Modeling Students' Procrastination in Higher Education: Causes, Outcomes, and Prediction

Tianchen Sun

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2023

Reading Committee:

Ji-Eun Kim, Chair

Shuai Huang

Prashanth Rajivan

Program Authorized to Offer Degree:
Industrial and Systems Engineering

University of Washington

Abstract

Modeling Students' Procrastination in Higher Education: Causes, Outcomes, and Prediction

Tianchen Sun

Chair of the Supervisory Committee:
Ji-Eun Kim
Industrial and Systems Engineering

Students spend little time completing tasks when deadlines are far off; however, they tend to increase their work amounts as deadline approaches. This phenomenon, which is called deadline rush, can be modeled by exponential distributions. Deadline reactivity, represented by a rate parameter of the exponential distribution, parameterizes individual differences in procrastination. That is, an individual with high reactivity to deadlines procrastinates more than an individual with low deadline reactivity. While the phenomenon and parametric models of individual differences in procrastination have been investigated, practical applications in the classroom setting have garnered little attention from researchers. Past research on procrastination has not much considered its relationships with learning environment factors and academic performance, with a lack of objective measurements and heavy reliance on self-reported questionnaires.

My dissertation will respond to this gap in the research by modeling students' individual procrastination in university classroom settings, paying close attention to factors influencing the students' procrastination as well as the effects of procrastination on performance. In particular, the dissertation will answer the following three research questions: (1) Do learning environments (i.e., online learning, task complexity, and time in the academic term) affect students' procrastination? (2) Does procrastination affect individual and team per-

formance? and (3) How can procrastination be predicted through physiological responses (i.e., eye movement, heart rate, electrodermal activity, and skin temperature)? The first two research questions have been answered by longitudinal field studies, while controlled laboratory experiments are conducted to answer the third research question. My findings shed light on how objective modeling and prediction of procrastination can be applied in the classroom setting. In particular, the findings will provide instructors, researchers, and online learning platforms with practical strategies to better design classes and interventions of procrastination for improvements in students' performance.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Table Number Page

GLOSSARY

A: Original, non-discounted value of the outcome in deadline rush model

AAWC: Anti-Air Warfare Coordinator

ANOVA: Analysis of Variance

APA: American Psychological Association

API: Aitken Procrastination Inventory

API: Application Programming Interface

AUC: Area Under Curve

BVP: Blood Volume Pulse

$^\circ C$: Degree Celsius

χ^2: Chi squared index

CFI: Comparative Fit Index

CI: Confidence Interval

D: Time to deadline in deadline rush model

DT: Decision Tree

EDA: Electrodermal Activity

ERP: Event-Related Potential

f^2: Effect Size Index

FN: False Negative

FP: False Positive

GPA: Grade Point Average

GTZERO: Score of anagram difficulty

HR: Heart Rate

k: Deadline reactivity in deadline rush model

KNN: K-Nearest Neighbor

LMS: Learning Management System

M: Mean

MBTI: Myers-Briggs Type Indicator

μS: Micro Siemens

MSE: Mean Squared Error

No.: Number

nW: Nano Watt

OSHA: Occupational Safety & Health Administration

p: P value

PACED: Pacing Action Categories of Effort Distribution

PASS: Procrastination Assessment Scale-Students

PPG: Photoplethysmography

R^2: Coefficient of Determination

ρ: Correlation coefficient

RBF kernel: Radial Basis Function kernel for support vector machine

RF: Random Forest

RMS: Root Mean Square

RMSEA: Root Mean Square Error of Approximation

ROC: Receiver Operating Characteristic

RQ: Research Question

SCL: Skin Conductance Level

SCR: Skin Conductance Response

SD: Standarad Deviation

SE: Standard Error

SEM: Structural Equation Modeling

SRMR: Standardized Root Mean Squared Residual

STEM: Science, Technology, Engineering, and Math

SVM: Support Vector Machine

TEMP: Skin Temperature

TLI: Tucker-Lewis Index

TMT: Temporal Motivation Model

TN: True Negative

TP: True Positive

V: Current discounted value of the outcome in deadline rush model

W: Mann-Whitney U test W statistic

WHO: World Health Organization

ACKNOWLEDGMENTS

First and foremost, I would like to express my deepest gratitude to my advisor, Dr. Ji-Eun Kim, for your invaluable patience and mentoring throughout my journey pursuing PhD. Your expertise and selfless guidance steered me in the right direction through this journey over the years. You have set a role model as a researcher, mentor, and instructor that I will always look up to.

I could not have undertaken this journey without my committee members, Dr. Shuai Huang, Dr. Prashanth Rajivan, and Dr. Andrea Stocco. I am grateful for your insightful comments and for taking the time to serve as my committee members.

I recognize my research collaborators over the years, Auður Anna Jónsdóttir, Dr. Ameer Hamza Shakur, and others. I also owe thanks to all the undergraduate researchers that I have worked with, Roger Huynh, Jessica Hwee, Glenndi Tjuandi, and Yaocheng Zhuang. This endeavor would not have been possible without your invaluable contributions.

Additionally, I would like to extend my appreciation to my friends and colleagues, Yiting Chen, Jiaxin Li, Lun Li, Pariyakorn Maneekul, Wengeng Pan, Xiaonan Sun, Jinlin Xiang, Yilun Xing, Tianhao Xu, and Xinyi Xu. I deeply thank all of you for your support and assistance to me in my research and in my life over the years.

Lastly, I am deeply indebted to my parents, Baolin Sun and Yu Yu, for your unconditional and endless support. I would never attain the achievements without your continuous encouragement. Special thanks to my cat, Huskie, for all the joy and emotional support you provided.

Chapter 1

INTRODUCTION

Almost all of us can recall a time in our lives that we have procrastinated (Ackerman and Gross, 2007; Fernie et al., 2017). Students may procrastinate and submit an exam late (Hong et al., 2021); faculty members may delay when writing manuscripts and proposals (Ackerman and Gross, 2007); customers frequently postpone the completion of their online purchases (Negra et al., 2008); and individuals wait until the last moment to file their tax returns (Martinez et al., 2022). Procrastination is an easily observable and prevalent phenomenon in the general population (Akram et al., 2019; Chu and Choi, 2005; Ferrari et al., 1992; Steel, 2007; Solomon and Rothblum, 1984). Researchers have investigated individual differences in procrastination (Gevers et al., 2015, 2016), finding that although some individuals may start working as soon as tasks are assigned, others may not start until right before the due date.

Procrastination can usually be observed on individuals who have poor time management and indecisiveness (Steel, 2007; Knaus, 2000). Workers who constantly procrastinate would assign the majority of workload close to task deadlines, which would result in an increase of time pressure due to high workload under limited deadline. Consequently, such time pressure is one of the leading causes of work-related stress (Michie, 2002). According to Occupational Safety & Health Administration (OSHA), approximately 83% of US workers reported to be suffered from work-related stress in 2022 (OSHA, 2022). Among these workers, 54% indicated that stress have negatively impacted their daily lives (OSHA, 2022). Additionally, around a third of the US adults surveyed by American Psychological Association (APA) reported that they were stressful with daily decision making during the pandemic (APA, 2021). Stress not only negatively impacts individuals' work performance, but would also lead to consequences of poor mental and physical health (Michie, 2002; Rose et al., 2017), resulting in trillions of dollars costs for companies globally per year according to World Health Organization (WHO; WHO, 2022). Therefore, it is necessary to study individuals' procrastination to reduce their work-related stress for their general well-being and improvements in work performance.

While chronic procrastination affects approximately 15% to 25% adult population worldwide (Harriott and Ferrari, 1996; Nguyen et al., 2013; Steel, 2007), studies have reported that procrastination is especially prevalent among college students. Approximately 46% to 95% of college students have experienced procrastination (Gafni and Geri, 2010; Fernie et al., 2017; Özer et al., 2009; Solomon and Rothblum, 1984). Students may see direct consequences of procrastination in their course grades and grade-point averages (GPAs). Procrastination can cause serious trouble for students since their calendars are usually full of assignment deadlines. Failure to complete assignments and projects in a timely manner significantly and negatively impacts students' academic performance, which then lowers their self-efficacy and leads to more procrastination (Steel, 2007). By studying procrastination, researchers can develop better interventions for reducing individuals' procrastination and hence improving their productivity and performance.

Chapter 2

LITERATURE REVIEW

In this chapter, I first introduce a theoretical framework that defines and explains procrastination. For the rest of the chapter, I summarize the past literature about the causes, outcomes, and quantification methods of procrastination. At the end of each section, I describe the research gaps in each corresponding research area and introduce the research questions that I aim to answer in this dissertation.

2.1 Theoretical framework of procrastination

Procrastination is defined as the tendency to postpone an intended course of action related to starting or completing tasks (Chu and Choi, 2005; Ferrari et al., 1992; Steel, 2007). Procrastination can be represented by individuals' work pace under deadlines. Individuals given shorter deadlines accomplish tasks at a faster rate than those given tasks with longer deadlines (Bryan and Locke, 1967). Studies have shown that individuals' work pace increases as deadlines approach. For instance, some researchers found that the participants paid more attention to the time and increased their work pace as the time remaining to deadlines decreases (Waller et al., 2002).

A theoretical framework that explains procrastination is the temporal motivation model (TMT; Steel and König, 2006; Steel, 2007). This model is derived from the consolidation of expectancy theory and hyperbolic discounting. Equation 2.1 illustrates the formula for TMT. Inside the formula, utility represents how motivated an individual is to perform a course of actions. Higher utility shows greater motivation. $(E \times V)$ shows the perceived values an individual expects towards a task, considering both the individual's expectancy and perceived values towards the tasks. Z is a constant indicating the rewards of the tasks.

$(T-t)$ describes the delay of the tasks between "time now" and "time to receive rewards." Finally, Γ is a parameter referring to individuals' reactivity to the delay of the tasks. TMT assumes that individuals' motivation (utility) on completing tasks changes depending on their expectancy (E) and value V towards the tasks. When multiple tasks were available, individuals' expectancy (E) and perceived values V on the tasks may differ. In addition, such motivation would be weakened by the delay of the tasks $(T-t)$. As a deadline approaches, the delay of the task shortens and their motivation for the task would raise. Individuals will prefer to conduct the task whose overall utility is higher. TMT considers individual differences in sensitivity to time delay (Γ) to be a factor representing various procrastination tendencies (Steel, 2007; Wu et al., 2016).

$$\text{Utility} = \frac{E \times V}{Z + \Gamma(T-t)} \tag{2.1}$$

2.2 Causes of students' procrastination

Extensive research on the factors influencing individuals' procrastination has been conducted. The factors can be generally categorized as individual differences and task characteristics (Steel, 2007). Individual differences include personality traits, characteristics, and demographics. Researchers have found that individuals' traits and characteristics are significant factors influencing their procrastination. For instance, some individual personality types are significantly related to individual procrastination (Ferrari et al., 1992; Dewitte and Schouwenburg, 2002; Schouwenburg and Lay, 1995). Ferrari et al. (1992) found that procrastination is negatively correlated with the judging personality type and positively correlated with the perceptive personality type according to the Myers-Briggs Type Indicator (MBTI) scale. Schouwenburg and Lay (1995) surveyed 352 participants and concluded that their procrastination levels were negatively related to conscientiousness and neuroticism based on the Big Five factors of personality. Researchers have also argued that self-efficacy and self-esteem are directly linked to procrastination (Chu and Choi, 2005; Owens and Newbegin, 1997; Steel, 2007; Tuckman, 1991). The relationship between gender and procrastination is controversial. While some researchers have found that women have lower procrastination tendencies than men (Kim and Nembhard, 2019b; Özer et al., 2009), others have found no significant differences in procrastination by gender (Ferrari et al., 1992; Kachgal et al., 2001; Solomon and Rothblum, 1984).

The characteristics of tasks themselves are also known to affect individuals' procrastination. Task characteristics refer to environmental factors that influence procrastination. According to researchers, the time to deadline has been found to be a significant factor influencing individuals' procrastination (Ackerman and Gross, 2007; Levy et al., 2013; Steel, 2007; Paden and Stell, 1997). Less-distant tasks are usually considered more important and urgent than tasks with later deadlines (Paden and Stell, 1997). As a result, individuals put more effort into completing the less-distant tasks. Task complexity is another factor related to procrastination (Ackerman and Gross, 2007; Kim et al., 2016; Paden and Stell, 1997). On the one hand, Ackerman and Gross (2007) surveyed 244 faculty members and reported that the faculty members were more willing to procrastinate on tasks when they perceived the tasks to be more difficult. On the other hand, Kim et al. (2016) analyzed participants' procrastination on an Anti-Air Warfare Coordinator (AAWC) task and found a negative correlation between task complexity and procrastination.

Previous studies have investigated the factors influencing procrastination in terms of individual differences and task characteristics. However, few studies have considered the causes of procrastination and how procrastination manifests in the university classroom setting. The learning environments and instruction technologies in university classrooms have advanced rapidly in recent decades. The effects of learning environments on individual procrastination remain underexplored. Therefore, I propose the first research question.

RQ I: What factors affect students' procrastination in university classrooms?

2.3 Outcomes of students' procrastination

Extensive research has been done on the effects of procrastination. One of the topics that has most interested researchers is the effects of procrastination on performance. While some researchers have stated that procrastination can be treated as a performance-enhancing strategy (Chu and Choi, 2005; Steel, 2007; Tice and Baumeister, 1997), other researchers have shown that procrastination is negatively related to individual academic performance (i.e., grades and GPAs) in the university classroom setting (Elvers et al., 2003; Kim and Nembhard, 2019b; Michinov et al., 2011; Owens and Newbegin, 1997; Widjaja and Tarigan, 2017; You, 2015). In addition to displaying weaker academic performance, students with higher levels of procrastination have been reported to have lower satisfaction with their classes (Elvers et al., 2003) and a greater likelihood to drop out of the classes in the middle of a semester (Michinov et al., 2011).

Although multiple studies have uncovered a negative relationship between procrastination and students' individual academic performance, very few studies have considered team performance. It remains unclear whether the findings regarding procrastination negatively influencing individual performance can be generalized to team performance. Therefore, I propose a second research question.

RQ II: Does procrastination affect students' individual and team performances?

2.4 Quantification of procrastination

2.4.1 Measurement of procrastination

Researchers have heavily relied on subjective measurements to identify individual differences in procrastination. Most of the existing research measures procrastination using self-reported questionnaires, including the Aitken Procrastination Inventory (API; Aitken, 1982; Moon and Illingworth, 2005; Owens et al., 2008; Pychyl et al., 2000), the Procrastination Assessment Scale–Students (PASS; Ferrari et al., 1992; Kachgal et al., 2001; Owens and Newbegin, 1997; Özer et al., 2009; Solomon and Rothblum, 1984), and the Pacing Action

Categories of Effort Distribution (PACED; Gevers et al., 2015). Although the reliability of such questionnaires has been discussed and supported by several studies over the past few decades, self-reported questionnaires present a methodological limitation (Ackerman and Gross, 2005, 2007; Chu and Choi, 2005; Lay, 1986; Schouwenburg and Lay, 1995). The results based on such questionnaires may be inaccurate due to self-reported bias (Yi et al., 2023): namely, participants may report what they would like to do instead of what they actually do.

Another commonly used measurement of procrastination is the length of time between the time of task completion and the task deadline (Gafni and Geri, 2010; Janssen and Carton, 1999; Levy et al., 2013; Moon and Illingworth, 2005). Intuitively, an individual is considered to have procrastinated less if the individual completes a task relatively early. However, the time proximity of task completion to deadline is a discrete measurement and cannot accurately track individuals' actual work behavior before task completion.

2.4.2 Modeling of procrastination: deadline reactivity

A reliable objective modeling approach that explains procrastination was introduced as the deadline rush model (Kim et al., 2016; Konig and Kleinmann, 2005), which was built upon the theoretical framework of TMT (Chapter 2.1). The model assumes that individuals discount the value of future tasks, applying a higher discount rate to later tasks and a lower discount rate to earlier tasks. When the time to deadline for a task is long, an individual may assign low value to the task and choose to delay working on it. On the other hand, when the deadline is approaching, an individual may assign increased value to the task and prioritize it. According to the model, an individual would procrastinate completing tasks if the perceived values of tasks are lower than the values for the other activities, such as entertainment (Figure 2.1). The exponential deadline rush model shown in Equation (2.2) describes the relationship between procrastination and the value an individual assigns to a particular task. V indicates the current discounted value of the outcome; A indicates the original, non-discounted value of the outcome; D is the deadline; and k is each individual's deadline reactivity, which represents their procrastination. The larger the k value, the more an individual procrastinates. The deadline rush model provides a quantifiable parameter of k that indicates individual differences in procrastination based on continuous observations of behavior.

The deadline rush model can be applied to model procrastination when the tasks had fixed deadlines. Konig and Kleinmann (2005) introduced the modeling approach to computing students' deadline reactivity through their longitudinal learning activities captured by course webpage from 97 students in four different courses. Besides the task scenarios of college learning environments, the deadline rush model has also been applied by researchers in determining individuals' procrastination when completing tasks in a self-training program (Häfner et al., 2014), and in an Anti-Air Warfare Coordinator task (Kim et al., 2016).

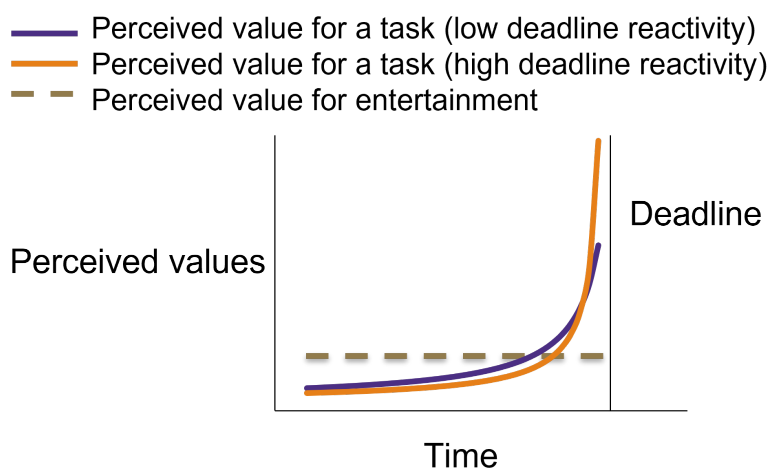

Figure 2.1: Illustration of perceived values by deadline rush model.

$$V = Ae^{-kD} \tag{2.2}$$

To simplify the computation in Equation (2.2), a modified deadline rush model is introduced in Equation (2.3). This simplified model allows us to calculate a student's deadline reactivity by fitting their daily activity (Kim and Nembhard, 2018; Sun and Kim, 2022a). Both the original (Equation (2.2)) and the simplified (Equation (2.3)) versions of the deadline rush model are based on exponential probability distribution. The difference between the models is that the original model sets two free parameters, A (non-discounted value of the outcome) and k (deadline reactivity), while the modified model sets one free parameter, k. Recent studies have shown that k values calculated using the simplified model are not significantly different from those calculated using the original model and that the simplified model maintains small mean square errors of data fitting (Kim and Nembhard, 2018; Sun and Kim, 2022a).

$$V = ke^{-kD} \tag{2.3}$$

2.4.3 Prediction of procrastination

The current methods of measuring and modeling procrastination have their advantages and drawbacks. Self-reported questionnaires can capture individual procrastination instantaneously, but the subjective results may not accurately capture the participants' thoughts (Yi et al., 2023). The task completion time as well as the deadline rush model may objectively reveal an individual's behavioral procrastination, but the computation is asynchronous and may require a large dataset, which is costly. Moreover, the measurement and modeling approaches could only identify individual differences in procrastination after task completion, which is limited in practical applications if interventions are required to reduce procrastination. Therefore, a prediction model for procrastination in real-time prior to task completion is worth exploring.

Physiological responses have been frequently applied in prediction of individuals' cognitive states, behaviors and performance in many studies (McDonald et al., 2020; Smyth et al., 2021; Wu et al., 2021; Yi et al., 2023). In recent years, the development of sensors and wearable devices has made the acquisition of physiological responses much easier and more affordable (Carter and Luke, 2020). Compared with surveys that are distributed in the middle of experiments, collection of physiological data is less intrusive and the continuous data collected may carry objective and longitudinal information (Yi et al., 2023). In addition, prediction models using physiological responses are advantageous because the prediction can be applied in real-time (Baltaci and Gokcay, 2016). However, approaches using physiological responses to predict procrastination has lacked explored. Thus, I propose the third research question.

RQ III: How can procrastination be predicted prior to task completion?

2.5 Research purpose

In this dissertation, I will fill the research gaps identified and answer the research questions proposed in the aforementioned chapters. In particular, I will model students' procrastination in the university classroom setting, focusing on the influences of learning environments on procrastination and the effects of procrastination on individual and team performance.

The chapters of the dissertation proposal are organized as follows and displayed in Table 2.1. **RQ I** is addressed in Chapter 3 and 4, and **RQ II** is answered in Chapter 5. Chapter 6 summarizes the findings for both **RQ I** and **RQ II**. In Chapter 7, I introduce an exploratory study that answers **RQ III**.

Table 2.1: Chapters and corresponding research questions addressed

Chapter	**Research question addressed**
3. Causes: online learning and task complexity	RQ I
4. Causes: online learning and time in the academic term	RQ I
5. Outcomes: individual and team performance	RQ II
6. Integrated model of causes and outcomes	RQ I, II
7. Quantification: physiological modeling	RQ III

Chapter 3

CAUSES: ONLINE LEARNING AND TASK COMPLEXITY

Online learning has been gaining the attention of students and instructors in recent years, a phenomenon that has only increased now that most courses are being delivered online due to the ongoing COVID-19 pandemic. However, challenges remain for students to properly manage their time and achieve satisfactory grades in the online learning environment. In this chapter, I answer to the **RQ I** and investigate how the online learning environment and varying levels of task complexity affect students' procrastination and academic performance. I collected 157 college students' online activity data. The results show that the students procrastinated more in the online learning environment than in the face-to-face setting; they likewise procrastinated more when completing high-complexity tasks than low-complexity tasks. A significant interaction effect shows that the students reacted differently to deadlines for high- and low-complexity tasks in the online learning environment than they did to deadlines in the face-to-face setting. My findings highlight the importance of taking into account course assignment complexity when designing face-to-face and online classes.

3.1 Introduction

Online learning has become increasingly popular in recent years, especially now that most courses are being delivered online due to the COVID-19 pandemic (Adnan and Anwar, 2020; Dhawan, 2020; Hodges et al., 2020; Rapanta et al., 2020). Online learning can be defined as a learning experience accessed via the internet and that occurs in either a synchronous or asynchronous environment without any constraints on students' physical location (Dhawan, 2020). To facilitate social distancing during the pandemic, both instructors and students are taking advantage of teaching and learning via the online environment

(Hodges et al., 2020).

