918 | 20.098 | 1.868 |
| Low income | 0.392 | 0.489 | 0.401 | 0.492 | 0.383 | 0.488 |
| First generation | 0.515 | 0.501 | 0.519 | 0.501 | 0.512 | 0.501 |
| No English at home | 0.336 | 0.473 | 0.383 | 0.488 | 0.290 | 0.455 |
| Freshman | 0.194 | 0.396 | 0.204 | 0.404 | 0.185 | 0.390 |
| Sophomore | 0.519 | 0.500 | 0.531 | 0.501 | 0.506 | 0.502 |
| Junior | 0.179 | 0.384 | 0.167 | 0.374 | 0.191 | 0.395 |
| Senior or above | 0.108 | 0.311 | 0.099 | 0.299 | 0.117 | 0.323 |
| *N* | 324 | | 162 | | 162 | |
| *Pre-course survey items* | | | | | | |
| Expected grade on 4.0-scale | 3.807 | 0.334 | 3.805 | 0.357 | 3.810 | 0.312 |
| Expected hours per week | 8.961 | 4.217 | 8.895 | 3.880 | 9.025 | 4.536 |
| Keep record of assignments | 3.949 | 0.994 | 3.974 | 1.008 | 3.924 | 0.984 |
| Plan work in advance | 3.353 | 1.093 | 3.362 | 1.137 | 3.345 | 1.053 |
| Work where there are no distractions | 4.243 | 0.959 | 4.190 | 0.959 | 4.294 | 0.960 |
| Ignore distractions around me | 4.145 | 0.963 | 4.095 | 0.995 | 4.193 | 0.932 |
| Willing to use email to ask questions | 4.183 | 0.968 | 4.233 | 0.954 | 4.134 | 0.982 |
| Lazy or bored so quit (Reversed) | 2.396 | 0.974 | 2.431 | 0.971 | 2.361 | 0.981 |
| Work hard even if do not like task | 3.868 | 0.908 | 3.888 | 0.940 | 3.849 | 0.880 |
| When difficult give up (Reversed) | 2.838 | 0.982 | 2.819 | 1.043 | 2.857 | 0.923 |
| Even when dull keep working | 3.698 | 0.937 | 3.655 | 0.979 | 3.739 | 0.897 |
| *N* | 235 | | 116 | | 119 | |

All demographic variables except age are dummy variables. All pre-survey items except expected grade and expected hours per week are measured on a Likert scale from 1 (strongly disagree) to 5 (strongly agree). Two effort regulation items have been reverse-coded. None of the demographic of pre-course survey variables were found to be significantly different across treatment and control.

Table 3.6. Academic Performance Summary Statistics, Study 3.2

| | Full Sample (n = 324) | | Control (n = 162) | | Treatment (n = 162) | |
|---|---|---|---|---|---|---|
| | Mean | SD | Mean | SD | Mean | SD |
| Course grade (4.0-scale) | 2.964 | 1.057 | 2.831 | 1.107 | 3.097 | 0.989 |
| Week 1 quiz score (5 max) | 4.625 | 0.909 | 4.611 | 0.914 | 4.639 | 0.906 |
| Week 2 quiz score (5 max) | 4.610 | 0.828 | 4.509 | 1.030 | 4.710 | 0.542 |
| Week 3 quiz score (5 max) | 4.511 | 1.302 | 4.435 | 1.393 | 4.586 | 1.202 |
| Week 4 quiz score (5 max) | 4.454 | 1.229 | 4.395 | 1.253 | 4.512 | 1.206 |
| Week 7 quiz score (5 max) | 4.397 | 1.422 | 4.309 | 1.525 | 4.485 | 1.309 |
| Week 9 quiz score (5 max) | 4.403 | 1.440 | 4.349 | 1.518 | 4.457 | 1.360 |
| Week 10 quiz score (5 max) | 4.278 | 1.643 | 4.290 | 1.630 | 4.265 | 1.661 |
| Total homework score (100 max) | 85.514 | 19.197 | 85.259 | 19.477 | 85.769 | 18.971 |
| Total discussion score (100 max) | 84.566 | 16.928 | 83.778 | 17.650 | 85.355 | 16.191 |
| Presentation (30 max) | 27.245 | 6.055 | 27.201 | 6.205 | 27.290 | 5.920 |
| Research paper (75 max) | 64.601 | 19.178 | 63.773 | 20.230 | 65.429 | 18.090 |
| Midterm score (60 max) | 50.312 | 10.277 | 49.321 | 11.301 | 51.302 | 9.065 |
| Final exam score (60 max) | 45.324 | 11.353 | 43.735 | 12.032 | 46.914 | 10.427 |

Scores were collected at the end of the course. Course grade and total homework score have been adjusted to exclude extra credit awarded from the weekly surveys.