For many students, time management is one of the biggest impediments to successful online learning. Students who are used to face-to-face learning may find planning and managing their studies difficult in the online learning environment (You, 2015). Students in the online learning environment are usually given substantial time and flexibility on assignments; however, some students cannot adequately manage their time to finish the work (Dhawan, 2020). For example, Muilenburg and Berge (2005) conducted a survey of 1,056 students and found that one of the most frequently mentioned barriers to online learning was a lack of time and support for studying. Elvers and colleagues (2003) similarly concluded that students' procrastination in online learning courses was related to low satisfaction in the classes.

3.1.1 Procrastination in the online learning environment

Procrastination in the online learning environment has garnered much attention recently, but very few studies have directly compared students' procrastination levels in face-to-face and online learning environments based on the deadline rush model. Instead, several studies have used self-reported questionnaires or discrete temporal information. For example, using the self-reported Academic Procrastination Scale with a sample of 88 students, Yilmaz (2017) found no significant differences in students' procrastination levels between the face-to-face and online learning environments. Elvers et al. (2003) calculated students' procrastination levels by measuring the amount of time between when students first accessed particular course material and when they took the corresponding exam and found that the students' procrastination levels were not significantly different across learning environments. Such nonsignificant differences in procrastination across the two learning environments may be due to a lack of rigor in measuring procrastination. Indeed, although previous studies have failed to find differences in procrastination across the two learning environments, several studies have reported students' reduced efficiency and satisfaction in the online learning environment because of the students' difficulty with time management

(Elvers et al., 2003; Hong et al., 2021). This shows the need for a direct comparison of students' procrastination levels in the face-to-face and online learning environments using the deadline rush model.

3.1.2 Procrastination and task complexity

One reason that previous studies have found only nonsignificant differences in procrastination across the two learning environments may be the different course types the studies have investigated and the studies' failure to account for various levels of task complexity. Understanding the importance of task complexity to procrastination requires a clear and cogent definition of task complexity. Among the various ways to define task complexity are two especially frequently used approaches. The first way to define task complexity is to determine the number of elements in a given task (Campbell and Gingrich, 1986; Liu and Li, 2012; Wood, 1986). A task is considered complex if it can be split into multiple subtasks or if several distinct actions need to be executed to complete the task. The second way to define task complexity is to identify the expected volume of output associated with the task (Bonner, 1994; Funke, 1991; Liu and Li, 2012). A task that has substantial expected output is considered more complex than a task with little expected output.

Very few studies have considered how task complexity affects procrastination. Kim and colleagues (2016) used an Anti-Air Warfare Coordinator (AAWC) simulator to design decision-making tasks with different levels of complexity and found that the participants procrastinated more when reacting to low-complexity compared to high-complexity targets. As the task environment in the aforementioned study relied on an anti-air warfare coordinator simulator, however, whether the findings can be generalized to the online learning environment is uncertain.

3.1.3 Gaps in previous studies and research goal

Despite the importance of understanding procrastination in online learning, there remain critical research gaps. The first research gap is that to the best of the authors' knowledge, little consideration has been given to the influence of both the learning environment and task complexity on academic procrastination. The aforementioned studies that found a relationship between task complexity and procrastination did not consider the online learning environment.

The objective of this chapter is to investigate the effects of both the learning environment (face-to-face and online learning) and task complexity (low and high) on students' procrastination levels and academic performance. Rather than considering one learning environment or one level of task complexity, this chapter aims to understand how students' learning is impacted by diverse learning environments as well as levels of task complexity. In this chapter, I rely on the objective measurements of course webpage activity and assignment submission time to model procrastination.

3.2 Methodology

3.2.1 Participants

One hundred and fifty-seven undergraduate students (92 males, 65 females) from the University of Washington participated in this study. Of those students, 102 (66 males, 36 females) took the study's selected courses in 2019 when the instruction was delivered face-to-face; 55 students (26 males, 29 females) took the courses in 2020 when the instruction was delivered entirely online. This research was approved by the Institutional Review Board at the University of Washington. I obtained consent from each student at the beginning of the respective academic quarter.

3.2.2 Study design

The independent and dependent variables are listed in Table 3.1. This chapter used a 2x2 mixed design in which the between-subject variable was the learning environment and the within-subject variable was task complexity. To measure the dependent variables, I collected students' daily page views, as well as their submission time and grades for all assignments. These data were collected from course pages hosted by the Canvas learning management system (LMS) through its application programming interface (API).

Table 3.1: Independent and dependent variables

Independent variable	**Levels**
Learning environment (between-subject)	Face-to-face vs. online
Task complexity (within-subject)	Low vs. high complexity
Dependent variable	
Deadline reactivity	
Time proximity of submission to deadline	
Academic performance	

Learning environment

The learning environment was defined as the delivery method for the course content: face-to-face or online. The study took place in the context of two undergraduate-level courses: "Human Factors in Design" (coded as Course A) and "Statistical Quality Control" (coded as Course B) in the years 2019 and 2020. Both courses were offered as a face-to-face in 2019 and an online learning course in 2020. While the instructor who taught the Course A was different from the instructor who taught the Course B, the same instructor taught the same classes in the years 2019 and 2020. For both courses, students were required to attend classes for approximately 4.5 hours per week; they met in the classroom for the face-to-face setting and accessed Zoom and recorded videos for the online setting. In both learning environments, the students were able to browse their respective course pages on the Canvas

LMS and access the announcements, syllabus, instructional slides, assignment descriptions, and their grades, among other content. Students who missed an assignment deadline were penalized in their grade on that particular assignment.

Task complexity

Students completed assignments with both low and high task complexity. I followed existing research in defining task complexity in accordance with either the number of elements of a given task (Campbell and Gingrich, 1986; Liu and Li, 2012; Wood, 1986) or the number of the expected volume of output (Bonner, 1994; Funke, 1991; Liu and Li, 2012).

In the Course A, I identified 11 homework assignments (five in the face-to-face setting and six in the online learning environment) that each required a one-page case study report. I also identified seven project assignments (three in the face-to-face setting and four in the online learning environment) that each required the deliverables of a project proposal, final project presentation, and final project report. I coded the homework assignments as low-complexity tasks and the project assignments as high-complexity tasks because there were fewer task elements associated with the homework assignments (three required components) than the project assignments (more than 10 required components; Campbell and Gingrich, 1986; Liu and Li, 2012; Wood, 1986). Moreover, the expected volume of output for the homework assignments (one page per assignment) was smaller than the expected volume of output for the project assignments (usually more than three pages per assignment; Bonner, 1994; Funke, 1991; Liu and Li, 2012). I used the five days immediately prior to each of the homework deadlines and the seven days immediately prior to each of the project assignment deadlines for analysis. I discarded one homework assignment in the face-to-face setting and three homework assignments in the online learning environment that were fewer than five days away from an exam because I anticipated that students may have used this period to simultaneously work on the assignment and study for the exam.

Students enrolled in the Course B had 15 homework assignments (seven in the face-to-face setting and eight in the online learning environment). Each homework assignment included several quality control questions requiring statistical calculations, thus marking the assignment as a high-complexity task. Students also had two project assignments (one in the face-to-face setting and one in the online learning environment), each of which was a presentation summarizing a scholarly paper. The project assignments were considered less complex than the homework assignments since the project assignments required fewer task elements and less output (five task elements and approximately 10 pages of slides) than the homework assignments (usually more than 10 task elements and more than five pages of written calculations plus lines of R code; Bonner, 1994; Campbell and Gingrich, 1986; Funke, 1991; Liu and Li, 2012; Wood, 1986).

Deadline reactivity

In this study, I used a simplified deadline rush model in Equation (2.3) to calculate each student's deadline reactivity by fitting their online page view frequency. This model assumes that A, the non-discounted value assigned to the task based on the original deadline rush model in Equation (2.2), is identical to the value of k, the deadline reactivity. Recent studies have shown that the k values calculated using the simplified model are similar to those calculated using the original model while maintaining small mean square errors of data fitting (Kim and Nembhard, 2018; Sun and Kim, 2022a).

Time proximity of submission to deadline

Another common measurement of procrastination is the time proximity of submission to deadline, representing the time gap in minutes between the deadline and the time that a student submitted their last attempt at the assignment (Akram et al., 2019; Gafni and Geri, 2010; Levy et al., 2013). Time proximity of submission to deadline is calculated by taking the average amount of time between a student's final submission time and the deadline of a given assignment. It measures how much time that the student allots for the assignment before the assignment is due. Intuitively, the smaller the amount of time between the submission time and the deadline, the later the final submission is, suggesting more procrastination. Time proximity of submission to deadline was used as a secondary measurement of procrastination.

Academic performance

Academic performance is the average number of points out of 100 that a student received for their assignments. It is measured by taking the summation of the points the student received for all the low- or high-complexity assignments and dividing that number by the maximum possible points for all these assignments. The assignments were graded by the course instructors using standard grading rubrics.

3.3 Results

To determine whether ANOVA is suitable for use in the data analysis, I checked the normality assumption for the dependent variables using the Shapiro-Wilk test. All the variables violated the assumption under certain experimental conditions. Therefore, I adopted a nonparametric approach to data analysis, using the 'nparLD' package in R (Noguchi et al., 2012).

The nonparametric mixed-design variance analysis showed a significant main effect of learning environment ($F = 17.30$, $p < 0.001$) and task complexity ($F = 5.95$, $p < 0.05$) on deadline reactivity and a significant interaction effect between learning environment and task complexity ($F = 20.37$, $p < 0.001$), as shown in Figure 3.1. That is, students' deadline reactivity was significantly higher in the online learning environment than in the face-to-face setting; it was likewise significantly higher when conducting high-complexity tasks than low-complexity tasks. The difference in students' procrastination between high- and low-complexity tasks was greater in the online learning environment than in the face-to-face learning environment.

Figure 3.2 shows a significant main effect of task complexity ($F = 10.37$, $p < 0.01$)

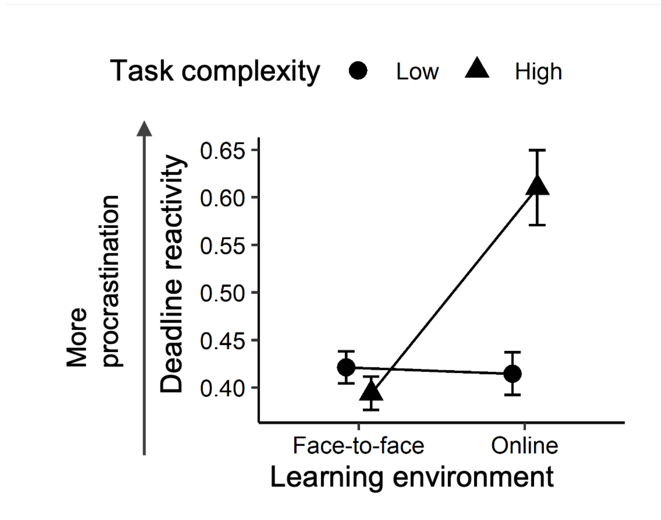

Figure 3.1: Interaction plot of deadline reactivity with regard to learning environment and task complexity. Error bars indicate standard error.

and a significant interaction effect of task complexity and learning environment ($F = 3.87$, $p < 0.05$) on time proximity of submission to deadline. Students submitted assignments much closer to the deadline for high-complexity tasks than low-complexity tasks. The difference in students' time proximity of submission between high- and low-complexity tasks was greater in the online learning environment compared to the face-to-face learning environment. The effect of the learning environment did not show a significant effect on time proximity of submission to deadline at $\alpha = 0.05$.

I found a significant main effect of task complexity only on academic performance ($F = 48.68$, $p < 0.001$), as shown in Figure 3.3. Students received higher grades on low-complexity tasks than high-complexity tasks regardless of the learning environment. However, the effect of the learning environment did not show any significant effect on academic performance at $\alpha = 0.05$. The interaction term for task complexity and learning environment did not reach significance at $\alpha = 0.05$.

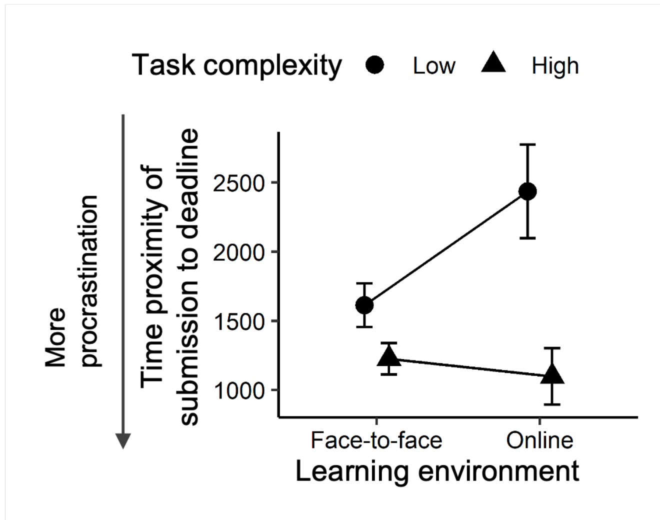

Figure 3.2: Interaction plot of time proximity of submission to deadline with regard to learning environment and task complexity. Error bars indicate standard error.

3.4 Discussions and conclusions

3.4.1 Effect of learning environment and task complexity on procrastination

When measured by deadline reactivity, students' procrastination levels were found to be higher in the online learning environment than in the face-to-face setting (Figure 3.1). However, this relationship was not significant when procrastination was measured by time proximity of submission to deadline (Figure 3.2). The results related to time proximity of submission to deadline are similar to those found in previous research that determined procrastination based on the amount of time between when students first accessed particular course materials and when the students took the exam (Elvers et al., 2003) and by self-reported questionnaires (Yilmaz, 2017). I note that unlike the time proximity to deadline, which is a discrete dataset by nature such that one data point is available under one deadline, deadline reactivity measures procrastination using a series of continuous data points before assignment and project deadlines over a period of time. Measurements of discrete time gaps alone may not capture as much procrastination information as deadline reactivity, thereby

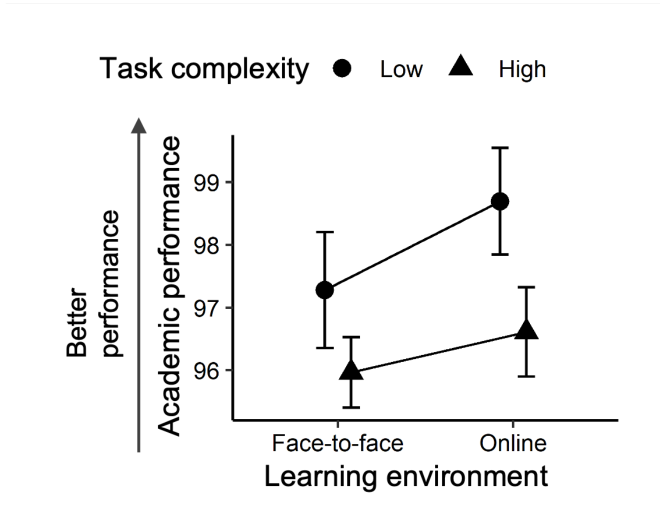

Figure 3.3: Interaction plot of academic performance with regard to learning environment and task complexity. Error bars indicate standard error.

explaining why past research has not found significant differences in procrastination across the face-to-face and online learning environments.

When I measured procrastination by both deadline reactivity (Figure 3.1) and time proximity of submission to deadline (Figure 3.2), I found that students procrastinated more when completing high-complexity tasks than low-complexity tasks. This finding contradicts those of Kim and colleagues (Kim et al., 2016). Kim and colleagues found a negative relationship between task complexity and procrastination. Possible reasons for this inconsistency in findings are the studies' different task types as well as the time available to complete tasks. In my study, students had either days or weeks to complete the course assignments; in other words, they had sufficient time to put off high-complexity tasks.

The finding that differences in procrastination for low- and high-task complexity tasks are larger in the online learning environment compared to the face-to-face learning environment (Figures 3.1 and 3.2) implies that course instructors might consider additional aids for online learners who are completing high-complexity tasks. For example, instructors might offer extended deadlines or issue multiple reminders of difficult tasks to help online learners better allocate their time to succeed in the classes.

3.4.2 Effect of learning environment and task complexity on academic performance

The finding that students received better grades on low-complexity tasks than high-complexity tasks regardless of whether they were taking the face-to-face or online version of a course is in line with the findings of previous studies (Bonner, 1994; Cho, 2018; Kim et al., 2016). That is, students' performance increased as task complexity decreased. This implies that students may have had enough time to prepare for the low-complexity tasks, which led to better academic performance, given that the low-complexity tasks included fewer elements and a smaller expected volume of output following both definitions of task complexity. The consistency in findings between this study and previous studies supports this study's design of task complexity in both the face-to-face and online learning environments.

Although I did not find any significant main effects of learning environment on students' academic performance when tasks were analyzed by complexity, when I conducted a Mann-Whitney U Test ($W = 1118$, $p < 0.001$), I discovered that students' total cumulative grades, as calculated by the weighted summation of points on all assignments and exams, were significantly higher in the online learning environment than they were in the face-to-face environment. This might be explained by the fact that both courses were intended to be delivered face to face and the learning environment was shifted online due only to the COVID-19 pandemic. Instructors may have lower expectations for students' volume of work and the quality of deliverables in the online learning environment and therefore be more generous during final grading (Johnson et al., 2020). Another possible explanation is that the quizzes and exams in the face-to-face setting were closed-book and closed-notes, while in the online setting, students were allowed access to all the course materials. The freedom to access all course materials while completing quizzes and exams may have decreased the difficulty of the assessments, thereby contributing to the differences in academic performance observed in the two learning environments.

3.4.3 Relationship between procrastination and academic performance

I additionally computed the repeated-measure correlation between the two dependent variables, i.e., procrastination and academic performance. I found a significant positive relationship between time proximity of submission to deadline and academic performance ($r = 0.217$, $p < 0.01$). That is, students who submitted assignments early tended to earn higher grades on their assignments compared to those who submitted assignments close to the deadlines. This finding is in line with those of previous studies indicating that more procrastination leads to poorer academic performance (Michinov et al., 2011; Yilmaz, 2017). Importantly, when I divided the time proximity of submission to deadline data into two groups based on the face-to-face and online learning environments, I found a significant relationship between time proximity of submission to deadline and academic performance for those students in the online learning environment ($r = 0.345$, $p < 0.01$), but not for those in the face-to-face learning environment ($r = 0.155$, $p > 0.05$). This result implies that students' submission time may be a predictor of their academic performance, especially in the online learning environment.

3.4.4 Study Limitations and Future Work

The limitations of this study suggest future research directions. Firstly, despite our quantitative measurement and modeling of procrastination, I were limited in our data collection process by our inability to track the learning behaviors of those participants who may have downloaded the materials from the Canvas LMS all at once and worked on the assignments offline. Secondly, most of our participants were students majoring in the same department of the same university, and the generalizability of the results might be limited. The diversity of participants' backgrounds was likewise relatively low. Thirdly, there were some students who took both courses in this study, and I were unable to model the potential effects of this overlap in the samples under different conditions. I note that there were five students taking one of the two courses in the face-to-face setting and the other course in the online learning environment. Due to the small sample size, I were not able to analyze these students' differences in procrastination across the two environments. Lastly, the instructors' different teaching styles may have influenced the students' procrastination and performance. I suggest that future researchers use larger and more diverse samples of participants to conduct controlled experiments that manipulate students' procrastination levels in both the face-to-face and online learning environments, thereby helping course designers to create courses that facilitate students' success.

Chapter 4

CAUSES: ONLINE LEARNING AND TIME IN THE ACADEMIC TERM

Past research investigated the changes in students' procrastination primarily in face-to-face settings and reported mixed findings relying on self-reported questionnaires and discrete measurements such as submission times. In this chapter, I build on past research, considering both face-to-face and online learning environments when investigating changes in procrastination over an academic term (**RQ I**). My findings show that in both the face-to-face and online learning environments, students' procrastination increased as the term progressed. I also construct predictive models of deadline reactivity with time proximity of submission to deadline, learning environments, and time in the academic term. My results suggest that the instructors' efforts to intervene in students' procrastination would be more required in the second half of the academic term when procrastination is highest.

4.1 Introduction

Research has investigated how various task-related factors influence students' procrastination. For example, distant deadlines (Paden and Stell, 1997), high task complexity (Ackerman and Gross, 2007), and online learning environments (Sun and Kim, 2022b) have been shown to be indicators of greater procrastination. However, such research treats procrastination as a trait of students that is constant over time. That is, the studies assume that students procrastinate at a similar level from the beginning to the end of an academic term.

4.2 Related works

Procrastination has been treated as a unique and consistent characteristic that is associated with individuals' personality traits (Steel, 2007; van Eerde and Klingsieck, 2018; Van Hooft and Van Mierlo, 2018). For instance, Schouwenburg and Lay (1995) conducted a survey of 352 students and found that the students' procrastination levels were negatively related to their conscientiousness and neuroticism, traits that are based on the Big-Five personality structure. Lower self-efficacy, self-esteem, and self-regulation were also found to be associated with greater procrastination (Michinov et al., 2011; Steel, 2007). Given that adults' Big-Five personality traits remain consistent over a four-year period (Cobb-Clark and Schurer, 2012), university students' procrastination levels over an academic term likewise have been assumed to be consistent, with researchers computing the levels using data across multiple terms (Kim and Nembhard, 2019b; Konig and Kleinmann, 2005; Sun and Kim, 2022a; Yilmaz, 2017).

Very few studies have investigated changes in students' procrastination over time. Of those that that have, there have been mixed findings. Özer and Saricaoglu (2014) surveyed 114 students on their procrastination scales at the beginning and the end of the semester and found no significant differences. Other researchers have found that students' self-reported procrastination is significantly higher at the end of the semester than at the beginning or middle of the semester (Fincham and May, 2021; Scheunemann et al., 2022). Moon and Illingworth (2005) tracked students' procrastination longitudinally based on the time a take-home test was made available to students and the time the students completed the test. They found that students' procrastination rose during the first four exams and then dropped for the last exam.

Studies about changes in procrastination have been conducted primarily in face-to-face instruction settings without due consideration for the online learning environment (Moon and Illingworth, 2005; Özer and Sarıcaoğlu, 2014). Our previous findings have shown that students tend to procrastinate more when in an online learning environment than in a face-to-face setting (Sun and Kim, 2022b). However, there is little evidence indicating how the learning environment influences students' temporal changes in procrastination. Based on this gap in the existing literature, I pose the first research question for this chapter (**RQ4.1**):

RQ4.1: Do learning environments affect temporal changes in students' procrastination?