Table 3.7. Treatment Effect Estimates on Weekly Study Times, Study 3.2

| | Earliest Click over Week | | Click Spacing | | Earliest Click in a Day | |
|---|---|---|---|---|---|---|
| | M1 | M2 | M1 | M2 | M1 | M2 |
| Week 1 | -0.027 | -0.072 | 0.032 | 0.122 | 0.027 | -0.006 |
| | (0.110) | (0.106) | (0.109) | (0.106) | (0.113) | (0.134) |
| Week 2 | 0.044 | 0.005 | -0.075 | -0.060 | 0.201+ | 0.106 |
| | (0.108) | (0.107) | (0.110) | (0.114) | (0.112) | (0.137) |
| Week 3 | 0.067 | 0.013 | -0.136 | -0.036 | -0.026 | 0.038 |
| | (0.111) | (0.111) | (0.112) | (0.119) | (0.111) | (0.131) |
| Week 4 | -0.017 | 0.026 | -0.039 | -0.075 | -0.057 | 0.065 |
| | (0.113) | (0.121) | (0.113) | (0.116) | (0.112) | (0.130) |
| Week 5 | -0.069 | -0.037 | 0.161 | 0.252* | 0.039 | 0.106 |
| | (0.111) | (0.130) | (0.112) | (0.120) | (0.112) | (0.135) |
| Week 6 | -0.087 | -0.089 | -0.061 | 0.024 | 0.057 | 0.120 |
| | (0.112) | (0.111) | (0.111) | (0.119) | (0.112) | (0.135) |
| Week 7 | -0.044 | -0.014 | -0.009 | 0.019 | 0.162 | 0.135 |
| | (0.112) | (0.119) | (0.112) | (0.125) | (0.112) | (0.133) |
| Week 8 | -0.018 | -0.047 | -0.019 | 0.050 | 0.006 | -0.038 |
| | (0.111) | (0.118) | (0.113) | (0.123) | (0.113) | (0.135) |
| Week 9 | 0.073 | 0.114 | -0.150 | -0.147 | 0.128 | 0.165 |
| | (0.108) | (0.125) | (0.108) | (0.124) | (0.112) | (0.133) |
| Week 10 | -0.088 | -0.101 | 0.110 | 0.133 | 0.104 | 0.086 |
| | (0.109) | (0.121) | (0.109) | (0.128) | (0.112) | (0.137) |

Each cell reports the standardized ITT estimate from a regression of the treatment on how early students clicked in the course over the week, how spaced students worked on the online course, and how early students clicked in the course in a given day. Earliest click is the time of the earliest click in a given week or day. Click spacing is defined as the standard deviation of student clicks in a given week. Regression models M1 include 7 student-level covariates from institutional data (N = 324). Regression models M2 include 18 student-level covariates from institutional data and pre-course survey responses (N = 235).
*$p < 0.05$