As discussed in Chapter 2.4, the modeling methods using deadline rush model have its pros and cons. Despite its strength in quantifying individual differences in procrastination (Kim et al., 2016; Kim and Nembhard, 2018, 2019b; Sun and Kim, 2022b), the deadline rush model is limited by its complexity in modeling and the huge amount of data required to compute reliable parameters. Previous studies have adopted data from an entire academic term (more than 10 weeks) to model students' deadline reactivities (Kim et al., 2016; Sun and Kim, 2022b). However, such data collection and modeling procedures are incapable of modeling procrastination in real-time, which is useful for providing instant feedback to and developing interventions for learners in academic settings.

Identifying a predictive relationship between deadline reactivity and time proximity of submission to deadline could help researchers obtain a rough estimate of deadline reactivity using a less time-consuming quantification approach. Our previous findings showed that time proximity of submission to deadline can distinguish differences in procrastination between low- and high-complexity tasks. However, this measurement was not as effective as deadline reactivity when comparing students' procrastination between the face-to-face and online learning environments. A predictive relationship could mitigate the limitations of the time proximity of submission to deadline measurement to represent students' longitudinal academic procrastination and simultaneously reduce the amount of data and the computational steps required for modeling students' deadline reactivity. To investigate the possibility of such an approach, I propose our second research question (**RQ4.2**):

RQ4.2: Can students' deadline reactivities be predicted by time proximity of submission to deadline?

This chapter investigates the dynamic nature of students' academic procrastination while taking into account the course delivery mode. I conducted a field study to measure students' procrastination tendencies using students' learning activity data during the first and second halves of an academic term that was either completely in-person or completely online. Uncovering the relationship between students' procrastination and time in the academic term could provide insights useful for the design of practical procrastination interventions.

4.3 Methods

4.3.1 Participants

A total of 104 students from two undergraduate classes at the University of Washington (N = 58 in Course A with 23 females and 35 males; N = 46 in Course B with 13 females and 33 males) participated in our 2019–2021 field study. All of the participants were sophomores or juniors with an engineering major. I recruited participants with similar ages and educational backgrounds to minimize the potential nuisance effects of uncontrolled variables on our results (Sun and Kim, 2022a). Both Course A and Course B are offered during the 10-week spring term from April to June. At the beginning of each term, informed consent was obtained from each participant. Each student received an extra credit on their final 100-point course grade for their participation in the study. The study was approved by the Institutional Review Board at the University of Washington.

4.3.2 Data collection

At the end of each academic term, I retrieved the participants' activity data on the course websites created through the Canvas learning management system (LMS). The Canvas LMS was used as an integrated platform that stored and displayed to students all course materials, including announcements, assignments, and grades. During the academic term, students needed to log on to the appropriate course website to check for course updates and submit their assignments. The activity data that I collected were the participants' daily number of views of their particular course website and their submission time for each assignment. I assumed that the students' daily course website log-on activity represented their daily learning behaviors in the course (Sun and Kim, 2022a).

4.3.3 Data analysis

Independent variable 1: learning environment

The learning environment was either face-to-face or online. The classes were instructed completely in person in 2019 (N = 71 with 24 females and 47 males). In 2020 and 2021, the classes were delivered fully online via Zoom (N = 33 with 12 females and 21 males) due to COVID-19.

Independent variable 2: time in the academic term

Time in the academic term was either the first or second half of the term. I purposely split the 10-week academic term into halves because splitting the data into more than two segments would have resulted in an insufficient amount of data in each segment for robust deadline rush model results. For consistency in our data analysis, I used the participants' activity around homework assignments that were due before Week 4 to model their deadline reactivity in the first half of the term. Students' activity around homework assignments that were due after Week 6 were used to model their deadline reactivity in the second half of the term. Activity related to homework assignments due between Week 4 and Week 6 was not considered in the modeling since the midterm exams were scheduled during this period. The activity data of the last homework assignment for each course were discarded for all three years of the study because other course assignments' deadlines and exam dates overlapped with those of the homework assignments.

Dependent variable 1: deadline reactivity

I adopted the deadline rush model (Equation 2.2) to compute participants' deadline reactivities, which represent individual differences in procrastination under a deadline (Sun and

Kim, 2022a). Given that the participants had approximately a week to complete each assignment, I fit each student's five-day historical number of page views recorded on the corresponding course website before a deadline to the deadline rush model. Each student's page-view datasets were transformed into a histogram indicating the frequency of page views of the course website before a deadline, as shown in Figure 4.1. Figure 4.1 illustrates a 2021 Course A participant's daily number of page views over the second half of the term. The histograms were then fit into Equation (2.2) to identify the deadline reactivity value (k) that yielded the minimum mean squared errors using the Microsoft Excel Solver built-in.

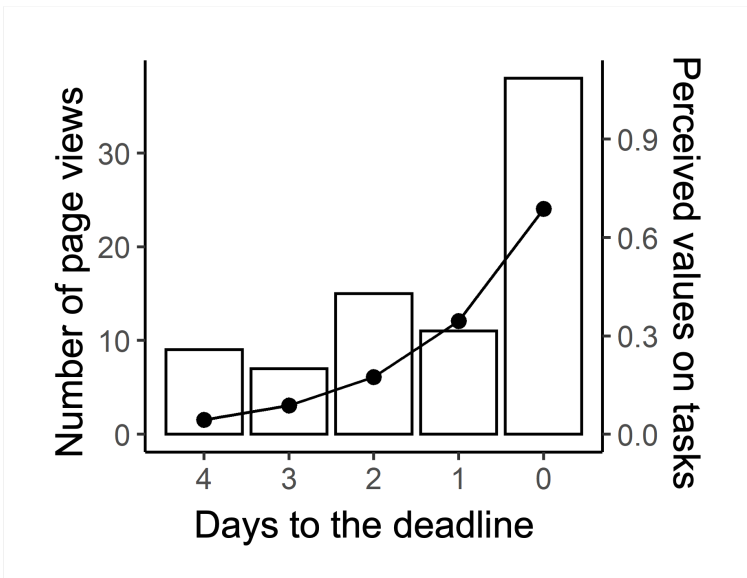

Figure 4.1: Illustration of data fit of the deadline rush model for a sample 2021 Course A participant in 2021 in the second half of the academic term ($k = 0.688$). The bars represent the number of page views each day (left y-axis). The line plot represents the participant's expected value of the task each day computed from the deadline rush model (right y-axis).

Dependent variable 2: time proximity of submission to deadline

A secondary measurement of students' academic procrastination was time proximity of submission to deadline, defined as the time gap between assignment submission and deadline (Equation 4.1). I computed the time proximity of submission to deadline using the same assignments that I used for modeling the participants' deadline reactivities in the first and second halves of the academic term. The average time proximity of submission to deadline for the selected assignments in each half of the academic term was calculated as a measurement of participants' procrastination over that half of the academic term. The later a participant submitted assignments, the smaller the value of time proximity of submission to deadline, which I took as an indicator of greater procrastination.

$$\text{Time proximity of submission to deadline } (minutes) = \text{Time of deadline} - \text{Time of submission} \tag{4.1}$$

Statistical analysis

Since the numbers of datasets in the two levels of both independent variables were unbalanced, I conducted a nonparametric mixed-design ANOVA to compare the students' deadline reactivities and time proximity of submission to deadline under the two learning environments and times in the academic term (i.e., first half and second half) using the 'nparLD' package (Noguchi et al., 2012) in R (R Core Team, 2020). To develop the predictive model for deadline reactivity, I constructed a linear regression model in R (R Core Team, 2020). A Type I error of 0.05 was used to determine statistical significance.

4.4 Results

4.4.1 Nonparametric mixed-design ANOVA

The results of the nonparametric mixed-design ANOVA illustrate the relationship between the dependent and independent variables. Our findings showed that deadline reactivity was significantly affected by both the time in the term ($F = 11.389$, $p < 0.001$) and learning environment ($F = 24.740$, $p < 0.001$). As shown in Figure 4.2, students had greater deadline reactivity in the second half of the term and when learning in the online learning environment. The results indicate that the students procrastinated more as the academic term was about to end than they did at the beginning of the term. The students also procrastinated more when studying online than taking in-person classes. The interaction effect of time in the term and learning environment on deadline reactivity was not significant ($F = 0.820$, $p = 0.365$).

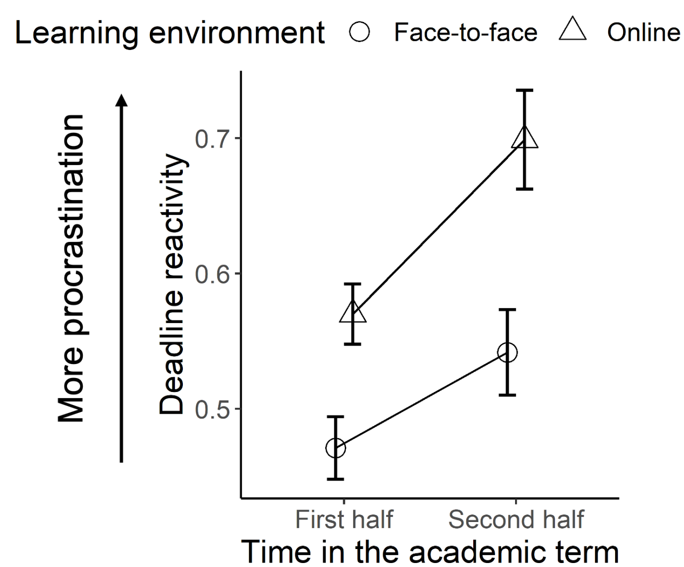

Figure 4.2: Interaction and box plots of students' deadline reactivities in each half of the term according to learning environment. The cross signs represent the average of the students' deadline reactivities in the given condition.

I also found a significant main effect of time in the term on time proximity of submission to deadline ($F = 24.765$, $p < 0.001$), indicating that students submitted assignments sooner in the first half of the term than in the second half of the term. Figure 4.3 shows the interaction and box plots of students' time proximity of submission to deadline in the first and second halves of the academic term for the face-to-face and online learning environments. I did not find a significant main effect of the learning environment on time proximity of submission to deadline ($F = 0.016$, $p = 0.897$). The interaction effect of time in the academic term and learning environment on time proximity of submission to deadline was not significant ($F = 0.862$, $p < 0.353$).

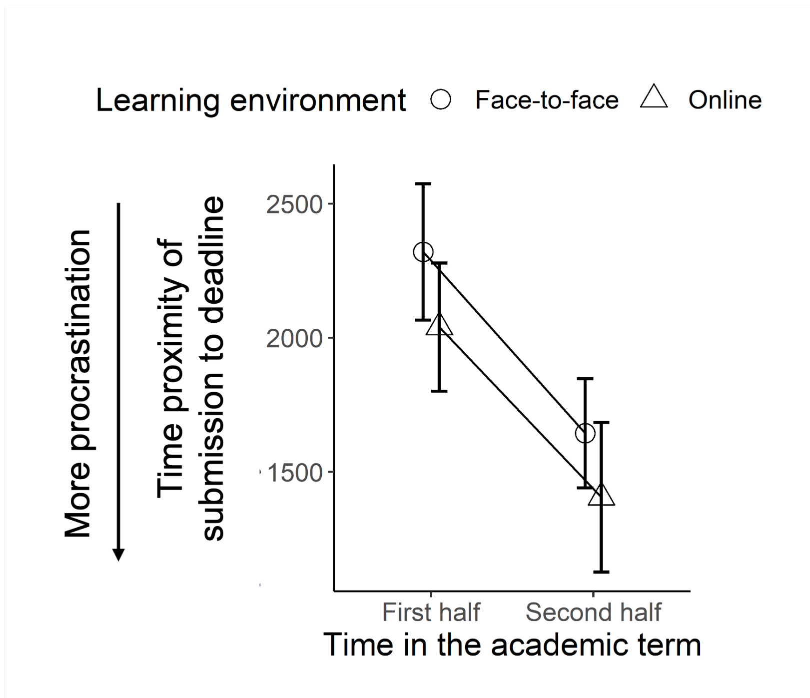

Figure 4.3: Interaction and box plots of students' time proximity of submission to deadline in the first and second halves of the academic term according to learning environment. The cross signs represent the average of the students' time proximity of submission to deadline in the given condition.

4.4.2 Linear regression

A linear regression model was constructed to predict students' deadline reactivities with time proximity of submission to deadline, time in the academic term, and learning environment. Time in the term was coded as either "1" or "2," representing the first or second half of the term, respectively. The learning environment was dichotomized as either "1" or "2," representing the in-person or online learning environment, respectively.

Table 4.1 shows the regression coefficients of the linear regression model ($F(3,204) = 24.27$, $p < 0.001$, $R^2 = 0.252$). The linear regression model was used to predict deadline reactivity using three variables: time proximity of submission to deadline, time in the term, and learning environment. Higher values of deadline reactivity were predicted when the submission time was close to the deadline in the second half of the term for students studying in the online learning environment. This pattern is consistent with the nonparametric variance analysis results in Chapter 4.4.1. The model's goodness-of-fit index was 0.252. That is, the independent variables in the model can explain approximately 25.2% of student deadline reactivity.

Table 4.1: Regression coefficients of linear regression model for predicting deadline reactivity.

	beta	**SE**	***t***	***p***
Time proximity of submission to deadline (10^{-5})	−4.990	0.761	−6.562	< 0.001
Time in the academic term	0.056	0.028	2.019	0.045
Learning environment	1.152	0.029	3.917	< 0.001

Notes: The baseline condition for "time in the academic term" is the first half of the term. The baseline condition for "learning environment" is the in-person setting. SE = standard error.

4.5 Discussions

The present study investigated the relationship between students' procrastination and time in the academic term with consideration for the learning environment. Our results indicate that students' procrastination, as measured by either deadline reactivity or time proximity of submission to deadline, increased as the academic term progressed. Such changes in procrastination over time were found in both the in-person and online learning environments. In response to **RQ4.1**, our findings show that students' procrastination was dynamic over time and that changes in procrastination could be observed during a 10-week term in both the in-person and online learning environments.

Several factors may have caused this increase in students' procrastination in the second half of the academic term. First, it is inevitable that students experience increasing workload as the academic term progresses (Ruiz-Gallardo et al., 2011) due to concurrent deadlines of homework assignments, course projects, and exams toward the end of multiple courses. Studies have shown that individuals tend to procrastinate more when the perceived workload is high (Ackerman and Gross, 2007). Second, students may be less motivated to study (Darby et al., 2013) and experience more stress (Fincham and May, 2021) as the academic term progresses. Individuals with lower levels of motivation have been reported to procrastinate more (Afrashteh and Koohneshin Seighalani, 2021; Darby et al., 2013).

The significant main effects of time in the academic term on deadline reactivity contradict the previous findings in literatures, whereas the significant main effects of learning environment on deadline reactivity corroborate our previous findings. On the one hand, our findings about the dynamic nature of procrastination contradict the findings in the literature that failed to capture temporal changes in procrastination (Özer and Sarıcaoğlu, 2014). This discrepancy may be due to differences in experimental designs and methods of quantifying procrastination. While I quantified students' procrastination with a modeling approach using students' learning activity before a deadline, other studies have measured procrastination using self-reported questionnaires (Özer and Sarıcaoğlu, 2014). On the other hand, our finding that students' deadline reactivities were higher in the online learning environment than in the in-person setting are in line with our previous findings (Sun and Kim, 2022a). As students may have more flexibility and reduced interaction with their classmates in the online learning environment, they may have more difficulty appropriately managing their time to complete assignments and study for exams (Dhawan, 2020; Gonda et al., 2021).

I answered **RQ4.2** by constructing a linear regression model to predict deadline reactivity using students' time proximity of submission to deadline. The model confirmed the variance analysis results that show students procrastinated more during the second half of the term and when learning in online learning environments. Our model serves as a preliminary approach to estimating student deadline reactivity with easily accessible submission time data. In addition, the model shows the potential of predicting students' deadline reactivity in the later period of the academic term through the data collected from the beginning of the academic term. Although the three variables significantly predicted deadline reactivity in our regression model, I acknowledge that the goodness of fit of our model ($R^2 = 0.252$) could be improved by incorporating individual traits such as gender (Kim and Nembhard, 2019b) and self-regulation (Zarrin et al., 2020) and contextual variables such as task complexity (Ackerman and Gross, 2007).

The findings from the present study imply that instructors may need to intervene more in the second half of an academic term when students' procrastination is known to be higher. I analyzed the effects of time in the term on students' academic performance, with academic performance defined by a student's average grades on assignments. In line with past studies (Michinov et al., 2011; Kim and Nembhard, 2019b; Sun and Kim, 2022b), students' performance was found to be significantly higher and their procrastination lower in the first half of the term ($F = 12.654$, $p < 0.001$). The proposed prediction model in the present study could help instructors to predict each student's changes in procrastination during the academic term. Customized intervention strategies could be applied to mitigate students' individual differences in changes in procrastination over time. For instance, instructors might consider providing stress management to help students stay motivated in learning (Kachgal et al., 2001; Van Hooft and Van Mierlo, 2018) and mediate the increase in their procrastination (Afrashteh and Koohneshin Seighalani, 2021). In addition, various types of calls to action, such as applying peer pressure and providing incentives, could assist students in reducing procrastination effectively (Huang et al., 2021).

The present study has a few limitations that suggest the need for future research. First, the courses in the online learning environment had both live lectures and self-paced, prerecorded lectures. It is not clear whether students would have procrastinated more if the online courses had been fully live or fully prerecorded. Future research might compare procrastination levels when students attend live lectures or watch prerecorded lectures. Second, modeling students' procrastination with the Canvas LMS data may not have fully represented students' procrastination behavior, especially for those who primarily downloaded all course materials and studied offline. I cannot be certain whether the students logged on the course website to work on the homework assignments or to study for the exams. Future research might explore quantifying procrastination by students' self-reported daily learning time. Lastly, due to the short length of the academic term, I were only able to split the students' activity data into halves. Segmenting activity data into more than two time periods would have resulted in insufficient data for the modeling of students' deadline reactivities. Acquiring course activity data for an extended period of time would allow researchers to segment data into more than two groups, which might lead to a more robust analysis of changes in procrastination.

Chapter 5
OUTCOMES: INDIVIDUAL AND TEAM PERFORMANCE

The goal of this chapter is to respond to the **RQ II** and investigate the effect of procrastination heterogeneity on team performance. While researchers have investigated the causes and correlates of individual differences in procrastination, very few studies have paid attention to how to form teams considering procrastination heterogeneity among team members to maximize team performance. Further, such research has relied on subjective questionnaires as a measurement of individual procrastination and team performance. In this chapter, I collected data on the daily page views of two course websites. Course A included 38 individuals comprising 15 dyads or triads, while Course B included 55 individuals comprising 20 dyads or triads. I fitted the students' page hits per day to a deadline rush model that quantified individuals' procrastination using an exponential function and employed a structural equation modeling (SEM) approach to investigate the relationships among individual procrastination, procrastination heterogeneity in teams, and team performance. The results show that homogeneous teams of procrastination perform better than heterogeneous teams, regardless of whether the homogeneous teams are composed of high-procrastinated individuals or low-procrastinated individuals. Team variation in procrastination fully mediated the relationship between individual procrastination and team performance. The findings can be applied to team formation mechanisms used by organizations, including schools and industries, to maximize overall team performance.

5.1 Introduction

Teamwork plays an important role in shaping organizational structure (Barrick et al., 1998). For this reason and in light of the increasingly diverse workforce, the formation of teams merits considerable investigation. An inadequate understanding of heterogeneity among diverse team members and in relation to team formation may result in team conflict and reduced team performance. Therefore, team-based organizations need to consider individuals' characteristics as well as the aggregate characteristics of the individuals who comprise a team (Mohammed and Angell, 2003).

Team composition research investigates the diverse characteristics of individuals in teams (Stewart, 2006). One area of team composition research considers whether the heterogeneity of team members' characteristics predicts team performance (Kramer et al., 2014; Stewart, 2006). Much of this research relies on variance in team members' characteristics as indicators of team composition, in accordance with the perspective that individual characteristics are not combined in a linear manner (Kristof-Brown and Stevens, 2001; Stewart, 2006). Researchers have explored demographic factors such as age and gender (Baugh and Graen, 1997; Clement and Schiereck, 1973; Timmerman, 2000), personality characteristics such as the Big Five traits and the Myers-Briggs Type Indicators (MBTI Boyle, 1995; Coe, 1992; Mohammed and Angell, 2003; Volkema Gorman, 1998), and skill sets such as roles and abilities (Barrick et al., 1998; Higgs et al., 2005; Hooper and Hannafin, 1991).

While researchers have considered the demographic factors, personalities, and skillsets of team members as indicators of team heterogeneity, few have investigated procrastination heterogeneity among team members and its effect on team performance. While researchers have investigated the causes and correlates of individual differences (Steel, 2007), they have paid relatively little attention to how to form teams considering procrastination homogeneity and heterogeneity to maximize team performance. I note that time-management related problems such as scheduling and time allocation have been found to cause interpersonal conflicts during teamwork, thereby resulting in reduced team performance (Mohammed and Nadkarni, 2011). This chapter aims to consider procrastination heterogeneity as team heterogeneity and investigate its influence on team performance.

5.1.1 The effect of member heterogeneity and homogeneity on team performance

There are mixed findings as to whether homogeneous or heterogeneous teams produce better team performance (Bowers et al., 2000; Van Knippenberg and Schippers, 2007). Teams' homogeneity and heterogeneity in various aspects such as team members' demographic factors and personalities have been found to be related to teams' performance (Bowers et al., 2000; Van Knippenberg and Schippers, 2007). Some researchers have found that homogeneity in age, gender, or ability produces better team performance. In a study that surveyed teams composed of individuals from different functional areas of a company, the various team members and leaders were asked about team composition and the effectiveness with which their respective teams completed their tasks (Baugh and Graen, 1997). Those in single-gender or single-race teams reported that they performed more effectively than members of mixed-gender or mixed-race teams reported that they did. Another study experimented on groups that were either single-gender or mixed-gender in composition (Clement and Schiereck, 1973). The members were asked to complete forced-choice visual signal detection tasks. The results showed that the all-male and all-female groups performed equally well and that both groups had higher correctness rates on the signal detection tasks than the mixed-gender groups did. Hooper and Hannafin (1991) found that team performance is related to homogeneity in team members' abilities. Homogeneous dyad teams composed of two high-ability students had better scores than heterogeneous dyad teams composed of one high-ability student and one low-ability student.

In contrast, some studies have found heterogeneous teams to be more productive and effective. For instance, Mello and Ruckes (2006) developed a model that when information was shared within teams, the heterogeneous teams composed of individuals with different ages, gender, and social background, had the potential to outperform the homogeneous teams (Mello and Ruckes, 2006). Hooper and Hannafin (1991) similarly found that for low-ability students who performed poorly on mathematics tests, heterogeneous teams performed better than homogeneous teams. In Van Offenbeek (2001)'s study, diversity in team members' attitudes toward tasks is effective in facilitating learning. In this study, teams with greater diversity in team members' attitudes showed higher self-rating scores on how much they had learned in a case study than teams with lower diversity in their attitudes.

In addition to the mixed results in terms of the effect of member heterogeneity on team performance, there are research gaps in the literature regarding research methods and the lack of studies on procrastination heterogeneity. The first gap is a lack of objective observations in studies in favor of subjective questionnaires as a measurement of team performance (Baugh and Graen, 1997; van Offenbeek, 2001). Using questionnaires precludes researchers from making objective comparisons among teams' outcomes. For instance, individuals' survey-based self-evaluations of their teams' performance may inadvertently lead to self-report bias. It is thus hard to draw conclusions about whether the subjective scores retrieved using survey questions accurately represent how well teams perform. The second gap is that very few studies have investigated the effects of teams' homogeneity and heterogeneity on procrastination. Van Hooft and Van Mierlo (2018) found that team members' average levels of trait procrastination had a positive relationship with team procrastination, which in turn negatively and indirectly affected team performance mediated by team members' stress levels. However, the findings were based on a self-reported questionnaire using Lay's (1986) General Procrastination Scale and did not answer the question regarding how different combinations of individuals in terms of procrastination impact team performance.

5.1.2 Research goal

The goal of this chapter was to investigate the effect of procrastination heterogeneity on team performance. I set my research hypothesis that teams composed of members with similar levels of procrastination perform better than teams composed of members with a variety of procrastination levels. This hypothesis was developed based on previous studies indicating that conscientiousness has an inverse relationship with procrastination (Barrick et al., 1998; Kichuk and Wiesner, 1998; Lay, 1997; Lay and Brokenshire, 1997; Scher and Osterman, 2002). That is, conscientious individuals are known to procrastinate less, as they are goal-oriented, planful, and able to delay gratification (Roberts et al., 2009). The team with low conscientiousness variance among team members showed better team performance than the team with high conscientiousness variance (Humphrey et al., 2007). In a team with high conscientiousness variance, team members with high conscientiousness want to put in significant effort as they are achievement-oriented, whereas team members with low conscientiousness expend less effort on the team project (Humphrey et al., 2007). As individuals with low levels of conscientiousness are expected to perform worse (Barrick et al., 1998), highly conscientious team members consider the low-conscientious members to be free riders who put less effort into teamwork than they do individual work (Humphrey et al., 2007; LePine and Van Dyne, 2001; Williams and Karau, 1991). Such different goals for teamwork make high-conscientious individuals upset, which results in team conflict (Humphrey et al., 2007). The findings in this chapter provide guidance for forming teams considering member homogeneity/heterogeneity in procrastination.

5.2 Methodology

5.2.1 Participants

Thirty-eight students (23 males, 15 females) comprising 15 dyads or triads from a human factors in design course (Course A) and 55 students (37 males, 18 females) comprising 20 dyadic or triadic teams from a statistical quality control course (Course B) participated in this study. When I recruited participants, I tried to meet two competing goals: (1) including students from more than one course to generalize our findings and (2) recruiting students who were homogeneous in age (undergraduate students) and educational background (engineering majors) to avoid the potential influence of other sources of variation on our results (Kim et al., 2016; Kim and Nembhard, 2018, 2019b). To meet these goals, I recruited students from two different courses that had been randomly chosen from the same department. All of the participants were enrolled at the university as sophomore, junior, or senior students. At the beginning of their respective courses, participants were asked to form teams of two or three members within two weeks. Any participant who was unable to find a teammate was randomly assigned to a team by their instructor. Each participant received extra credit on their final grade for participating this study. This research was approved by the Institutional Review Board at the University of Washington. Informed consent was obtained from each participant at the beginning of the academic quarter.

5.2.2 Data collection

I collected daily page views and grades from two courses' (Courses A and B) pages hosted by the Canvas Learning Management System (LMS) through its application programming interface (API). Both courses were taught face-to-face during the 11-week spring quarter of 2019. During the academic quarter, each participant was able to access their respective course's slides, assignments, syllabus, grades, and announcements, among other content, on the corresponding course page in the Canvas LMS. Participants were required to submit all homework and project assignments online through the Canvas LMS before the assigned deadlines.

Course A covers the basic concepts of human factors and the importance of considering human capabilities and limits to understand design techniques. Course A includes (1) five individual homework assignments and (2) three team projects. The individual homework assignments in Course A used case studies that guide students to think about how human factors engineering principles can be applied to real systems. Students were required to read a case and write a one-page summary of it, including information on the accident(s), human factor reasons for the accident, and ways to design the system to avoid the accident. The average time available for participants to complete each homework assignment was approximately five days. Of these five individual assignments, four assignments were used to compute individuals' deadline reactivity. I discarded one assignment that was due fewer than seven days before from an upcoming exam because I anticipated that students' online learning behaviors for the exam would overlap with their learning behaviors for this assignment. The team projects of Course A gave students the opportunity to practice performing usability evaluations of existing products and services. The three team projects were writing a project proposal, delivering a final presentation, and writing a final report, all with the same team members.

Course B covers theoretical knowledge to address quality control problems, including statistical process control design, control charts for attributes and variables, process capability analysis, statistical tolerance design, and quality management. Course B includes (1) eight homework assignments and (2) a team project. The homework assignments in Course B were composed of several quality control questions that required knowledge of statistics and R coding. Five out of the eight homework assignments related to statistical analysis, and mathematical computations were used to determine individuals' deadline reactivity. The average time available for participants to complete each homework assignment was approximately seven days. The project assignment in Course B required the students to read a recent journal article related to statistical quality control and give a presentation summarizing their reflections while reading the article.

5.2.3 Data analysis

To calculate individual deadline reactivity, I fit individual historical page view data into the exponential model based on previous studies (Kim and Nembhard, 2018; Konig and Kleinmann, 2005). To do so, I first accrued individual page views from the Canvas course website. As the format of raw login data from the course website is the number of page views per day, I converted the number of page views each day before the deadline for each individual task. That is, I segmented the number of page views by the four individual assignments for Course A and the five individual assignments for Course B. Then, I generated historical data for each day before the deadline, and I combined the four sets of historical data for Course A and five sets of data for Course B. This method allowed us to minimize any potential individual variance in deadline reactivity during each course. Subsequently, I fit the historical data of the four assignments from Course A and of the five assignments of Course B into Equation (2.3) to yield k values. Figure 5.1 illustrates an example of the number of page views for one participant as the deadline for the homework assignment approached, and the dashed line represents the fitted exponential function with the individual k value. I used the Solver built-in in Microsoft Excel by setting the k value as a changing variable. This enabled us to minimize the MSE between the actual cumulative page view values and the cumulative probability function using Equation (2.3). The average MSE for the data fitting was 0.028 ($SD = 0.016$) for Course A and 0.026 ($SD = 0.021$) for Course B.

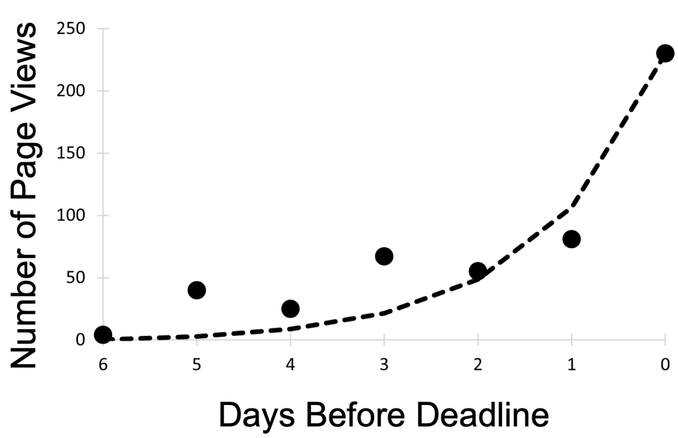

Figure 5.1: An example of one participant's page views. The dashed line represents the fitted exponential function with individual deadline reactivity.

5.2.4 Structural equation modeling

I employed a structural equation modeling (SEM) approach to investigate relationships among all the variables. SEM is advantageous for its ability to show the relationships among multiple variables simultaneously (Kline, 2015; Vinodh and Joy, 2012) and to provide direct, indirect, and mediating effects among variables (e.g., In'nami and Koizumi, 2013; Kim and Nembhard, 2019a; Macht and Nembhard, 2015; Macht et al., 2014; MacKinnon et al., 2007). Table 5.1 summarizes the variables alongside the definitions and measurements that I used to model the relationship among procrastination, team heterogeneity, and team performance. Deadline reactivity represents each participant's procrastination and was measured by the k value in Equation (2.3). Team variation measured the heterogeneity of teams by computing the standard deviation of team members' Deadline Reactivities within each team. I assumed that the teams with small variation in members' procrastination represent homogeneous teams and the teams with large variation represent heterogeneous teams. Team performance and individual performance were measured by numeric grades on participants' team assignments and individual work. The numerical grades on Team performance ranged from 0 to 100 for Course A and from 0 to 215 for Course B. The numerical grades on individual performance ranged from 0 to 4 for Course A and from 0 to 100 for Course B.

Table 5.1: Definitions and measurements for variables used in the analysis.

Variable	**Definition**	**Measurement**
Deadline reactivity	Each individual's reactivity to deadlines, representing their procrastination	k value fitted to page views per day
Team variation	Variability in team members' deadline reactivities within teams (i.e., small variation representing homogeneity and large variation representing heterogeneity)	The standard deviation of k values from team members within each team
Team performance	Academic performance on team assignments	Numerical grade on team assignments
Individual performance	Academic performance on individual assignments	Numerical grade on individual course grades

Based on the literature review, I generated a conceptual SEM without latent variables, which is shown in Figure 5.2. I used R (R Core Team, 2020) with the package 'sem' (Fox et al., 2020) to test the significance of the model. The chi-square test, root mean square error of approximation (RMSEA), Tucker-Lewis index (TLI), comparative fit index (CFI), and standardized root mean squared residual (SRMR) values were calculated. Also, the effect size of the SEM indicates the proportion of variance accounted for by each endogenous variable, i.e., team variation, team performance, and individual performance (Kraus et al., 2020; Schreiber et al., 2006). I calculated the effect size index f^2 using the equation $f^2 = R^2/(1 - R^2)$, where R^2 is the proportion of variance for each variable (Cohen, 1992).

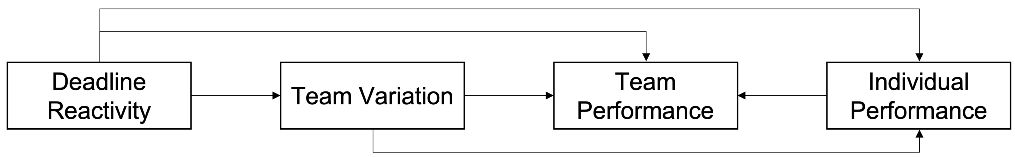

Figure 5.2: Conceptual structural equation model.

5.3 Results

5.3.1 Correlation coefficients

I calculated the correlation coefficients of all the variables to obtain a general understanding of how the variables related to each other, as shown in Table 2. In both courses, I found significant correlations between variables. For both Courses A and B, as an individual team member's deadline reactivity increases, the variability in team members' deadline reactivity increases and team performance worsens. Moreover, as the variability in team members' deadline reactivity increases, team performance worsens. Although I found consistent bivariate correlations in both Courses A and B, bivariate correlations regarding individual performance showed different relationships between Courses A and B.

Also, the results of the bivariate correlations in Table 5.2 are to detect multicollinearity for generating SEM (Grewal et al., 2004; Teo et al., 2013). Each of the correlation coefficients among the variables shown in Table 5.2 was smaller than 0.85, which indicates only a small to medium correlation among the variables (Kline, 2015). Since multicollinearity is only severe and may cause high Type-II errors when correlations among variables are high, there is minimal multicollinearity in this study.

Table 5.2: Definitions and measurements for variables used in the analysis.

Course A						
Variable	Mean	SD	95% CI	1	2	3
1. Deadline reactivity	0.38	0.18	(0.32, 0.44)			
2. Team variation	0.12	0.13	(0.08, 0.16)	0.54***		
3. Team performance	92.13	5.33	(90.38, 93.88)	-0.35*	-0.39*	
4. Individual performance	3.49	0.32	(3.39, 3.59)	-0.36*	-0.30	0.34*
Course B						
Variable	Mean	SD	95% CI	1	2	3
1. Deadline reactivity	0.42	0.16	(0.37, 0.46)			
2. Team variation	0.09	0.06	(0.07, 0.10)	0.57***		
3. Team performance	194.22	12.32	(190.89, 197.55)	-0.35**	-0.58***	
4. Individual performance	94.05	5.45	(92.58, 95.53)	0.08	0.06*	-0.08

Notes: *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

5.3.2 Structural equation model

To construct multiple relationships among the variables deadline reactivity, team variation, individual performance, and team performance, I generated empirical SEMs for Course A and for Course B, which are illustrated in Figure 5.3(a) and Figure 5.3(b), respectively. The variables from the two courses were analyzed separately due to the courses' unique content and grading criteria. I adopted this approach based on a previous study that collected page hits from four different courses and reported deadline reactivity separately for each course (Konig and Kleinmann, 2005).

As shown in Figure 5.3, both models indicated consistent relationships among the variables. Deadline reactivity positively predicted team variation, and team variation negatively predicted team performance. Team variation fully mediated the relationship between deadline reactivity and team performance. Based on the similarities between the two models, I can infer that for participants in both courses, lower deadline reactivity resulted in lower team variation, which yielded a higher team performance. In other words, less procrastinated individuals were more likely to form homogeneous teams in terms of procrastination, which resulted in better team performance. In addition to the common relationships

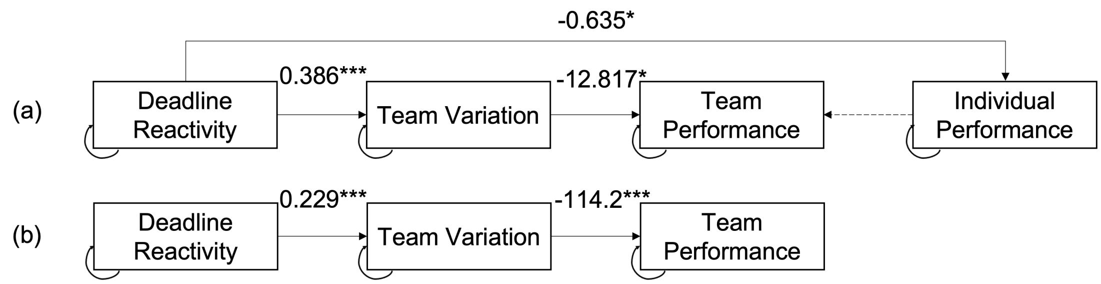

Figure 5.3: (a) Empirical SEM for Course A; (b) empirical SEM for Course B. Numeric values represent coefficients of the relationships; solid lines represent significant relationships and dotted lines represent insignificant relationships. $*p < 0.05$, $**p < 0.01$, $***p < 0.001$.

mentioned in Figure 5.3, the SEM for Course A showed a significant relationship between deadline reactivity and individual performance. Deadline reactivity negatively predicted individual performance, which means that procrastinators in Course A had lower individual grades than students who did not procrastinate.

The SEMs in Figure 5.3 were significant for both Course A ($\chi^2 = 1.354$, $p = 0.508$, $RMSEA = 0 < 0.06$, $TLI = 1.09 > 0.96$, $CFI = 1.0 > 0.96$, $SRMR = 0.05 < 0.09$) and Course B ($\chi^2 = 0.048$, $p = 0.83$, $RMSEA = 0 < 0.06$, $TLI = 1.07 > 0.96$, $CFI = 1 > 0.96$, $SRMR = 0.008 < 0.09$) using the values recommended by Hu and Bentler (1999) and Hooper et al. (2008). For the effect size of the SEM, in Course A, the SEM accounted for 28.63% of the variance of team variation ($f^2 = 0.40$), 19.23% of the variance of team performance ($f^2 = 0.24$), and 13.14% of the variance of individual performance ($f^2 = 0.15$). In Course B, the model accounted for 32.85% of the variance of team variation ($f^2 = 0.49$) and 27.92% of the variance of team performance ($f^2 = 0.39$). Cohen (1992) suggested that an f^2 of 0.35 represents a large effect size, f^2 of 0.15 represents a medium effect size, and f^2 of 0.02 represents a small effect size. Taken together, the SEMs for Courses A and B accounted for medium to large effects, for all the variables with significant relationships.

5.4 Discussions and conclusions

The results of this chapter show the mediating effect of team variation between individual deadline reactivity and team performance. My findings support our hypothesis that the teams that were homogeneous in terms of procrastination performed better than the heterogeneous teams. My findings, however, contradict the findings of previous studies that have highlighted the importance of temporal diversity within teams. For example, Mohammed and Harrison (2013) theorized that a homogeneous team composed of individuals who procrastinate less may be at risk when the task demand becomes high close to the deadline. This contradiction in findings may be due to the consistency in task demand across all the assignments in our study. I note that the results in my study were based on data gleaned under consistent five- or seven-day deadlines. Moreover, the task complexity of homework assignments was consistent throughout each course, and no changes were made in terms of assignment deadlines or task materials.

The main focus of this chapter was to model multiple and interrelated relationships between variables at once using SEM. Two common relationships between variables, i.e., between deadline reactivity and team variation and between team variation and team performance, were found via SEM for both Courses A and B, while only one relationship, i.e., between deadline reactivity and individual performance, was found for Course A. One possible reason for the single inconsistent relationship in Courses A and B is the courses' different grading policies. Course A's final grading policy indicates that individual work (such as in-class assignments, quizzes, homework assignments, and exams) constitutes 65% of the student's final grade and teamwork constitutes 35%. In contrast, a student's final grade for Course B consists of individual work for 60% and teamwork for 40%. As Course A's individual course activities constitute a greater proportion (i.e., 65%) of the final grade compared to those of Course B (i.e., 60%), students in Course A may pay more attention to individual work than students in Course B. This greater attention to individual work in Course A may result in less reactivity to deadline (i.e., smaller deadline reactivity, as summarized in Table 2) in Course A compared to Course B and a significant relationship between deadline reactivity and individual performance in Course A only.

In addition to having different final grading policies, the courses had different instructor expectations and levels of assignment difficulty, which are reflected in the different individual and team performance scores indicated in Table 2. Such different learning environments may affect the inconsistent bivariate correlations between Courses A and B. However, I note that all the inconsistent bivariate correlations between Courses A and B in Table 2, including the relationships between team performance and individual performance in Course A and team variation and individual performance in Course B, are not significant in the SEM results. One inconsistent relationship between the two courses, i.e., between deadline reactivity and individual performance, remains significant in Course A's SEM result, and a possible reason for this is described in the previous paragraph.

I followed the assumption that individual procrastination represented by deadline reactivity is consistent, as it is an individual personality trait (Schouwenburg and Lay, 1995; Van Hooft and Van Mierlo, 2018). That is, individuals' procrastination is an indication of their predisposition to put off conducting and completing tasks (Schouwenburg and Lay, 1995). Researchers have reported that trait procrastination and the Big Five personality traits are highly related (Schouwenburg and Lay, 1995). Given that the Big Five personality traits are consistent over time for adults (Cobb-Clark and Schurer, 2012), I assumed that an individual's procrastination is also consistent across the length of an academic quarter (approximately three months). This assumption is corroborated by the consistent deadline reactivity shown by the 25 individuals who participated in both course settings. There were no significant differences between these participants' deadline reactivity values for Course A and those for Course B ($p > 0.05$).

Given that the SEM results showed that homogeneous teams outperformed heterogeneous teams, I investigated whether homogeneous teams composed of low-procrastinated individuals or high-procrastinated individuals yield better performance. To answer this, I divided 35 teams (15 from Course A, 20 from Course B) into three groups—18 heterogeneous teams (eight from Course A, 10 from Course B), 12 homogeneous teams with proactive members (six from Course A, six from Course B), and five homogeneous teams with procrastinators (one from Course A, four from Course B) by applying a median split on team variation and individuals' Deadline Reactivities. If most members of a homogeneous team identified as having lower team variation had low Deadline Reactivities, I defined the team as a homogeneous team with proactive members. The rest of the homogeneous teams were homogeneous teams with procrastinators. I did not find significant differences in team performance among the three team types when conducting Kruskal-Wallis tests, which contradicts previous research showing that teams composed of proactive members perform better than teams composed of procrastinators (Van Hooft and Van Mierlo, 2018). The insignificant difference between the two types of homogeneous teams may be due to small sample sizes. There was only one team categorized as homogeneous with procrastinators in Course A; the sample sizes of other team types were likewise very small. In addition, there could be hidden variables in team composition that may affect the effects of team homogeneity in procrastination on team performance. For instance, some teams in the courses were formed by the students themselves; but some others were grouped by the instructors randomly because they were not able to find teammates by the team formation deadline. Future research can investigate the effects of the autonomy of team composition on the relationships among the individual level of procrastination, team level of procrastination, and team performance.

I remark that the results of this chapter were rooted in objective observation and quantitative modeling, making the chapter distinct from the majority of extant procrastination research that relies on self-reported questionnaires such as the Procrastination Assessment Scale-Student (PASS; Solomon and Rothblum, 1984) and the Pacing Action Categories of Effort Distribution (PACED; Gevers et al., 2015). To model individuals' procrastination more objectively, I collected participants' online activity data and fit it into the deadline rush model. Using data sets gathered from different task settings and from different courses helped us model the relationships among individual procrastination, procrastination heterogeneity, and team performance under different task types. Although the course contents were different between Course A and Course B, the significant relationships in the SEMs for both courses were consistent.

Despite such advantages, this chapter has some limitations that direct future research. First, as this project was an observational study, the class settings were uncontrollable by the authors. The uncontrollable factors included the team size of dyads and triads. Some researchers have found that team size impacts team collaboration and performance (Mao et al., 2016), while others have found that team size has little effect on team performance (Hackman and Vidmar, 1970; Van Hooft and Van Mierlo, 2018). To test whether our data can support the argument on the relationship between various team size's influence on team performance, I compared team performance between dyads and triads in each course. A Mann-Whitney U test indicated that there were no significant differences in team performance between dyads and triads in either Course A or Course B. In addition to calculating team performance, I had participants complete subjective peer evaluations of their teammates for the team projects in Course A. I noted that all students in dyads gave their teammates full scores ($Mean = 5$, $SD = 0$), while eight students in triads did not give their teammates full scores ($Mean = 4.43$, $SD = 1.13$). The Mann-Whitney U test indicated that there were significant differences in students' satisfaction levels with their teammates between the dyads and triads in Course A ($W = 592$, $p = 0.02$). Such differences suggest that some students' satisfaction with team members in triads may be lower than in dyads. Future studies might consider such factors in more detail with larger sample sizes.