Table 3.8. Treatment Effect on Academic Outcomes, Study 3.2

| | No Covariates (n = 324) | M1 (n = 324) | M2 (n = 235) |
|---|---|---|---|
| Course grade | 0.251* | 0.199* | 0.254* |
| | (0.110) | (0.099) | (0.099) |
| W1 quiz | 0.028 | 0.035 | -0.025 |
| | (0.101) | (0.100) | (0.113) |
| W2 quiz | 0.201* | 0.186* | 0.146 |
| | (0.091) | (0.091) | (0.100) |
| W3 quiz | 0.151 | 0.132 | 0.167 |
| | (0.145) | (0.144) | (0.150) |
| W4 quiz | 0.117 | 0.104 | 0.325* |
| | (0.137) | (0.134) | (0.157) |
| W7 quiz | 0.176 | 0.184 | 0.195 |
| | (0.158) | (0.155) | (0.165) |
| W9 quiz | 0.108 | 0.104 | 0.110 |
| | (0.160) | (0.161) | (0.175) |
| W10 quiz | -0.025 | -0.025 | 0.002 |
| | (0.183) | (0.182) | (0.209) |
| Quiz total | 0.146 | 0.139 | 0.177 |
| | (0.111) | (0.108) | (0.110) |
| Homework total | 0.027 | 0.011 | 0.089 |
| | (0.111) | (0.109) | (0.099) |
| Discussion total | 0.093 | 0.057 | 0.098 |
| | (0.111) | (0.106) | (0.099) |
| Presentation | 0.015 | -0.008 | 0.014 |
| | (0.111) | (0.109) | (0.105) |
| Paper | 0.086 | 0.065 | 0.126 |
| | (0.111) | (0.110) | (0.113) |
| Midterm | 0.193+ | 0.141 | 0.161 |
| | (0.111) | (0.106) | (0.101) |
| Final exam | 0.280* | 0.230* | 0.241** |
| | (0.110) | (0.093) | (0.093) |

Each cell reports the standardized ITT estimate from a regression of the treatment on scores indicated by each row. Regression models M1 include 7 student-level covariates from institutional data. Regression models M2 include 18 student-level covariates from institutional data and pre-course survey responses.
+$p < 0.10$, *$p < 0.05$, **$p < 0.01$

Table 3.9. Post-Course Survey, Study 3.2

| | ITT | n |
|---|---|---|
| Keep record of assignments | -0.198+ | 214 |
| | (0.115) | |
| Plan work in advance | -0.126 | 212 |
| | (0.122) | |
| Work where there are no distractions | -0.230+ | 214 |
| | (0.129) | |
| Ignore distractions around me | -0.016 | 214 |
| | (0.128) | |
| Willing to use email to ask questions | -0.123 | 213 |
| | (0.136) | |
| Lazy or bored so quit (Reversed) | -0.019 | 210 |
| | (0.127) | |
| Work hard even if do not like task | -0.006 | 210 |
| | (0.129) | |
| When difficult give up (Reversed) | 0.065 | 209 |
| | (0.124) | |
| Even when dull keep working | 0.041 | 205 |
| | (0.126) | |

Each cell reports the standardized ITT estimate from a regression of the treatment on post-survey responses indicated by each row. Post-survey items were measured on a Likert scale from 1 (strongly disagree) to 5 (strongly agree). Two effort regulation items have been reverse-coded. Regression models include 18 student-level covariates from institutional data and pre-course survey responses. N varied slightly for each model because not all students answered all post-course survey questions.

+p < 0.10

Table 4.1. Grade and Course Difficulty Summary Statistics

| | Full (n = 309) | | Control (n = 154) | | Treatment (n = 155) | |
|---|---|---|---|---|---|---|
| | Mean | SD | Mean | SD | Mean | SD |
| Subsequent GPA | 2.835 | 0.774 | 2.738 | 0.831 | 2.931 | 0.701 |
| Grade relative to classmates | -0.199 | 0.693 | -0.316 | 0.709 | -0.083 | 0.659 |
| Average grade 1 year prior | 3.015 | 0.315 | 3.045 | 0.333 | 2.985 | 0.294 |
| Average grade 2 years prior | 2.961 | 0.290 | 2.991 | 0.293 | 2.931 | 0.284 |

Grade relative to classmates is calculated by subtracting the average course grade from the students' actual grade. Each student has one score averaged from all courses. Average grade 1 and 2 years prior are the average grades awarded in the same courses in prior terms.