The second limitation of this chapter is that I may not have captured the full range of participants' activity if they downloaded class materials all at once and worked offline or if they accessed the materials through a smartphone or tablet application. This is because I relied on participants' page views per day on their respective course websites. To handle such potential limitations, I combined the page view data from at least four assignments for both Courses A and Course B to calculate individuals' procrastination throughout the quarter instead of relying on the data from any single assignment. In addition, I followed previous studies that used students' online page views to estimate students' deadline reactivity (Kim and Nembhard, 2018; Konig and Kleinmann, 2005). Advanced software that can capture each student's specific type of work at the exact time of the day regardless of learning devices would provide more robust data.

My last limitation is the small number of participants that prevents our advanced analysis. Firstly, more participants would have been required if I had included latent variables, which is why I only focused on observable variables. Researchers have suggested that the required sample size should be at least ten times the number of variables (Gefen et al., 2000; Westland, 2010) to construct a convincing SEM. Given that our SEM includes four observable variables for Courses A and B, about 40 participants were required (ten times four variables). Future studies with larger sample sizes will allow for the inclusion of latent variables. Secondly, the results of this chapter do not show any significant gender differences for the dependent variables. Conflicting results were reported regarding whether gender is related to procrastination (Kim and Nembhard, 2019b; Özer et al., 2009) or not (Ferrari et al., 1992; Gafni and Geri, 2010). By using a larger sample size than that of the current study, future researchers could better elucidate the significance of gender differences. Thirdly, the small number of data points for some variables prevented my application of an autoregressive model which is useful for analyzing time-series data and could be used to analyze within-individual differences (MacCallum and Austin, 2000).

Chapter 6

INTEGRATED MODEL OF CAUSES AND OUTCOMES

Learning environment, such as online/in-person learning and time in an academic term, has known to contribute to the increase in students' academic procrastination and reduce in their academic performance. However, the role of academic procrastination remains unclear regarding its relationship between learning environment and academic performance. The objective of this chapter was to investigate the multivariate relationships among learning environments including online/in-person classrooms and time in academic term, academic procrastination, and academic performance simultaneously in an integrated model (**RQ I** and **RQ II**). A longitudinal field study comprised of 120 undergraduate participants was conducted from 2019 to 2022. A structural equation model (SEM) was constructed to test the relationships among variables. The results showed that in the second half of academic quarter in online learning environments, students procrastinated more and submitted assignments close to the deadline, which resulted in low academic performance. Students' academic procrastination mediated the relationship between learning environment and academic performance.

6.1 Introduction

Extensive research has concluded that academic procrastination was negatively associated with students' academic performance. Such decline in academic performance with increased procrastination has been reflected by lower grades for assignments and exams (Kim and Nembhard, 2019a; Kim and Seo, 2015; Sun and Kim, 2022b), lower course satisfaction (Elvers et al., 2003), and higher desire to dropout (Michinov et al., 2011). Procrastinators showed poor time management skills and would stack the workload to the last moment (Lay and Schouwenburg, 1993). As a result, it was harder for procrastinators to complete the assignments and prepare for the exams with good quality.

Besides academic procrastination, learning environmental has known to impact students' academic performance. For instance, students who took courses online were found to receive higher grades for exams and assignments compared to when they took in-person instructions (Clark et al., 2021; Sun and Kim, 2022b). In addition, time in academic terms was found highly associated with students' academic performance. Scheunemann et al. (2022) found that students' GPA and dropout intentions were higher in the later period of the semester than in the beginning of the semester.

Our research in Chapter 3 and 4 previously found that online learning and time in academic term impact students' academic procrastination. Students in the online learning environment showed higher tendencies to put off learning activities and submit assignments than those who attended classes in person (Sun and Kim, 2022b). Online learning demands greater self-regulation and time management skills, which were negatively correlated with academic procrastination (Hong et al., 2021; Melgaard et al., 2022). In terms of the temporal factor, students' academic procrastination was found increasing as the academic term progress in both online and in-person settings (Sun et al., 2022).

There is a lack of research integrating all the factors in a single model, investigating the role of academic procrastination in the relationship between learning environments and academic performance. Instead, there is fragmented research indicating that learning environment and procrastination affect academic performance. In this chapter, I aim to investigate the multivariate relationship among learning environment variables, academic procrastination, and academic performance.

6.2 Methodology

6.2.1 Participants

Our study included 120 participants (70 males and 50 females) from the University of Washington longitudinally from 2019 to 2022. Participants were all sophomores and juniors with a background in engineering. I recruited participants from the same course which were held in the 10-week spring quarter of 2019, 2020, 2021, and 2022.

6.2.2 Data collection

Students' learning activity data and grades were retrieved from Canvas LMS for the course from which I recruited participants. Canvas LMS integrated course announcements, assignments, and learning materials altogether. Students' number of page views on the course websites prior to the homework assignments, submission time and grades of homework assignments were recorded. Table 6.1 introduces the variables that were used in the analysis.

Table 6.1: Definitions and measurements for variables used in the analysis.

Variable		**Definition**
Learning environment		
(1)	Online/in-person classroom	Course instruction type: online learning/in-person setting.
(2)	Time in term	Time period in the academic term: 1st half/2nd half.
Academic procrastination		
(3)	Deadline reactivity	Individual's sensitivity to procrastinate on tasks with deadlines, representing individuals' procrastination.
(4)	Time proximity of submission to deadline	Time gap (minutes) between assignment submissions to assignment deadlines, representing individuals' procrastination.
Academic performance		
(5)	Task grade	Assignment grades received for the assignments used to compute (3) and (4).

Online/in-person classroom

Online/in-person classroom includes two levels, namely online learning and in-person setting. Classroom environment that was delivered completely online in 2020 and 2021 during pandemic considered as online classroom setting. There were 48 students who took the online courses. Classroom environment in 2019 and 2022 was completely in-person lectures. Seventy-two participants enrolled in the in-person courses.

Time in academic term

Time in academic term includes two levels, i.e. first half and second half of term. I split the academic terms into halves and utilized the data of assignments before week 5 for the first half of term, and data from assignments later than week 6 as the second half of term.

Deadline reactivity

Deadline reactivity was a measure of students' academic procrastination according to deadline rush model (Konig and Kleinmann, 2005). In this study, I adopted a simplified version of the deadline rush model (Equation 2.3) for computations of deadline reactivity (Sun and Kim, 2022a,b). The model assumes that individuals' perceived value toward a task V would raise as the time to deadline D declines. Individuals would procrastinate when they value the tasks less than other activities. Deadline reactivity k represents how sensitive an individual is to raise the perceived values as time progresses. In other words, the higher the deadline reactivity k, the more an individual would procrastinate on a task.

The students' daily number of page views activities data were identified as their perceived values towards the soonest assignment. The data were fit into the simplified deadline rush model and the students' deadline reactivity was obtained by minimizing the mean squared errors of data fit (Sun and Kim, 2022a,b).

Time proximity of submission to deadline

Time proximity of submission to deadline was a secondary measurement of students' academic procrastination. It was computed by finding the time differences in minutes between the students' submission time and the corresponding assignments' deadlines. Students with higher academic procrastination would submit their homework closer to the deadlines, and hence showing smaller values of time proximity of submission to deadline.

Task grade

Task grade is an indicator of students' academic performance. It is computed by the average grades in a 100% scale on the certain assignments that were previously used to compute deadline reactivity and time proximity of submission.

6.2.3 Data analysis

I constructed a correlation matrix and a Structural Equation Modeling (SEM) to examine and analyze the relationships among variables. Pearson's correlation among variables was computed by the 'Hmisc' package (Harrell Jr and Harrell Jr, 2019) in R (R Core Team, 2020). The assessment of the multivariate relationships based on multiple regression analyses was conducted through SEM by the 'sem' package (Fox et al., 2020). SEM could show the relationships among multiple variables simultaneously, discovering the direct, indirect, and mediating effects (Kline, 2015; Kim and Nembhard, 2019a; Peixoto et al., 2021). The quality of data fit to the model was examined using chi-square statistic, root mean square error of approximation (RMSEA), Tucker-Lewis index (TLI), comparative fit index (CFI), and standardized root mean squared residual (SRMR), according to the cutoff criteria recommended by Hu and Bentler (1999) and Hooper et al. (2008).

Figure 6.1 demonstrates the conceptual model showing the hypothetical relationships among variables according to the literature review. For the online/in-person classroom variable, the in-person setting is coded as '0' and online learning is coded as '1.' The time in academic term variable is coded such that '0' represents the first half of the term, and '1' represents the second half of the term. I hypothesized that procrastination (i.e., variables deadline reactivity and time proximity of submission to deadline) mediates the relationship between learning environment (i.e., variables online/in-person classroom and time in academic term) and students' academic performance (i.e., variable task grade).

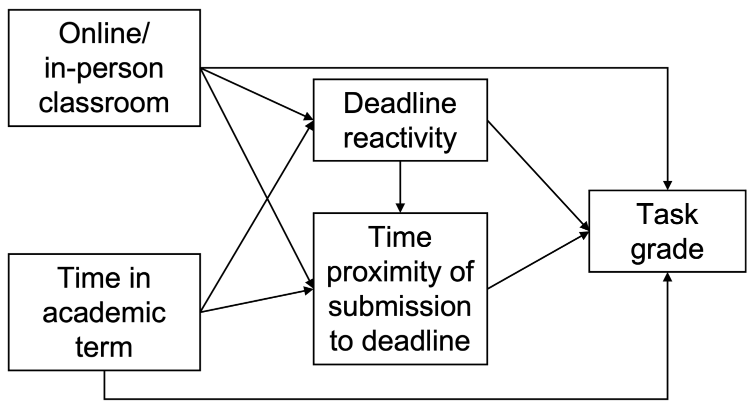

Figure 6.1: Hypothetical model of relationships among variables.

6.3 Results

6.3.1 Correlation covariance

Table 6.2 presents the descriptive statistics of the variables and the correlation coefficients among variables. Students' deadline reactivity was higher in the online learning environment than in the in-person setting. Time in academic term was positively correlated with students' deadline reactivity, and negatively correlated with time proximity of submission to deadline and task grades. Students' deadline reactivity was also negatively correlated with their time proximity of submission to deadline. Time proximity of submission to deadline was positively correlated with task grades.

Table 6.2: Descriptive statistics and correlation matrix of the variables.

	M (SD)	(1)	(2)	(3)	(4)
(1) Online/in-person classroom	-				
(2) Time in academic term	-	-			
(3) Deadline reactivity	0.532 (0.384)	0.398 ***	0.151 *		
(4) Time proximity of submission to deadline	2584.790 (2699.650)	0.117	-0.158 *	-0.149 *	
(5) Task grade	94.150 (17.416)	0.047	-0.161 *	-0.001	0.137 *

Notes: * $p < 0.05$; *** $p < 0.001$.

6.3.2 Structural equation model

Figure 6.2 illustrates significant SEM where both online/in-person classroom and time in academic term positively predicted deadline reactivity. Time proximity of submission to deadline was positively predicted by online/in-person classroom, and negatively predicted by time in academic term and deadline reactivity. Time proximity of submission to deadline positively predicted task grades.

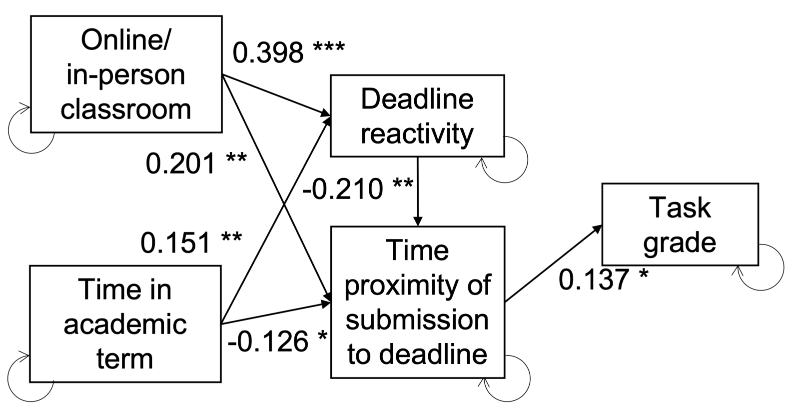

Figure 6.2: Empirical SEM of relationships among variables. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

Students' procrastination, as measured by deadline reactivity and time proximity of submission to deadline, fully mediated the relationship between online/in-person classroom and time in term, and their task grades. In other words, when studying in online environment and in the second half of the term, students would tend to procrastinate more, resulting in lower grades on the assignments. Additionally, I observed a partial mediation of students' deadline reactivity in the relationships among online/in-person classroom, time in academic term, and time proximity of submission to deadline. Following the suggestion of Hu and Bentler (1999) and Hooper et al. (2008), The model was significant with good data fit ($\chi^2 = 3.464$, $p = 0.325$, $RMSEA = 0.025 < 0.06$, $TLI = 0.977 > 0.96$, $CFI = 0.993 > 0.96$, $SRMR = 0.031 < 0.09$).

6.4 Discussions

The results of the SEM provided an understanding of the multivariate relationships among the learning environment, academic procrastination, and academic performance. Increased academic procrastination was observed when students were taking courses in an online learning environment and when they were in the second halves of the academic terms, which was in line with our preliminary research (Sun and Kim, 2022a,b; Sun et al., 2022). Our results also showed consistent findings that students received lower grades on the assignments as a consequence of increasing academic procrastination (Kim and Seo, 2015; Melgaard et al., 2022; Steel, 2007).

The academic procrastination variables fully mediated the relationship between learning environment and academic performance, showing that academic procrastination could explain how online learning and time in academic term impacted students' academic performance. The findings suggest that that by applying intervention and lessen the increase in students' academic procrastination, it is possible to mitigate the effects of online/in-person learning and time in academic term on students' grades on the assignments and exams, and further help students achieve academic success.

I noted that students' deadline reactivity predicted time proximity of submission to deadline. The later further mediated the relationship between deadline reactivity and task grade.

One of the possible reasons for the relationship was that deadline reactivity explained the students' longitudinal learning activity, while time proximity of submission to deadline explained the submission time for assignments. It was reasonable to infer that students' learning activity data carried the information of when they attempted to submit assignments. In other words, models could be constructed to predict students' submission time based on their longitudinal learning activities.

One limitation of the current study includes that the participants were collected from a single course from a cohort of population with similar backgrounds in terms of ages and majors. Similar backgrounds of participants would minimize the nuisance and unknown effects to our results, but reduce the generalizability of the findings. Future research may include datasets obtained from diverse class settings to generalize the findings and could include other qualitative measurements such as individuals' self-regulation and self-efficacy.

Chapter 7
QUANTIFICATION: PHYSIOLOGICAL MODELING

Measurement of procrastination is a critical step in research to discover how procrastination impacted individuals. However, past research frequently adopted methodologies to quantify procrastination after the task deadlines. There is a lack of models and strategies to predict procrastination while individuals are still working on the tasks. In this chapter, I conducted an exploratory study that investigates the prediction of individual procrastination from eye movement and physiological responses prior to task completion to answer the **RQ III**. Eye movement and physiological data were recorded from 20 participants during lab experiments. Participants' procrastination was computed based on their behavioral data for the tasks. I constructed and tested the prediction power of various machine learning models, including decision tree, random forest, support vector machine, and k-nearest neighbors. Random forest and support vector machine outperformed other models in predicting participants' procrastination and feature importance of the model was discussed. Our findings directed future research on modeling procrastination using combinations of eye movement and physiological responses with individual customizations.

7.1 Introductions

Objective and continuous measurements of procrastination merit accurate prediction of individual differences in students' procrastination. With an objective prediction of procrastination, researchers can accurately identify patterns of how variables affect procrastination, and how procrastination impacts individuals' performance. Objective quantification of procrastination in past research included measurement approaches such as task completion time (Elvers et al., 2003) and modeling approaches such as deadline reactivities (Sun and Kim, 2022a,b). However, these methodologies had limitations that the researchers could only determine individuals' procrastination tendencies after task completions. Therefore, there is an urge in research to explore methodologies and models to predict individual procrastination while completing tasks in real-time. For example, the recent availability and evolution of eye tracking and sensor technologies allowed researchers to measure individual differences in changes of cognitive states in the human factors field (Ahmadi et al., 2022).

Supervised machine learning methods have been frequently used in predicting individual behaviors, performance and cognitive states (Baltaci and Gokcay, 2016; McDonald et al., 2020). In a supervised learning model, both input features and output labels were provided to train the algorithm and estimate the relationships between the inputs and outputs (Zhang and Kaber, 2016). Supervised learning models were beneficial in exploring the linear and nonlinear relationships between the continuous time-series predictor and predicting variables (Kuhn et al., 2013). Therefore, well validated models could be applied to predict human behaviors and cognitive states with real-time data collected. However, limited studies exist regarding the applications of supervised machine learning approaches in predicting individuals' procrastination.

7.1.1 Eye movement and physiological responses

Few studies have directly related individuals' procrastination with their eye movement or physiological responses. Kulatilake et al. (2022) proposed a prediction model to determine individuals' procrastination indirectly through depression detection by eye-aspect ratio and facial emotions. However, the study had limitations that the authors did not provide information about how procrastination was measured, and the connection between depression detection and procrastination prediction was unclear. Despite prediction of procrastination, some other researchers found that low and high procrastination groups of individuals had significant differences in the event-related potential (ERP) on the P2 component during decision making of whether to procrastinate (Wu et al., 2016). These research showed the potential of predicting procrastination from eye movement and physiological responses.

Although limited research studied the direct correlations of eye movement and physiological responses with procrastination, there has been more literature discussing how eye movement and physiological responses data connected with anxiety (Berggren et al., 2012; Nath and Thapliyal, 2021), stress (Ancillon et al., 2022; Ren et al., 2014), and indecisiveness (Patalano et al., 2010) which are strongly associated with procrastination. Researchers found that individuals who procrastinated more tended to have more anxiety (Melgaard et al., 2022; Zarrin et al., 2020), stress (Nicholson and Scharff, 2007; Paden and Stell, 1997; Steel, 2007), and became indecisive when conducting tasks (Harriott and Ferrari, 1996; Negra et al., 2008). Since the variables associated with procrastination can be predicted by eye movement and physiological responses, the possibility of predicting procrastination with the same variables is worth exploring.

Eye tracking is a form of direct and objective measurement of individuals' eye movement patterns and visual attention (Conley et al., 2020). While processing visual information, human eyes are voluntarily fixed on the area of interest for 180-330 milliseconds (fixations) and they make quick movements between visual targets (saccades) to switch visual attention (Carter and Luke, 2020). Past research found that individuals showed different fixation patterns between low and high anxiety groups (Berggren et al., 2012), and between decisive and indecisive groups (Patalano et al., 2010). Pupil dilation is another form of eye movement that is not subject to voluntary control (Carter and Luke, 2020), which can reflect changes in individuals' stress levels (Baltaci and Gokcay, 2016; Ren et al., 2014).

Electrodermal activities (EDA) is a physiological response that measures the conductivity of skin (Shukla et al., 2019). EDA signal is composed of two components, namely the tonic and phasic components. Tonic component of EDA (skin conductance level; SCL) represents the general trend of skin conductance and changes slowly over time. Phasic component of EDA (skin conductance responses; SCR) takes place when there is a sharp increase in skin conductance. EDA is frequently used in research detecting anxiety (Nath and Thapliyal, 2021; Sebastião, 2021) and stress (Ancillon et al., 2022; Setz et al., 2009; Zontone et al., 2019) in combination with other physiological sources of data.

Blood volume pulse (BVP) measures the blood volume changes by the photoplethysmography (PPG) sensors (Val-Calvo et al., 2020). BVP signals consist of systolic peaks and notches to estimate individuals' inter-beat intervals, which can be further used to compute heart rates (HR; Ahmadi et al., 2022; Val-Calvo et al., 2020). Under conditions when human blood distribution changes, their skin temperature (TEMP) also fluctuates (Marks et al., 2009). Features from BVP, HR, and TEMP were found related to individuals' anxiety (Brandt and Dieterich, 2020; Mauriz et al., 2020) and stress (Herborn et al., 2015; Taelman et al., 2009; Yamakoshi et al., 2008). Therefore, recent studies have been actively explored to detect anxiety (Ancillon et al., 2022; Ihmig et al., 2020) and stress (Albaladejo-González et al., 2022; Sebastião, 2021) with BVP, HR, and TEMP signals.

7.1.2 Supervised learning with eye movement and physiological responses

Supervised machine learning approaches have advantages in predicting and classifying individuals' cognitive states with complex data input because they can learn from labeled training data to predict future unlabeled data. Most commonly used supervised learning methods for classifying human performance that have been applied in the human factors research include random forest (RF), support vector machine (SVM), decision tree (DT), and k-nearest neighbor (KNN), described in the rest of this section.

RF is an ensemble method built upon DT, which is a rule-based algorithm that partitions the data based on conditions of the predictor variables. RF fits multiple DTs into subsets of data and outputs the classification results based on the overall predictions of the DTs (McDonald et al., 2020; Yi et al., 2023). RF is superior to DT because it is more robust in avoiding overfitting the training data and usually yield better prediction accuracy on the unseen testing data (James et al., 2013; McDonald et al., 2020). Researchers have demonstrated that RF with physiological response input produced superior accuracy in predicting anxiety (Shaukat-Jali et al., 2021) and stress (Baltaci and Gokcay, 2016; Nath and Thapliyal, 2021; Ren et al., 2014) compared to DT and KNN.

SVM is a kernel-based algorithm that maps data features into a high-dimensional space, and separates data by a hyper-plane (Zhang and Kaber, 2016). SVM is suitable in human factors research because it is effective in training high dimensional data with non-Gaussian distributions and it is promising in resist overfitting (Zhang and Kaber, 2016). It has been implemented in anxiety detection (Ancillon et al., 2022; Ihmig et al., 2020; Shaukat-Jali et al., 2021) and stress detection (Attaran et al., 2018; Kyriakou et al., 2019; Setz et al., 2009; Sriramprakash et al., 2017; Zontone et al., 2019) with physiological response inputs producing outstanding prediction accuracy.