Table 4.2. Treatment Effect on Grades

| | GPA | | | Differential grade | | |
|---|---|---|---|---|---|---|
| | No covar. | Dem. included | Dem. & survey | No covar. | Dem. included | Dem. & survey |
| Treatment | 0.249* | 0.253* | 0.142 | 0.336** | 0.337** | 0.257+ |
| | (0.113) | (0.111) | (0.137) | (0.112) | (0.111) | (0.132) |
| Male | | -0.232+ | -0.351* | | -0.161 | -0.271 |
| | | (0.131) | (0.177) | | (0.131) | (0.171) |
| URM | | -0.338** | -0.383* | | -0.340** | -0.291+ |
| | | (0.125) | (0.162) | | (0.125) | (0.156) |
| Age | | 0.079 | -0.045 | | 0.001 | -0.136 |
| | | (0.063) | (0.089) | | (0.062) | (0.086) |
| Sophomore | | 0.388* | 0.516* | | 0.355* | 0.500* |
| | | (0.152) | (0.206) | | (0.151) | (0.199) |
| Junior | | 0.107 | 0.247 | | 0.006 | 0.140 |
| | | (0.190) | (0.258) | | (0.189) | (0.249) |
| Senior | | 0.325 | 0.565+ | | 0.046 | 0.218 |
| | | (0.233) | (0.328) | | (0.232) | (0.317) |
| Expected PH grade | | | 0.153* | | | 0.121+ |
| | | | (0.069) | | | (0.067) |
| *n* | 309 | 309 | 220 | 309 | 309 | 220 |

Demographic variables included are gender, underrepresented minority status, age, low-income status, first-generation status, whether English is spoken at home, and year in school. Survey covariates include 11 question items related to students' expected hours in public health, expected grades, and reported self-regulation skills.

+$p < 0.10$, *$p < 0.05$, **$p < 0.01$

Table 4.3. Treatment Effect on Course Enrollment

| | Average 1 year ago | | Average 2 years ago | | Number of Courses | |
|---|---|---|---|---|---|---|
| | Dem. included | Dem. & survey | Dem. included | Dem. & survey | Dem. included | Dem. & survey |
| Treatment | -0.177 | -0.263* | -0.202+ | -0.284* | -0.048 | -0.020 |
| | (0.109) | (0.133) | (0.109) | (0.133) | (0.077) | (0.093) |
| Male | -0.239+ | -0.367* | -0.181 | -0.303+ | -0.041 | 0.086 |
| | (0.129) | (0.171) | (0.127) | (0.171) | (0.091) | (0.120) |
| URM | -0.109 | -0.328* | -0.051 | -0.260+ | -0.255** | -0.295** |
| | (0.123) | (0.157) | (0.122) | (0.156) | (0.087) | (0.110) |
| Age | 0.217*** | 0.148+ | 0.241*** | 0.193* | -0.094* | -0.168** |
| | (0.062) | (0.087) | (0.061) | (0.087) | (0.043) | (0.061) |
| Sophomore | 0.081 | 0.166 | 0.189 | 0.386+ | 0.099 | 0.109 |
| | (0.149) | (0.200) | (0.150) | (0.201) | (0.105) | (0.140) |
| Junior | 0.329+ | 0.539* | 0.505** | 0.716** | 0.286* | 0.369* |
| | (0.187) | (0.249) | (0.187) | (0.251) | (0.132) | (0.175) |
| Senior | 0.472* | 0.972** | 0.634** | 1.081*** | -0.121 | 0.070 |
| | (0.229) | (0.317) | (0.228) | (0.318) | (0.162) | (0.222) |
| Expected PH grade | | 0.122+ | | 0.117+ | | 0.056 |
| | | (0.067) | | (0.067) | | (0.047) |
| *n* | 308 | 219 | 304 | 217 | 309 | 220 |

Demographic variables included are gender, underrepresented minority status, age, low-income status, first-generation status, whether English is spoken at home, and year in school. Survey covariates include 11 question items related to students' expected hours in public health, expected grades, and reported self-regulation skills. Sample size varies slightly for historical grades due to students taking newly added courses.

$+p < 0.10$, $*p < 0.05$, $**p < 0.01$, $***p < 0.001$

Table 4.4. Heterogeneous Effects of the Treatment

| | Without Survey Covariates | | | With Survey Covariates | | |
|---|---|---|---|---|---|---|
| | GPA | Diff. Grade | n | GPA | Diff. Grade | n |
| Male | 0.538* | 0.677** | 78 | 0.007 | 0.135 | 46 |
| | (0.224) | (0.222) | | (0.460) | (0.432) | |
| Female | 0.128 | 0.191 | 231 | 0.088 | 0.210 | 174 |
| | (0.128) | (0.126) | | (0.147) | (0.143) | |
| Freshman | 0.634* | 0.856** | 59 | 0.583 | 0.588 | 36 |
| | (0.262) | (0.268) | | (0.452) | (0.448) | |
| Soph. or above | 0.135 | 0.193 | 250 | -0.006 | 0.118 | 184 |
| | (0.121) | (0.119) | | (0.142) | (0.133) | |
| URM | 0.047 | 0.121 | 101 | 0.015 | 0.077 | 68 |
| | (0.229) | (0.220) | | (0.282) | (0.255) | |
| non-URM | 0.358** | 0.447*** | 208 | 0.217 | 0.343* | 152 |
| | (0.127) | (0.128) | | (0.162) | (0.159) | |
| Expected "A" | 0.229+ | 0.321* | 228 | 0.188 | 0.294+ | 150 |
| | (0.125) | (0.125) | | (0.159) | (0.151) | |
| Exp. below "A" | 0.517* | 0.566* | 81 | 0.334 | 0.405 | 70 |
| | (0.241) | (0.247) | | (0.275) | (0.292) | |
| Plan work (high) | 0.144 | 0.253+ | 181 | 0.134 | 0.256+ | 181 |
| | (0.144) | (0.142) | | (0.146) | (0.144) | |
| Plan work (low) | 0.341 | 0.459 | 39 | 0.357 | 0.423 | 39 |
| | (0.309) | (0.293) | | (0.518) | (0.483) | |
| Schedule high (high) | 0.095 | 0.238 | 176 | 0.047 | 0.206 | 176 |
| | (0.148) | (0.146) | | (0.150) | (0.149) | |
| Schedule work (low) | 0.657+ | 0.630+ | 44 | 0.806+ | 0.767+ | 44 |
| | (0.329) | (0.311) | | (0.428) | (0.385) | |

Each cell represents the ITT estimate of the treatment on the outcome variable indicated by each column. Rows indicate the subsample that regression models were built on.
+$p < 0.10$, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$

Table 4.5. Treatment Effect on Major

| | No Covariates | Demographics Included | Demographics & Survey |
|---|---|---|---|
| PH Major Start of Year | 0.019 | 0.014 | -0.013 |
| | (0.035) | (0.032) | (0.038) |
| PH Major End of Year | 0.108* | 0.090+ | 0.047 |
| | (0.051) | (0.047) | (0.058) |
| Changed to PH Major | 0.102* | 0.089* | 0.076 |
| | (0.045) | (0.043) | (0.056) |
| Changed Majors | 0.108* | 0.095+ | 0.146* |
| | (0.053) | (0.051) | (0.062) |
| *n* | 309 | 309 | 220 |

Each cell represents the ITT estimate of the treatment on the outcome variable indicated by each row. Columns indicate covariates that were include. In total there were 7 demographic variables (gender, underrepresented minority status, age, low-income status, first-generation status, whether English is spoken at home, and year in school) and 11 survey items related to students' expected hours in public health, expected grades, and reported self-regulation skills.
+$p < 0.10$, *$p < 0.05$

Table A4.1. Sample Demographic Summary Statistics

| | Continued | | | | | | Did not continue | |
|---|---|---|---|---|---|---|---|---|
| | All (n = 309) | | Control (n = 154) | | Treatment (n = 155) | | All (n = 15) | |
| | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| Male | 0.252 | 0.435 | 0.266 | 0.443 | 0.239 | 0.428 | 0.267 | 0.458 |
| URM | 0.327 | 0.470 | 0.344 | 0.477 | 0.310 | 0.464 | 0.133 | 0.352 |
| Age | 20.092 | 1.881 | 20.145 | 1.917 | 20.038 | 1.850 | 21.293 | 1.770 |
| Low-income | 0.405 | 0.492 | 0.422 | 0.496 | 0.387 | 0.489 | 0.133 | 0.352 |
| First generation | 0.508 | 0.501 | 0.519 | 0.501 | 0.497 | 0.502 | 0.667 | 0.488 |
| No English at home | 0.330 | 0.471 | 0.383 | 0.488 | 0.277 | 0.449 | 0.467 | 0.516 |
| Freshman | 0.191 | 0.394 | 0.195 | 0.397 | 0.187 | 0.391 | 0.267 | 0.458 |
| Sophomore | 0.537 | 0.499 | 0.552 | 0.499 | 0.523 | 0.501 | 0.133 | 0.352 |
| Junior | 0.184 | 0.388 | 0.169 | 0.376 | 0.200 | 0.401 | 0.067 | 0.258 |
| Senior | 0.087 | 0.283 | 0.084 | 0.279 | 0.090 | 0.288 | 0.533 | 0.516 |

No significant differences were found between treatment and control groups. All but age are dummy variables.