KNN classifies and separates data based on the Euclidean distance among data points (Sriramprakash et al., 2017). Though frequently adopted in human factors research, KNN is less robust due to its low tolerance on noises and outliers in data (Kotsiantis et al., 2007; Xie et al., 2022). Its prediction accuracy is usually lower than other approaches such as RF and SVM in detecting anxiety (Shaukat-Jali et al., 2021) and stress (Attaran et al., 2018; Ren et al., 2014; Sriramprakash et al., 2017).

Other deep learning algorithms such as artificial neural networks were also promising in detecting individuals' cognitive states such as anxiety (Ancillon et al., 2022) and stress (Albaladejo-González et al., 2022; Zontone et al., 2019). Artificial neural networks compute the possibility of data falling into each class through hidden layers. However, challenges remained in interpreting the hidden layers and explaining the relationships between inputs and outputs (Zhang et al., 2021). Due to its low interpretability, I did not consider artificial neural networks in the present study.

7.1.3 Research goal

In this chapter, I aim to propose a model that predicts individuals' deadline reactivity through eye movement, BVP, EDA, HR, and TEMP data (**RQ7.1**).To the best of the authors' knowledge, there has been no studies so far directly predicting individuals' procrastination with eye movement and physiological responses. With a classification model that successfully predicts deadline reactivity, researchers can gain more insights about which eye movement, BVP, EDA, HR, and TEMP features were important in predicting procrastination (**RQ7.2**). In addition, some researchers found that RF and KNN models using only EDA features yielded better prediction accuracy than the same models using the combination of eye movement and EDA (Ren et al., 2014), whereas others found that combinations of both eye movement and TEMP data improved the prediction accuracy for both DT and RF models (Baltaci and Gokcay, 2016). Inclusion of features that are not correlated with the predicting variables may harm the prediction accuracy (Sriramprakash et al., 2017). Therefore, I propose the **RQ7.3** to explore whether features from a single physiological source is sufficient in predicting deadline reactivity accurately.

RQ7.1: Can eye movement, BVP, EDA, HR, and TEMP predict individuals' deadline reactivity?

RQ7.2: Which features of physiological responses and eye movement are relatively more important for predicting deadline reactivity?

RQ7.3: Is the multimodal model using the combination of eye movement, BVP, EDA, HR, and TEMP better in prediction of deadline reactivity than models using features from a single sensor?

7.2 Methodologies

7.2.1 Participants

I recruited 20 participants for the present study through flyers and emails at the University of Washington. Table 7.1 shows the demographics of the participants. A screening questionnaire was filled prior to the experiment to determine eligibility to participate. I only recruited participants who were English as first language speakers so that they had sufficient vocabulary to complete the tasks. The participants had no training on completing the tasks prior to attending the experiments and had normal or corrected to normal eye sight. Participants received gift card compensation of a $10 base plus bonus according to the completion and correction of tasks during the experiments.

Table 7.1: Demographics of participants.

Number of participants	20
Age	
Mean	21.40
Standard deviation	5.71
Gender No. (%)	
Female	14 (70%)
Male	6 (30%)
Major No. (%)	
Arts and Science	5 (25%)
Engineering	12 (60%)
Others	3 (15%)
Ethnicity No. (%)	
Asian	9 (45%)
White and Caucasian	10 (50%)
Others	1 (5%)

7.2.2 Apparatus

The experiment was conducted on a desktop with a 27" monitor (1920 x 1080 pixel). Participants interacted with the task interface through a wired mouse and typed their responses with a wired keyboard to ensure the reliability of interactions. During the experiments, I used two non-invasive devices to acquire participants' physiological responses. A Tobii Pro X3-120 eye tracker (120 Hz) was attached below the monitor to track participants' eye movement on screen. The participants also wore an Empatica E4 wristband on their nondominant hand to record their blood volume pressure (BVP; 64 Hz), electrodermal activity (EDA; 4 Hz), heart rate (HR; 1 Hz), and skin temperature (TEMP; 4 Hz). Researchers monitored the data collected from the E4 wristband through the E4 real-time app on an iPad during the experiment to ensure the data quality.

7.2.3 Experiment procedures

The tasks that the participants completed were anagram tasks. For each anagram, participants needed to rearrange a list of random ordered letters to form them into an English word.

Anagram tasks were previously used by other researchers to identify individual differences in procrastination in limited experiment time (Nicholson and Scharff, 2007). I selected 64 5-letter anagrams that had single solutions from (Gilhooly and Johnson, 1978). The 64 anagrams were split into eight sessions, each containing eight anagrams. I purposely set the overall difficulty in solving the anagrams in the eight sessions to be consistent by controlling the GTZERO scores of the anagrams within the sessions ($M = 37.86$, $SD = 0.22$). GTZERO score was an index computed according to the positions and orders of the letters in the anagram that could serve as an measurement of the anagram difficulties (Gilhooly and Johnson, 1978). An anagram with a higher GTZERO score was considered more difficult to solve.

Upon participants' arrival, the researchers instructed the participants about the study goal and the experiment procedures. The participants had the opportunity to get familiar with the anagram tasks and the task interface for a 5-minute practice trial before the main experiments. Figure 7.1(a) demonstrates the task interface for the experiment. To eliminate the nuisance order effects of anagrams and sessions, I completely randomized the orders of the anagrams within and between sessions. To complete the task, participants needed to type the answers in the textbox next to each anagram, and click on the 'submit' button whenever they were confident with their answers. After submission, the participants could not make changes to the submitted answers anymore. The top right corner of the interface showed the current session number and the remaining time for completing the current session. For each session, participants were instructed to complete and submit the anagrams within a 10-minute deadline. I purposely set up the session such that participants could not move onto the next session before the 10 minutes were over, even if they completed all the anagrams before the deadline, so the participants could not rush to the next session during the experiment. Participants had a one-minute rest session between every two experiment sessions. I calibrated the eye tracker prior to the first experiment session , and during each of the rest sessions. Figure 7.1(b) demonstrates the flow of the experiment procedures. The experiment took approximately 1.5 hours to complete. The task interface recorded the time remaining in the session when participants clicked on the ‘submit’ button for each anagram and their submitted answers.

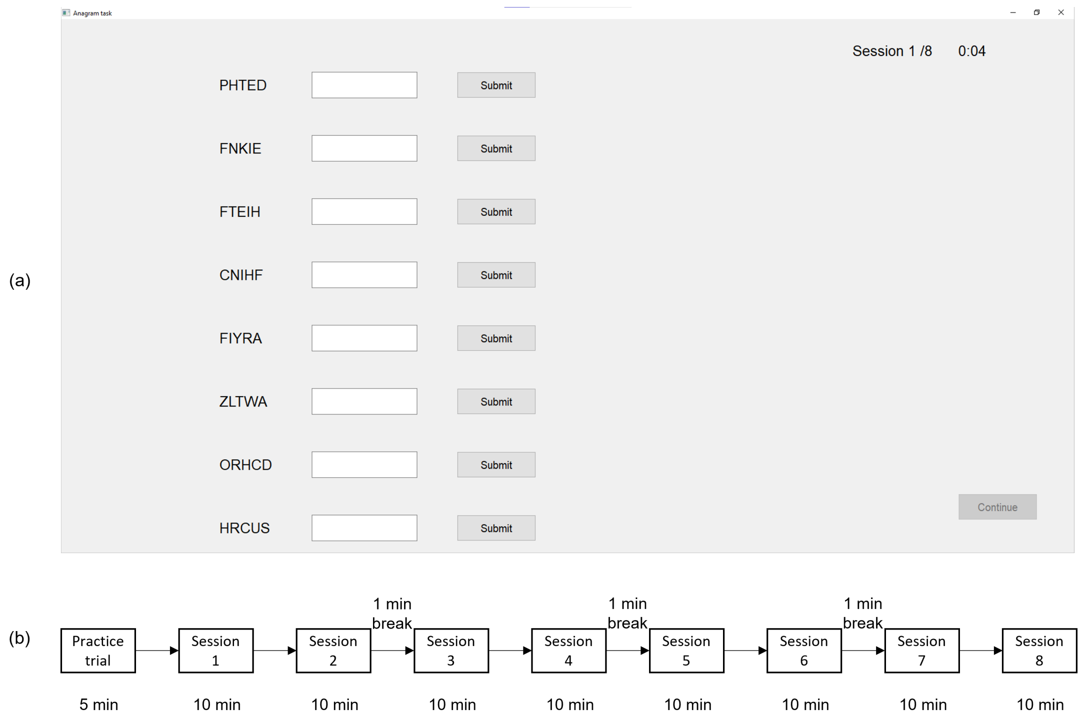

Figure 7.1: (a) Screenshot of task interface. (b) Experiment procedures.

7.2.4 Data analysis

Figure 7.2 demonstrates the schematic of the data analysis pipeline of this study. Since individuals’ procrastination cannot be directly observed, I used the deadline rush model to identify the individual differences of task procrastination through the participants’ behavioral data during the experiments as the ground truth labels for the supervised learning models. Raw eye movement data and physiological data collected from the E4 wristband were preprocessed and subjected to feature extraction. The combined features were inputted into the supervised learning models for model evaluation and comparison. Lastly, I computed the feature importance for the best performing model.

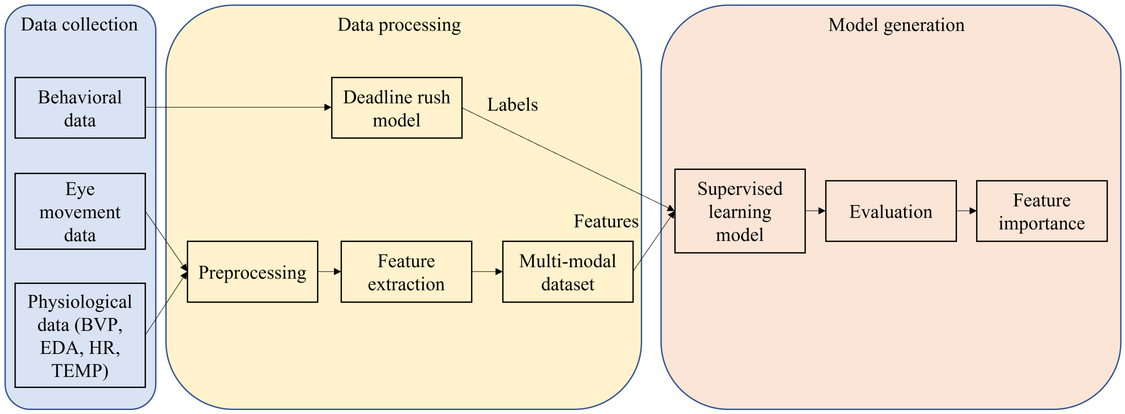

Figure 7.2: Schematic of data generation and modeling.

Behavioral data: deadline reactivity

Participants' procrastination while finishing the anagram tasks during the experiments was measured by their deadline reactivity (Equation 2.3) based on deadline rush model (Konig and Kleinmann, 2005). I assumed that participants' perceived values towards each anagram was linearly associated with the anagrams' difficulty levels measured by GTZERO scores. That is, a more difficult anagram accounted for greater values in completing the tasks before the deadline. Since the average difficulty of the anagrams in all the sessions were similar, I set the sum of participants' perceived values A towards the anagrams in a session as a constant 1. I fit the cumulative perceived values of the anagrams completed V and the time of each anagram submission D in the deadline rush model to retrieve a deadline reactivity value k for each participant at each session while minimizing the mean square errors. Figure 7.3 demonstrates the sample data fit for one of the participants. I retrieved eight deadline reactivity values for each participant based on their submission behaviors in the eight experiment sessions. As this project was an exploratory study for classification prediction, I labeled the participants' deadline reactivity data in each session as either high or low through a median split.

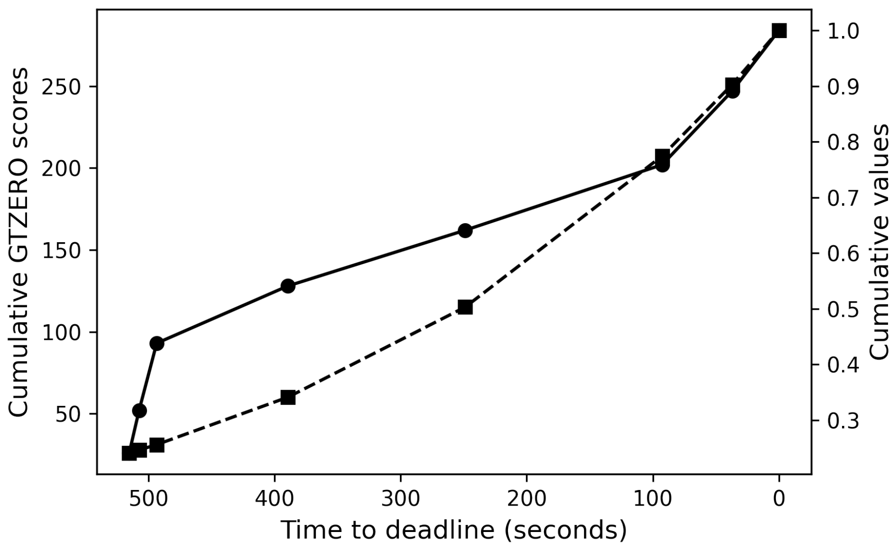

Figure 7.3: Example plot of data fit of deadline rush model to obtain deadline reactivity for a participant. Circles and the solid line represent the cumulative GTZERO scores of the anagrams they submitted versus the anagram submission time (left y-axis). Squares and the dashed line represent the fitted cumulative perceived values versus the anagram submission time (right y-axis).

Eye movements and physiological data: Data preprocessing

Both the eye movement data and the physiological data collected by E4 wristband were analyzed by the experiment sessions. Since the goal of the study was to find the relationship between physiological responses and participants' procrastination while conducting tasks, their behaviors and physiological data after they completed all the anagrams were not of our interests. Therefore, I cut the data at the time that participants submitted the last anagram in the session and used only the data prior to the last submission for analysis. For those who were not able to complete all anagrams before the deadline, the data for the entire session was used.

The raw eye tracking data was processed by the Tobii Pro Lab software to label the fixation and saccade instances. For the data collected from the E4 wristband, I first removed the baseline information from the BVP, EDA, HR, and TEMP data since huge individual differences could be observed from the magnitudes of physiological responses, and there could be session variations in the recorded physiological responses (Wu et al., 2021; Zhang et al., 2020) . I considered the participants' physiological responses collected during the one-minute eye movement calibration as a baseline level for the following two experiment sessions before the next rest session. The magnitude of participants' BVP, EDA, HR, and TEMP data were subtracted by the average of the corresponding physiological data collected during the latest eye calibration (Smyth et al., 2021).

After baseline removal, I applied filters to smooth the physiological data and remove noise artefacts. I applied a lowpass Butterworth filter with an order of 6 and cutoff frequencies of 0.5 Hz, 0.2 Hz, and 0.3 Hz for the BVP, EDA, and TEMP data respectively (Zhu et al., 2020). Since the HR data recorded were computed according to BVP by the E4 wristband and was sampled at 1 Hz, I did not apply filters on HR data.

Eye movements and physiological data: Feature extraction

As eye movement, BVP, EDA, HR, and TEMP data are time-series data, multiple approaches could be applied to input the data into the supervised learning models. The simplest approach is to directly set the time-series data as the input data (Dhaouadi and Khelifa, 2020). However, I purposely cut the eye movement and physiological data by the task completion time of the participants. The length of time series data for all sessions were inconsistent. Therefore, I generated a few time domain features from the data to characterize the individual differences in physiological changes while conducting the tasks (Zhu et al., 2020). Table 7.2 shows the list of features extracted from the eye movement, BVP, EDA, HR, and TEMP data prepared for model input.

For the eye movement data, I computed the total count, the average duration, and frequency of the eye fixations and saccades. I also computed the average pupil dilation for both left and right eyes. Entropy of the eye movement was calculated based on the eye scan path patterns across the screen regions of various elements of the task interface.

Table 7.2: Features extracted from the eye movement, BVP, EDA, HR, and TEMP data.

Eye movement	**BVP, EDA, HR, TEMP**
Total number of fixations	Time domain features:
Mean duration of fixations	Min
Frequency of fixations (number of fixations / task completion time)	Max
Total number of saccades	Mean
Mean duration of saccades	Standard deviation (SD)
Frequency of saccades (number of saccades / task completion time)	Root mean square (RMS)
Mean pupil dilation of left eye	Means of the absolute values of the first difference
Mean pupil dilation of right eye	Means of the absolute values of the second difference
Entropy	Additional EDA features:
	Mean amplitude of SCR
	Frequency of SCR (number of SCR / task completion time)

For BVP, EDA, HR, and TEMP data, I computed the time domain features, i.e. minimum, maximum, mean, standard deviation, root mean square, and means of the absolute values of the first difference and of the second difference (Zhu et al., 2020). For EDA only, I also extracted the SCR by a continuous decomposition analysis with an amplitude threshold of 0.1 muS using Ledalab V3.4.9 (Benedek and Kaernbach, 2010) in Matlab (The MathWorks Inc., 2022). The average amplitude of the SCR and the frequency of SCR were computed.

From the preprocessed eye movement, BVP, EDA, HR, and TEMP data, I collected a total of 39 features for model input: 9 (eye movement) + 7 (BVP) + 9 (EDA) + 7 (HR) + 7 (TEMP).

Model development and evaluation

I input the extracted features from eye movement, BVP, EDA, HR, and TEMP data into four supervised learning classification models, namely RF, SVM, DT, and KNN. Since the scale of the data values for different features varied a lot, I applied a z-score normalization for all the features across the participants to standardize the values. All the supervised learning models were constructed through the 'scikit-learn' package (Pedregosa et al., 2011) in Python. Table 7.3 illustrates the hyperparameters used for each model.

Table 7.3: Hyperparameter values for RF, SVM, DT, and KNN models.

Model	**Hyperparameter**	**Values**
RF	Number of trees	50
	Split criterion	Gini index
	Max tree depth	Split until all leaves are pure
SVM	Kernel	RBF kernel
DT	Split criterion	Gini index
	Max tree depth	Split until all leaves are pure
KNN	Number of neighbors	5

I used the precision, recall, balanced accuracy, F1 score, and area under curve (AUC) of the receiver operating characteristic (ROC) curve as the metrics to evaluate the model performance (Grandini et al., 2020; Rosset, 2004). The evaluation metrics are computed based on the ratios of the four possible outcomes of classification prediction, i.e. true positive (TP), true negative (TN), false positive (FP), and false negative (FN). Table 7.4 explains each of the evaluation metrics and the computation methods. All the evaluation metrics were ranged between zero and one, and higher values (closer to one) indicated better model performance. I utilized a 10-fold cross validation for the model validation. The dataset was cut randomly into 10 folds. In each iteration within the cross validation, nine folds of the dataset were used as the training data and the remaining fold was tested for computations of the evaluation metrics. The same 10-fold assignment was applied to train and test for all the models. I repeated the training and validating process for 50 times on 50 distinct random seeds (Yi et al., 2023). The average and standard deviation of the evaluation metrics were calculated for model performance comparisons.

Table 7.4: Evaluation metrics of machine learning models (Grandini et al., 2020; Rosset, 2004).

Evaluation metric	Definitions	Computations
Precision	Proportion of positive predicted instances that are actually positive. It tells how much the model can be trusted if an instance is predicted as positive.	$\frac{TP}{TP+FP}$
Recall	Proportion of positive predicted instances out of the positive class. It tells the model's ability to correctly predict the positive instances.	$\frac{TP}{TP+FN}$
Balanced accuracy	Average of recalls for both positive and negative classes. It computes how likely on average a single instance is predicted correctly for both positive and negative classes.	$\frac{1}{2} * (\frac{TP}{TP+FN} + \frac{TN}{FP+TN})$
F1 score	Harmonic average between precision and recall. It can spot the weakness in the prediction model if either precision or recall is extremely lower than the other.	$2 * \frac{Precision*Recall}{Precision+Recall}$
AUC-ROC curve	It shows how close the ROC curve is to the upper left corner of the coordinates. An ROC curve shows the relationship between TP rate and FP rate at various thresholds. AUC-ROC curve is a robust measurement to evaluate model performance in distinguishing between positive and negative instances.	Area under curve for the ROC curve.

7.3 Results

7.3.1 Descriptive statistics

Table 7.5 illustrates the descriptive statistics for the participants' deadline reactivity and the average of each sensor data collected during the experiments. The BVP, EDA, HR, and TEMP data included negative values because the baseline signal was subtracted.

Table 7.5: Descriptive statistics for deadline reactivity, eye movement, and physiological data with baseline removed.

	Mean	**SD**	**95% confidence interval**
Deadline reactivity	2.706	8.454	(1.385, 4.026)
Number of fixations (No.)	823.194	37.863	(739.213, 907.174)
Number of saccades (No.)	2104	1496.863	(1870.284, 2337.716)
Mean BVP (nW)	-0.057	0.509	(-0.137, 0.022)
Mean EDA (μS)	-0.144	1.450	(-0.371, 0.082)
Mean HR (No./second)	-2.041	10.335	(-3.655, -0.427)
Mean TEMP ($°C$)	0.751	1.193	(0.565, 0.938)

7.3.2 Model performance comparison

Table 3 shows the average and standard deviation of all the evaluation metrics for each of the supervised learning classification models during the 50 iterations of cross validations. Kruskal Wallis (K-W) rank sum test was conducted to compare each of the evaluation metrics among the models. All the evaluation metrics were significantly different among the models. I then conducted post hoc Dunn's tests with Benjamini-Hochberg adjustments. As shown in Figure 7.4 and Table 7.6, RF and SVM both had significantly better performance than DT and KNN in terms of all the five metrics. SVM had significantly higher precision than RF, whereas RF had significantly higher recall than SVM. However, the differences between RF and SVM in terms of balanced accuracy, F1 score, and AUC-ROC curve were insignificant. In addition, Figure 7.5 shows the mean ROC curves of DT, RF, SVM, and KNN models. Similar to what is shown by the Kruskal Wallis test, the ROC curves for RF and SVM were both above the ROC curves for DT and KNN at almost all the thresholds. The differences between ROC curves of SVM and RF models were barely visible as the two curves overlapped at most of the thresholds. Comprehensively considering the results of the five evaluation metrics, I selected both the RF and SVM models for further analysis.

Table 7.6: The average (standard deviation) of the evaluation metrics for RF, SVM, DT and KNN and their comparisons.

Model	**Precision**	**Recall**	**Balanced accuracy**	**F1 score**	**AUC-ROC curve**
RF	0.862 (0.016)	0.861 (0.025)	0.861 (0.015)	0.852 (0.017)	0.926 (0.012)
SVM	0.879 (0.014)	0.826 (0.013)	0.857 (0.009)	0.843 (0.011)	0.933 (0.009)
DT	0.788 (0.030)	0.789 (0.033)	0.787 (0.022)	0.776 (0.024)	0.789 (0.022)
KNN	0.786 (0.018)	0.813 (0.025)	0.796 (0.014)	0.787 (0.019)	0.869 (0.015)
K-W test	$\chi^2 = 153.01$, $p < 0.001$	$\chi^2 = 107.97$, $p < 0.001$	$\chi^2 = 150.55$, $p < 0.001$	$\chi^2 = 150.80$, $p < 0.001$	$\chi^2 = 170.24$, $p < 0.001$

7.3.3 Feature importance

I used permutation feature importance (Altmann et al., 2010; Qi, 2012) to compute the importance scores for all features in both the RF and SVM models. The permutation importance score for a feature is computed based on the average decrease in the model performance score when the values of the feature are randomly shuffled (Qi, 2012). A feature with a low importance score that is close to zero means that the feature is nearly independent when predicting the outcomes. Figure 7.6 demonstrates the top ten important features in the RF and SVM models averaged from the 50 iterations of 10-fold cross validations. For both RF and SVM models, the eye movement features such as the total number of fixations and saccades had the highest importance scores among all other features. Fixation rate, saccade rate, and standard deviation of EDA were also shared in both models as two of the most important features. In the RF model, features from TEMP (i.e., standard deviation of TEMP), EDA (i.e., standard deviation of EDA, and SCR frequency), and BVP (i.e., minimum BVP, and root mean square of BVP) had higher importance scores; whereas in the SVM model, features from HR (i.e., root mean square of HR) were more important in predicting deadline reactivity.

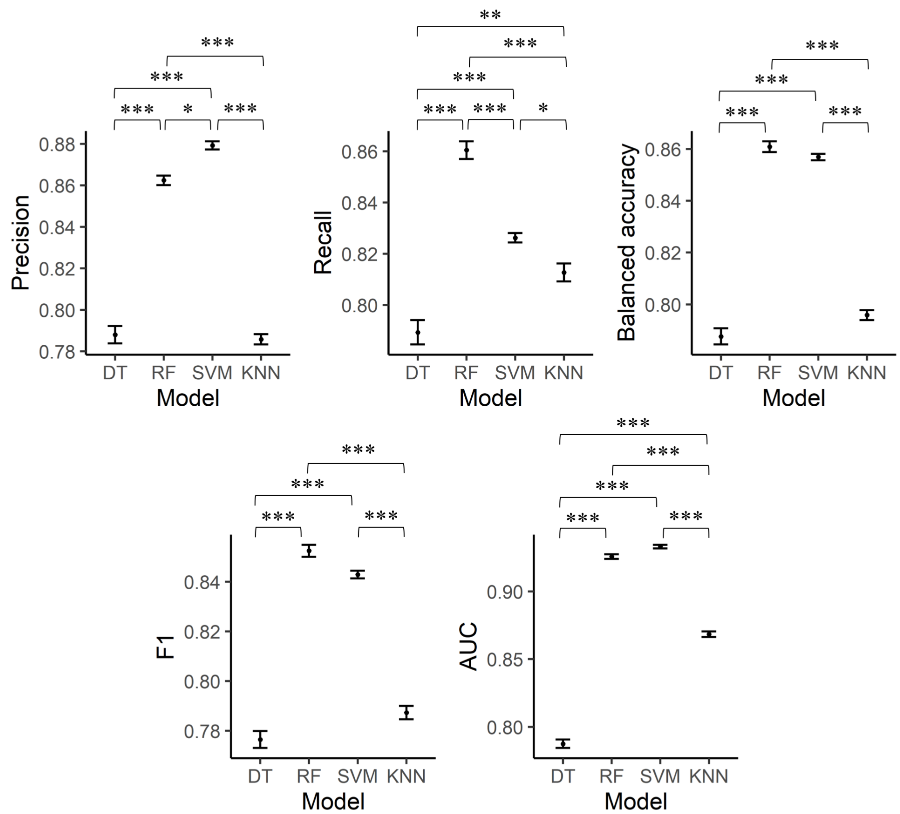

Figure 7.4: Pairwise comparisons of precision (top left), recall (top middle), balanced accuracy (top right), F1 (bottom left), and AUC-ROC curve (bottom right) among DT, RF, SVM, and KNN. Error bars represent the standard errors. Notes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

7.3.4 Multisensor vs. unisensor model performance

Since multiple eye movement features were ranked as the most important features in the RF and SVM models, I further evaluated the models' performance with various combinations of feature input. I compared the RF and SVM models that were trained upon (1) only the BVP, EDA, HR, and TEMP features collected from the E4 wristband, (2) only the eye movement features collected from the eye tracker, and (3) the combination of all features as shown in Chapter 7.3.1. The same training and evaluation approaches described

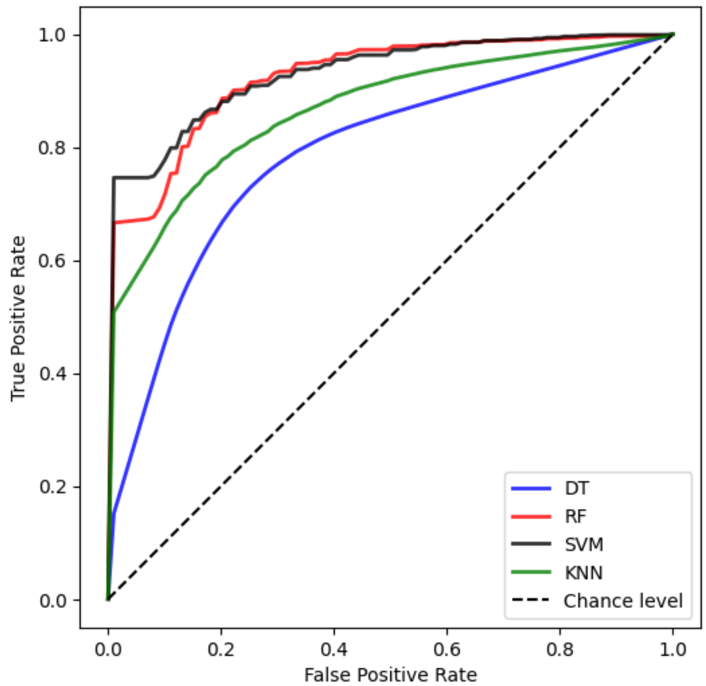

Figure 7.5: Graphical comparison of ROC curves for the DT, RF, SVM, and KNN models.

above were conducted on the models. Figure 7.7 shows the comparisons of model performance with the three sets of features and the detailed test statistics are displayed in Table 7.7. The results from RF models were similar to those from SVM models. Models trained solely by the eye movement features and models trained by the combination of all features showed better performance than the models with only the physiological features collected from the E4 wristband ($p < 0.05$ for all pairwise comparisons). In comparison between the two better performed models, the precision, balanced accuracy, and AUC-ROC curve of the model with all features were significantly higher than those of the model with only the eye movement features ($p < 0.05$ for all pairwise comparisons).

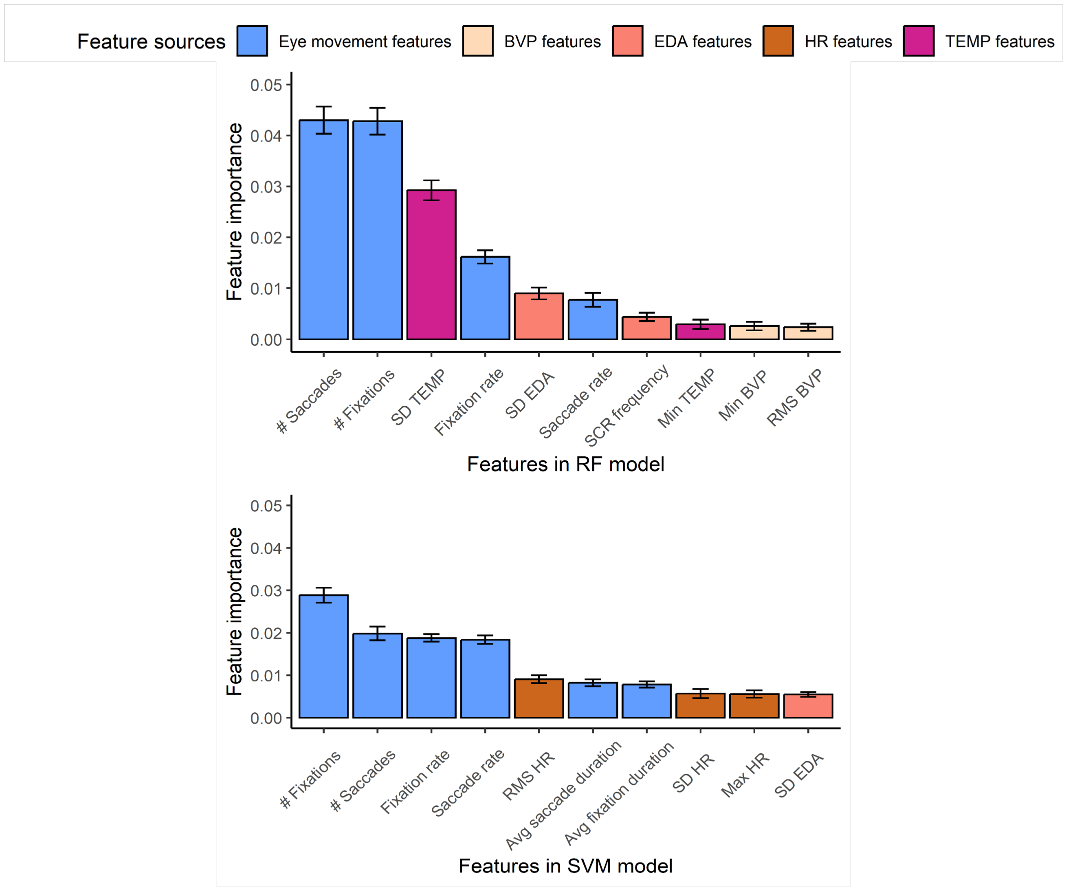

Figure 7.6: Top ten important features in the RF model (top) and in the SVM model (bottom). Features from the same sensor were coded in the same color.

7.4 Discussions

The present study explores the prediction of individuals' procrastination through eye movement and physiological responses including BVP, EDA, HR, and TEMP. I developed a controlled experiment to analyze participants' procrastination on task completion in a time sensitive environment. Our findings showed that supervised machine learning methods were effective in integrating both eye movement and physiological responses and accurately predict deadline reactivity towards tasks with deadlines.

The AUC-ROC curve of RF, SVM, DT, and KNN were all higher than 0.75, showing

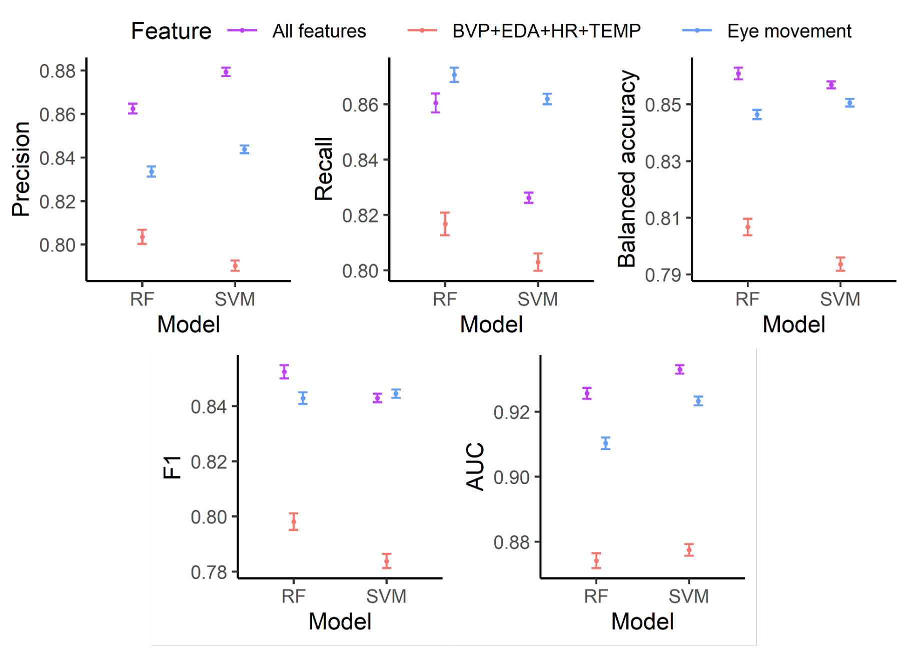

Figure 7.7: Comparisons of evaluation metrics between RF and SVM models using physiological features, eye movement features, and the combination of all features. Error bars represent the standard errors.

acceptable model performance as compared to the chance level (Mandrekar, 2010). Among the four models constructed, RF and SVM both showed better prediction performance than DT and KNN according to all the evaluation metrics. DT is less complex than RF and SVM and therefore may be more prone to overfitting (McDonald et al., 2020). However, DT is beneficial for its high interpretability by the non-machine learning experts (Kotsiantis et al., 2007; McDonald et al., 2020). Given that the DT model still yielded acceptable performance in predicting deadline reactivity, researchers could utilize DT as a tool to gain more insights of how eye movement and physiological features correlated with deadline reactivity. In line with other research predicting anxiety and stress, KNN was weaker in prediction than RF and SVM using features from various sensor data (Attaran et al., 2018; Ren et al., 2014; Shaukat-Jali et al., 2021; Sriramprakash et al., 2017). It is probably due to the nature of the KNN algorithm that it is considered intolerant to noises and outliers in data (Kotsiantis

Table 7.7: Test statistics of Kruskal Wallis tests for each of the model evaluation metrics in the RF and SVM models.

Model	Precision	Recall	Balanced accuracy	F1 score	AUC-ROC curve
RF	$\chi^2 = 103.130$, $p < 0.001$	$\chi^2 = 72.775$, $p < 0.001$	$\chi^2 = 99.644$, $p < 0.001$	$\chi^2 = 95.068$, $p < 0.001$	$\chi^2 = 104.990$, $p < 0.001$
SVM	$\chi^2 = 129.110$, $p < 0.001$	$\chi^2 = 107.980$, $p < 0.001$	$\chi^2 = 104.160$, $p < 0.001$	$\chi^2 = 98.855$, $p < 0.001$	$\chi^2 = 109.300$, $p < 0.001$

et al., 2007; Xie et al., 2022). However, huge individual differences could be observed in eye movement and physiological responses.

Interestingly, the balanced accuracy, F1 score, and AUC-ROC curve of RF and SVM in predicting deadline reactivity were similar to each other. However, I found that SVM had higher precision and lower recall than RF. That is, RF was more likely to identify individuals with high procrastination, while the high procrastination individuals classified by SVM had higher credibility. Practitioners could consider either RF or SVM models according to their expectations on detecting the high procrastination individuals and their tolerance to false alarms.

Our findings showed that eye movement, BVP, EDA, HR, and TEMP sensor data were capable of predicting deadline reactivity in a time sensitive environment. Although there has been little research connecting procrastination with eye movement and physiological responses directly, numerous studies have concluded that individuals' procrastination was highly correlated with their anxiety and stress (Melgaard et al., 2022; Paden and Stell, 1997; Steel, 2007; Zarrin et al., 2020), which could be effectively predicted by various combinations of eye movement and physiological responses (Attaran et al., 2018; Baltaci and Gokcay, 2016; Ihmig et al., 2020; Nath and Thapliyal, 2021; Shaukat-Jali et al., 2021; Kyriakou et al., 2019; Zontone et al., 2019). Among the eye movement, BVP, EDA, HR, and TEMP sensor data, eye movement features were dominating in the most important features to predict deadline reactivity in both RF and SVM models. For instance, the number of fixations and saccades were the top two important features in both models. Spearman's correlation showed that both number of fixations and saccades were positively correlated with the deadline reactivity class ($\rho = 0.514$, $p = 0.020$; and $\rho = 0.597$, $p = 0.006$, respectively). Intuitively, as an individual procrastinated in completing tasks, they tended to complete the anagrams closer to the deadlines, resulting in longer task completion time. Therefore, more fixations and saccades would be observed. In addition, individuals who procrastinate on completing tasks may have difficulties in the decision-making process about which task to start working on and how to proceed with the tasks (Negra et al., 2008). Researchers previously found that indecisive individuals showed different patterns of eye movements from the decisive individuals, specifically in the gaze transitions among areas of interest (Patalano et al., 2010). The indecisive participants may show more frequent gaze activities on the screen among the anagrams while deciding on which anagram to solve and whether to submit their answers to the anagrams. Therefore, the eye movement features, such as the number and rate of fixations and saccades, were found important in the prediction models of individual deadline reactivity.

The multimodal model evaluation highlights the importance of considering both eye movement and physiological responses in predicting procrastination, which is consistent with past research in stress detection (Baltaci and Gokcay, 2016; Nath and Thapliyal, 2021). Features from multisensor data may supplement information that was neglected by data from a single sensor signal. Although some researchers concluded that prediction models with subsets of features may perform better than considering all the features (Ren et al., 2014; Shaukat-Jali et al., 2021), a possible reason was that some particular physiological responses were not promising in prediction of the target predicting variables. For instance, RF and SVM models put higher weights on different combinations of physiological responses when predicting deadline reactivity. Since HR and BVP features were not listed in the top 10 important features in the RF and SVM models respectively, the exclusions of the unimportant sensors may reduce the complexity in data collection and model generation for researchers, and possibly further enhance the model performance. Future research could be conducted to explore the model performance with various combinations of the physiological responses input.

I note that the present study is an exploratory study for predicting deadline reactivity in real-time. Our findings showed the potential to identify individual differences in procrastination while completing tasks under deadlines with multisensor physiological data collected simultaneously. Practitioners may use the prediction models as a tool to identify procrastinators and nonprocrastinators prior to task completion. Intervention strategies to reduce procrastination may be applied on the procrastinators to assist them with time management and enhance task performance.

I acknowledge that the present study has a few limitations. I selected 39 features from the eye movement, BVP, EDA, HR, and TEMP physiological data. However, several features were highly correlated with each other, such as number of fixations and number of saccades ($\rho = 0.930$, $p < 0.001$), which may cause multicollinearity issues in the model generation. Multicollinearity may probably impact the performance and generalizability of the models (McDonald et al., 2020). Future research may consider various feature selection methods, such as clustering (Sriramprakash et al., 2017) and principal component feature selection (Lufimpu-Luviya et al., 2013; McDonald et al., 2020), to improve the model performance with subsets of the eye movement and physiological data. Secondly, the present study classified individuals' procrastination based on the median split of their deadline reactivity. The binary classification hinders the information of individual differences in procrastination that can be reflected by the continuous deadline reactivity values. Future research may explore regression models to predict the continuous deadline reactivity. With a regression prediction model of individuals' procrastination, practitioners can provide customized intervention strategies to reduce procrastination. Lastly, the importance of the eye movement features may be found if the scan path of switching attention among various task elements can be observed. For instance, the findings can be applied on the prediction of students' procrastination for remote learning because their learning activities were primarily on-screen using electronic devices (Clark et al., 2021). However, the findings may not be generalized to task scenarios when visual components of tasks are limited. Future research can be conducted to generalize the prediction model on tasks in diverse settings.

Chapter 8
GENERAL DISCUSSIONS AND CONCLUSIONS

In this dissertation, I modeled students' procrastination in the university classrooms setting to develop strategies for reducing their procrastination and improving academic performance. To achieve the research goals, I conducted a longitudinal field study from 2019 to 2022 at the University of Washington and investigated the learning environmental factors influencing students' procrastination (**RQ I**) as well as the effects of procrastination on their team performance (**RQ II**). I also conducted a controlled experiment and proposed a supervised machine learning approach to predict procrastination prior to task completion (**RQ III**). The chapters and their corresponding key summaries are displayed in Table 8.1. The rest of this chapter summarizes the overall findings and the contributions, and describes the limitations and future directions of research.

8.1 Review of findings

8.1.1 Causes of procrastination

The objective of the research was to investigate the learning environmental factors that could influence students' procrastination in university classrooms (Chapter 2, 3, and 6). Unlike other studies that heavily relied on subjective ratings to measure procrastination, I adopted a deadline rush model to obtain an objective measurement of students' procrastination using their longitudinal learning activities. Through a field study, I modeled students' procrastination by their interaction activities on the Canvas LMS in two undergraduate courses. Nonparametric mixed-design ANOVA was adopted to detect the relationships between students' procrastination and online/in-person classrooms, task complexity, and time in the academic term. The results showed that students showed higher procrastination

Table 8.1: Chapters and corresponding key summary

Chapter	RQ addressed	Key summary
3. Causes: online learning and task complexity	RQ I	Students procrastinate more in the online learning and when conducting high-complexity tasks.
4. Causes: online learning and time in the academic term	RQ I	Students procrastinate more in the online learning and second half of an academic term.
5. Outcomes: individual and team performance	RQ II	Teams composed of members with homogeneous procrastination perform better than those with heterogeneous procrastination.
6. Integrated model of causes and outcomes	RQ I, II	Procrastination mediates the relationship between learning environments and performance.
7. Quantification: physiological modeling	RQ III	Eye movement and physiological responses are capable of predicting procrastination with supervised learning approaches.

when studying in online learning classrooms, working on high-complexity tasks, and during the 2nd half of the academic term.

8.1.2 Outcomes of procrastination

The study aimed to investigate how team heterogeneity in procrastination impacted team performance (Chapter 4). I constructed SEMs to show the relationships among students' individual deadline reactivity, team variation in deadline reactivity, team performance, and individual performance. The empirical SEMs from both courses indicated that team variation in deadline reactivity mediated the effects of students' individual deadline reactivity on their team performance. Homogeneous teams composed of members with similar procrastination tendencies received better grades for their team assignments.

I further generated an SEM integrating the causes and outcomes of students' procrastination (Chapter 6). The model indicated that students' procrastination fully mediated the relationships between their task grades and both online/in-person classrooms and time in the academic term. Reducing students' procrastination could mitigate the negative effects of online/in-person learning and time in the academic term on students' academic performance, and further help students succeed in university courses.

8.1.3 Quantification of procrastination

I proposed a prediction methodology to determine students' procrastination prior to task completion using eye movement and physiological responses including BVP, EDA, HR, and TEMP (Chapter 7). I first designed a controlled experiment and collected students' eye movement, BVP, EDA, HR, and TEMP during anagram tasks in a time sensitive environment. Students' procrastination was modeled by their task completion during the experiments and labeled as either “high" or “low" by a median split. RF and SVM models were constructed and showed promising performance in predicting procrastination through features from eye movement and physiological responses. Although eye movement features were the most important features in both RF and SVM models, the multimodal RF and SVM models using the combination of features from all the sensors outperformed the models with only features from either eye movement or physiological sensors.