# APPENDIX C: BIG FIVE INVENTORY ITEMS BY WEEK

**Reference:**
John, O. P., & Srivastava, S. (1999). The Big Five trait taxonomy: History, measurement, and theoretical perspectives. *Handbook of personality: Theory and research*, 2(1999), 102-138.

The following items were administered at the beginning of each week. Students were asked to respond to each statement on a scale of 1 (disagree strongly) to 5 (agree strongly).

Week 1:
I see myself as someone who…

1. is talkative
2. tends to find fault with others
3. does a thorough job
4. is depressed, blue
5. is original, comes up with new ideas
6. is reserved
7. is helpful and unselfish with others
8. can be somewhat careless

Week 2:
I see myself as someone who…

1. is relaxed, handles stress well
2. is curious about many different things
3. is full of energy
4. starts quarrels with others
5. is a reliable worker
6. can be tense
7. is ingenious, a deep thinker
8. generates a lot of enthusiasm

Week 3:
I see myself as someone who…

1. has a forgiving nature
2. tends to be disorganize
3. worries a lot
4. has an active imagination
5. tends to be quiet
6. is generally trusting
7. tends to be lazy
8. is emotionally stable, not easily upset

Week 4:
I see myself as someone who…
1. is inventive
2. has an assertive personality
3. can be cold and aloof
4. perseveres until the task is finished c
5. an be moody
6. values artistic, aesthetic experiences
7. is sometimes shy, inhibited
8. is considerate and kind to almost everyone

Week 5:
I see myself as someone who…
1. does things efficiently
2. remains calm in tense situations
3. prefers work that is routine
4. is outgoing, sociable
5. is sometimes rude to others
6. makes plans and follows through with them
7. gets nervous easily
8. likes to reflect, play with ideas

Week 6:
I see myself as someone who…
1. has few artistic interests
2. likes to cooperate with others
3. is easily distracted
4. is sophisticated in art, music, or literature
5. is talkative (II)
6. tends to find fault with others (II)
7. does a thorough job (II)
8. is depressed, blue (II)

Week 7:
I see myself as someone who…
1. is original, comes up with new ideas (II)
2. is reserved (II)
3. is helpful and unselfish with others (II)
4. can be somewhat careless (II)
5. is relaxed, handles stress well (II)
6. is curious about many different things (II)
7. is full of energy (II)
8. starts quarrels with others (II)

Week 8:
I see myself as someone who…
1. is a reliable worker (II)
2. can be tense (II)
3. is ingenious, a deep thinker (II)
4. generates a lot of enthusiasm (II)
5. has a forgiving nature (II)
6. tends to be disorganize (II)
7. worries a lot (II)
8. has an active imagination (II)

Week 9:
I see myself as someone who…
1. tends to be quiet (II)
2. is generally trusting (II)
3. tends to be lazy (II)
4. is emotionally stable, not easily upset (II)
5. is inventive (II)
6. has an assertive personality (II)
7. can be cold and aloof (II)
8. perseveres until the task is finished c (II)

Week 10:
I see myself as someone who…
1. an be moody (II)
2. values artistic, aesthetic experiences (II)
3. is sometimes shy, inhibited (II)
4. is considerate and kind to almost everyone (II)
5. does things efficiently (II)
6. remains calm in tense situations (II)
7. prefers work that is routine (II)
8. is outgoing, sociable (II)

Week 11:
I see myself as someone who…
1. is sometimes rude to others (II)
2. makes plans and follows through with them (II)
3. gets nervous easily (II)
4. likes to reflect, play with ideas (II)
5. has few artistic interests (II)
6. likes to cooperate with others (II)
7. is easily distracted (II)
8. is sophisticated in art, music, or literature (II)