8.2 Contributions

The results of the dissertation contribute to the fields of engineering and education by providing insights of individual differences in procrastination in the university classrooms. The findings of **RQ I** (Chapter 3, 4, 6) showed that despite individual differences and task characteristics, learning environmental factors, such as online/in-person learning and time in the academic term, also affected students' procrastination. Considerations of the interaction effects of learning environments and other factors are necessary for researchers to draw more robust conclusions about individual differences in procrastination. Instructors and online learning platforms could be inspired by the findings to develop instructional designs and strategies to intervene and mitigate students' procrastination for courses in various learning environments.

The findings for **RQ II** (Chapter 4, 6) highlighted the importance of considering individual procrastination in team composition. When designing team projects, instructors and management team leaders can purposely form teams according to the members' homogeneity in procrastination for better team performance.

The prediction models of procrastination answering **RQ III** (Chapter 7) showed the potential of using continuous eye movement and physiological responses to determine levels of procrastination prior to task completion. The prediction models identified the features that contributed to predicting procrastination under a time sensitive environment. Researchers can modify the experiment design to determine effective intervention strategies for procrastination. Educators and online learning platforms can apply the proposed approaches to gather information on students' and workers' procrastination in real-time and provide customized interventions before poor performance is yielded.

The short-term impacts of the dissertation can benefit the general education society. Online learning platforms and remote teaching programs can apply the dissertation's findings to their products' designs to improve their users' satisfaction and learning outcomes, and hence improve the user engagement on the products. The dissertation's implications can be applied to the development of education and educators for science, technology, engineering, and math (STEM) and non-STEM students and assist the students to achieve academic success. The research scope can also be extended to a more diverse population, including students in elementary, middle, and high schools.

In long-term, the findings hold the potential to be incorporated into developments of training programs for workers in diverse systems. By modeling workers' learning and working behaviors, the organizations can gain insights of individual differences in work pace, providing personalized learning and training programs for workers to advance their performance while minimizing the risks of physical and mental illness.

8.3 Limitations and Future Research

There are limitations to this dissertation which guide future research directions. First, all the participants in my research were undergraduate and graduate students at the University of Washington and most of the participants had a homogeneous engineering background. The similarity in participants' backgrounds reduced the nuisance effects of uncontrolled factors on the findings, but compromised the generalizability of the results. Future research could be conducted on students with diverse backgrounds, i.e., students in different majors, from different universities, and at different age groups.

Second, individual differences in personality and characteristics may cause students to procrastinate (Ferrari et al., 1992; Dewitte and Schouwenburg, 2002; Schouwenburg and Lay, 1995). The present study had insufficient consideration of these factors when analyzing the learning environmental causes and outcomes of students' procrastination. In future research, questionnaires of students' personality, stress, and anxiety levels can be distributed during the academic terms to investigate the interaction effects of individual differences and learning environment factors on students' procrastination.

Finally, the exploratory model developed directed future research for more accurate, reliable, and real-time prediction of procrastination. The proposed model used eye movement and physiological features throughout the tasks prior to completion. Improvements to the feature generation and selection could be applied to the model for predicting procrastination in real-time. For instance, future research can split the time series physiological data into equal-width time windows and train the model with features generated from each time window. In addition, in future research, regression models can be explored in addition to the classification models to predict individual procrastination. The regression models can be applied to gain insights about individual differences in procrastination for customized intervention.

BIBLIOGRAPHY

Ackerman, D. S. and Gross, B. L. (2005). My instructor made me do it: Task characteristics of procrastination. *Journal of Marketing education*, 27(1):5–13.

Ackerman, D. S. and Gross, B. L. (2007). I can start that jme manuscript next week, can't i? the task characteristics behind why faculty procrastinate. *Journal of Marketing education*, 29(2):97–110.

Adnan, M. and Anwar, K. (2020). Online learning amid the covid-19 pandemic: Students' perspectives. *Online Submission*, 2(1):45–51.

Afrashteh, M. Y. and Koohneshin Seighalani, M. (2021). Comparing the effect of virtual and face-to-face education on learning motivation and procrastination in elementary school students during the coronavirus pandemic (covid-19). *Iranian Evolutionary and Educational Psychology Journal*, 3(3):225–233.

Ahmadi, N., Sasangohar, F., Nisar, T., Danesh, V., Larsen, E., Sultana, I., and Bosetti, R. (2022). Quantifying occupational stress in intensive care unit nurses: An applied naturalistic study of correlations among stress, heart rate, electrodermal activity, and skin temperature. *Human Factors*, 64(1):159–172.

Aitken, M. E. (1982). *A personality profile of the college student procrastinator*. University of Pittsburgh.

Akram, A., Fu, C., Li, Y., Javed, M. Y., Lin, R., Jiang, Y., and Tang, Y. (2019). Predicting students' academic procrastination in blended learning course using homework submission data. *IEEE Access*, 7:102487–102498.

Albaladejo-González, M., Ruipérez-Valiente, J. A., and Gómez Mármol, F. (2022). Evaluating different configurations of machine learning models and their transfer learning capabilities for stress detection using heart rate. *Journal of Ambient Intelligence and Humanized Computing*, pages 1–11.

Altmann, A., Toloşi, L., Sander, O., and Lengauer, T. (2010). Permutation importance: a corrected feature importance measure. *Bioinformatics*, 26(10):1340–1347.

Ancillon, L., Elgendi, M., and Menon, C. (2022). Machine learning for anxiety detection using biosignals: a review. *Diagnostics*, 12(8):1794.

APA (2021). Stress and decision-making during the pandemic. `https://www.apa.org/news/press/releases/stress/2021/october-decision-making`. Accessed: 2023-05-25.

Attaran, N., Puranik, A., Brooks, J., and Mohsenin, T. (2018). Embedded low-power processor for personalized stress detection. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 65(12):2032–2036.

Baltaci, S. and Gokcay, D. (2016). Stress detection in human–computer interaction: Fusion of pupil dilation and facial temperature features. *International Journal of Human–Computer Interaction*, 32(12):956–966.

Barrick, M. R., Stewart, G. L., Neubert, M. J., and Mount, M. K. (1998). Relating member ability and personality to work-team processes and team effectiveness. *Journal of applied psychology*, 83(3):377.

Baugh, S. G. and Graen, G. B. (1997). Effects of team gender and racial composition on perceptions of team performance in cross-functional teams. *Group & Organization Management*, 22(3):366–383.

Benedek, M. and Kaernbach, C. (2010). Decomposition of skin conductance data by means of nonnegative deconvolution. *psychophysiology*, 47(4):647–658.

Berggren, N., Koster, E. H., and Derakshan, N. (2012). The effect of cognitive load in emotional attention and trait anxiety: An eye movement study. *Journal of cognitive psychology*, 24(1):79–91.

Bonner, S. E. (1994). A model of the effects of audit task complexity. *Accounting, organizations and society*, 19(3):213–234.

Bowers, C. A., Pharmer, J. A., and Salas, E. (2000). When member homogeneity is needed in work teams: A meta-analysis. *Small group research*, 31(3):305–327.

Boyle, G. J. (1995). Myers-briggs type indicator (mbti): some psychometric limitations. *Australian Psychologist*, 30(1):71–74.

Brandt, T. and Dieterich, M. (2020). 'excess anxiety'and 'less anxiety': both depend on vestibular function. *Current Opinion in Neurology*, 33(1):136–141.

Bryan, J. F. and Locke, E. A. (1967). Parkinson's law as a goal-setting phenomenon. *Organizational Behavior and Human Performance*, 2(3):258–275.

Campbell, D. J. and Gingrich, K. F. (1986). The interactive effects of task complexity and participation on task performance: A field experiment. *Organizational behavior and human decision processes*, 38(2):162–180.

Carter, B. T. and Luke, S. G. (2020). Best practices in eye tracking research. *International Journal of Psychophysiology*, 155:49–62.

Cho, M. (2018). Task complexity, modality, and working memory in l2 task performance. *System*, 72:85–98.

Chu, A. H. C. and Choi, J. N. (2005). Rethinking procrastination: Positive effects of" active" procrastination behavior on attitudes and performance. *The Journal of social psychology*, 145(3):245–264.

Clark, A. E., Nong, H., Zhu, H., and Zhu, R. (2021). Compensating for academic loss: Online learning and student performance during the covid-19 pandemic. *China Economic Review*, 68:101629.

Clement, D. E. and Schiereck, J. J. (1973). Sex composition and group performance in a visual signal detection task. *Memory & cognition*, 1(3):251–255.

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

Coe, C. K. (1992). The mbti: Potential uses and misuses in personnel administration. *Public Personnel Management*, 21(4):511–522.

Cohen, J. (1992). A power primer. *Psychological bulletin*, 112(1):155.

Conley, Q., Earnshaw, Y., and McWatters, G. (2020). Examining course layouts in blackboard: Using eye-tracking to evaluate usability in a learning management system. *International Journal of Human–Computer Interaction*, 36(4):373–385.

Darby, A., Longmire-Avital, B., Chenault, J., and Haglund, M. (2013). Students' motivation in academic service-learning over the course of the semester. *College Student Journal*, 47(1):185–191.

Dewitte, S. and Schouwenburg, H. C. (2002). Procrastination, temptations, and incentives: The struggle between the present and the future in procrastinators and the punctual. *European Journal of personality*, 16(6):469–489.

Dhaouadi, S. and Khelifa, M. M. B. (2020). A multimodal physiological-based stress recognition: Deep learning models' evaluation in gamers' monitoring application. In *2020 5th International Conference on Advanced Technologies for Signal and Image Processing (ATSIP)*, pages 1–6. IEEE.

Dhawan, S. (2020). Online learning: A panacea in the time of covid-19 crisis. *Journal of Educational Technology Systems*, 49(1):5–22.

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

Fernie, B. A., Bharucha, Z., Nikčević, A. V., Marino, C., and Spada, M. M. (2017). A metacognitive model of procrastination. *Journal of affective disorders*, 210:196–203.

Ferrari, J. R., Parker, J. T., and Ware, C. B. (1992). Academic procrastination: Personality correlates with myers-briggs types, self-efficacy, and academic locus of control. *Journal of social Behavior and personality*, 7(3):495–502.

Fincham, F. and May, R. (2021). My stress led me to procrastinate: Temporal relations between perceived stress and academic procrastination. *College Student Journal*, 55(4):413–421.

Fox, J., Nie, Z., and Byrnes, J. (2020). *sem: Structural Equation Models*. R package version 3.1-11.

Funke, J. (1991). Solving complex problems: Exploration and control of complex systems. *Complex problem solving: Principles and mechanisms*, 4:185–222.

Gafni, R. and Geri, N. (2010). Time management: Procrastination tendency in individual and collaborative tasks. *Interdisciplinary Journal of Information, Knowledge, and Management*, 5(1):15–125.

Gefen, D., Straub, D., and Boudreau, M.-C. (2000). Structural equation modeling and regression: Guidelines for research practice. *Communications of the association for information systems*, 4(1):7.

Gevers, J., Mohammed, S., and Baytalskaya, N. (2015). The conceptualisation and measurement of pacing styles. *Applied Psychology*, 64(3):499–540.

Gevers, J., Rispens, S., and Li, J. (2016). Pacing style diversity and team collaboration: The moderating effects of temporal familiarity and action planning. *Group Dynamics: Theory, Research, and Practice*, 20(2):78.

Gilhooly, K. and Johnson, C. (1978). Effects of solution word attributes on anagram difficulty: A regression analysis. *Quarterly Journal of Experimental Psychology*, 30(1):57–70.

Gonda, D., Pavlovičová, G., Tirpáková, A., and Ďuriš, V. (2021). Setting up a flipped classroom design to reduce student academic procrastination. *Sustainability*, 13(15):8668.

Grandini, M., Bagli, E., and Visani, G. (2020). Metrics for multi-class classification: an overview. *arXiv preprint arXiv:2008.05756*.

Grewal, R., Cote, J. A., and Baumgartner, H. (2004). Multicollinearity and measurement error in structural equation models: Implications for theory testing. *Marketing science*, 23(4):519–529.

Grunschel, C. and Schopenhauer, L. (2015). Why are students (not) motivated to change academic procrastination?: an investigation based on the transtheoretical model of change. *Journal of College Student Development*, 56(2):187–200.

Hackman, J. R. and Vidmar, N. (1970). Effects of size and task type on group performance and member reactions. *Sociometry*, pages 37–54.

Häfner, A., Oberst, V., and Stock, A. (2014). Avoiding procrastination through time management: An experimental intervention study. *Educational Studies*, 40(3):352–360.

Harrell Jr, F. E. and Harrell Jr, M. F. E. (2019). Package 'hmisc'. *CRAN2018*, 2019:235–236.

Harriott, J. and Ferrari, J. R. (1996). Prevalence of procrastination among samples of adults. *Psychological reports*, 78(2):611–616.

Herborn, K. A., Graves, J. L., Jerem, P., Evans, N. P., Nager, R., McCafferty, D. J., and McKeegan, D. E. (2015). Skin temperature reveals the intensity of acute stress. *Physiology & behavior*, 152:225–230.

Higgs, M., Plewnia, U., and Ploch, J. (2005). Influence of team composition and task complexity on team performance. *Team Performance Management: An International Journal*.

Hodges, C. B., Moore, S., Lockee, B. B., Trust, T., and Bond, M. A. (2020). The difference between emergency remote teaching and online learning.

Hong, J.-C., Lee, Y.-F., and Ye, J.-H. (2021). Procrastination predicts online self-regulated learning and online learning ineffectiveness during the coronavirus lockdown. *Personality and individual differences*, 174:110673.

Hooper, D., Couglan, J., and Mullen, M. R. (2008). The effects of group composition on achievement, interaction, and learning efficiency during computer-based cooperative instruction. *Electronic Journal of Business Research Methods*, 6(1):53–60.

Hooper, S. and Hannafin, M. J. (1991). The effects of group composition on achievement, interaction, and learning efficiency during computer-based cooperative instruction. *Educational Technology Research and Development*, 39(3):27–40.

Hu, L.-t. and Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural equation modeling: a multidisciplinary journal*, 6(1):1–55.

Huang, N., Zhang, J., Burtch, G., Li, X., and Chen, P. (2021). Combating procrastination on massive online open courses via optimal calls to action. *Information Systems Research*, 32(2):301–317.

Humphrey, S. E., Hollenbeck, J. R., Meyer, C. J., and Ilgen, D. R. (2007). Trait configurations in self-managed teams: A conceptual examination of the use of seeding for maximizing and minimizing trait variance in teams. *Journal of Applied Psychology*, 92(3):885.

Ihmig, F. R., Neurohr-Parakenings, F., Schäfer, S. K., Lass-Hennemann, J., and Michael, T. (2020). On-line anxiety level detection from biosignals: Machine learning based on a randomized controlled trial with spider-fearful individuals. *Plos one*, 15(6):e0231517.

In'nami, Y. and Koizumi, R. (2013). Structural equation modeling in educational research: A primer. *Application of Structural Equation Modeling in Educational Research and Practice*, 23(51):23.

James, G., Witten, D., Hastie, T., and Tibshirani, R. (2013). *An introduction to statistical learning*, volume 112. Springer.

Janssen, T. and Carton, J. S. (1999). The effects of locus of control and task difficulty on procrastination. *The Journal of Genetic Psychology*, 160(4):436–442.

Johnson, N., Veletsianos, G., and Seaman, J. (2020). Us faculty and administrators' experiences and approaches in the early weeks of the covid-19 pandemic. *Online Learning*, 24(2):6–21.

Kachgal, M. M., Hansen, L. S., and Nutter, K. J. (2001). Academic procrastination prevention/intervention: Strategies and recommendations. *Journal of Developmental Education*, 25(1):14.

Kichuk, S. L. and Wiesner, W. H. (1998). Work teams: Selecting members for optimal performance. *Canadian Psychology/Psychologie canadienne*, 39(1-2):23.

Kim, J.-E. and Nembhard, D. A. (2018). Parametric empirical bayes estimation of individual time-pressure reactivity. *International Journal of Production Research*, 56(7):2452–2463.

Kim, J.-E. and Nembhard, D. A. (2019a). Eye movement as a mediator of the relationships among time pressure, feedback, and learning performance. *International Journal of Industrial Ergonomics*, 70:116–123.

Kim, J.-E. and Nembhard, D. A. (2019b). The impact of procrastination on engineering students' academic performance. *International Journal of Engineering Education*, 35(4):1008.

Kim, J.-E., Nembhard, D. A., and Kim, J. H. (2016). The effects of group size and task complexity on deadline reactivity. *International Journal of Industrial Ergonomics*, 56:106–114.

Kim, K. R. and Seo, E. H. (2015). The relationship between procrastination and academic performance: A meta-analysis. *Personality and Individual Differences*, 82:26–33.

Kline, R. B. (2015). *Principles and practice of structural equation modeling*. Guilford publications.

Knaus, W. J. (2000). Procrastination, blame, and change. *Journal of social Behavior and Personality*, 15(5):153.

Konig, C. J. and Kleinmann, M. (2005). Deadline rush: a time management phenomenon and its mathematical description relationships between critical thinking and attitudes toward women's roles in society. *The Journal of psychology*, 139(1):33–45.

Kotsiantis, S. B., Zaharakis, I., Pintelas, P., et al. (2007). Supervised machine learning: A review of classification techniques. *Emerging artificial intelligence applications in computer engineering*, 160(1):3–24.

Kramer, A., Bhave, D. P., and Johnson, T. D. (2014). Personality and group performance: The importance of personality composition and work tasks. *Personality and Individual Differences*, 58:132–137.

Kraus, J., Scholz, D., and Baumann, M. (2020). What's driving me? exploration and validation of a hierarchical personality model for trust in automated driving. *Human factors*, page 0018720820922653.

Kristof-Brown, A. L. and Stevens, C. K. (2001). Goal congruence in project teams: Does the fit between members' personal mastery and performance goals matter? *Journal of applied psychology*, 86(6):1083.

Kuhn, M., Johnson, K., et al. (2013). *Applied predictive modeling*, volume 26. Springer.

Kulatilake, T., Liyanage, P., Deemud, G., De Silva, U., Sriyaratna, D., and Kugathasan, A. (2022). Prodep: Smart social media procrastination and depression tracker. In *2022 17th International Workshop on Semantic and Social Media Adaptation & Personalization (SMAP)*, pages 1–6. IEEE.

Kyriakou, K., Resch, B., Sagl, G., Petutschnig, A., Werner, C., Niederseer, D., Liedlgruber, M., Wilhelm, F., Osborne, T., and Pykett, J. (2019). Detecting moments of stress from measurements of wearable physiological sensors. *Sensors*, 19(17):3805.

Lay, C. and Schouwenburg, H. (1993). Trait procrastination, time management, and academic behavior. *Journal of Social Behavior and Personality*, 8(4):647.

Lay, C. H. (1986). At last, my research article on procrastination. *Journal of research in personality*, 20(4):474–495.

Lay, C. H. (1997). Explaining lower-order traits through higher-order factors: the case of trait procrastination, conscientiousness, and the specificity dilemma. *European Journal of Personality*, 11(4):267–278.

Lay, C. H. and Brokenshire, R. (1997). Conscientiousness, procrastination, and person-task characteristics in job searching by unemployed adults. *Current Psychology*, 16(1):83–96.

LePine, J. A. and Van Dyne, L. (2001). Voice and cooperative behavior as contrasting forms of contextual performance: evidence of differential relationships with big five personality characteristics and cognitive ability. *Journal of applied psychology*, 86(2):326.

Levy, Y., Ramim, M. M., et al. (2013). An experimental study of habit and time incentive in online-exam procrastination. In *Proceedings of the chais conference on instructional technologies research*, pages 53–61.

Li, L., Gao, H., and Xu, Y. (2020). The mediating and buffering effect of academic self-efficacy on the relationship between smartphone addiction and academic procrastination. *Computers & Education*, 159:104001.

Liu, P. and Li, Z. (2012). Task complexity: A review and conceptualization framework. *International Journal of Industrial Ergonomics*, 42(6):553–568.

Lufimpu-Luviya, Y., Merad, D., Paris, S., Drai-Zerbib, V., Baccino, T., and Fertil, B. (2013). A regression-based method for the prediction of the indecisiveness degree through eye movement patterns. In *Proceedings of the 2013 Conference on Eye Tracking South Africa*, pages 32–38.

MacCallum, R. C. and Austin, J. T. (2000). Applications of structural equation modeling in psychological research. *Annual review of psychology*, 51(1):201–226.

Macht, G. A. and Nembhard, D. A. (2015). Measures and models of personality and their effects on communication and team performance. *International Journal of Industrial Ergonomics*, 49:78–89.

Macht, G. A., Nembhard, D. A., Kim, J. H., and Rothrock, L. (2014). Structural models of extraversion, communication, and team performance. *International Journal of Industrial Ergonomics*, 44(1):82–91.

MacKinnon, D. P., Fairchild, A. J., and Fritz, M. S. (2007). Mediation analysis. *Annu. Rev. Psychol.*, 58:593–614.

Mandrekar, J. N. (2010). Receiver operating characteristic curve in diagnostic test assessment. *Journal of Thoracic Oncology*, 5(9):1315–1316.

Mao, A., Mason, W., Suri, S., and Watts, D. J. (2016). An experimental study of team size and performance on a complex task. *PloS one*, 11(4):e0153048.

Marks, A., Vianna, D. M., and Carrive, P. (2009). Nonshivering thermogenesis without interscapular brown adipose tissue involvement during conditioned fear in the rat. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, 296(4):R1239–R1247.

Martinez, S.-K., Meier, S., and Sprenger, C. (2022). Procrastination in the field: Evidence from tax filing. *Journal of the European Economic Association*, page jvac067.

Mauriz, E., Caloca-Amber, S., and Vázquez-Casares, A. M. (2020). Effect of facial skin temperature on the perception of anxiety: a pilot study. In *Healthcare*, volume 8, page 206. MDPI.

McDonald, A. D., Ade, N., and Peres, S. C. (2020). Predicting procedure step performance from operator and text features: A critical first step toward machine learning-driven procedure design. *Human Factors*, page 0018720820958588.

Melgaard, J., Monir, R., Lasrado, L. A., and Fagerstrøm, A. (2022). Academic procrastination and online learning during the covid-19 pandemic. *Procedia computer science*, 196:117–124.

Mello, A. S. and Ruckes, M. E. (2006). Team composition. *The journal of Business*, 79(3):1019–1039.

Michie, S. (2002). Causes and management of stress at work. *Occupational and environmental medicine*, 59(1):67–72.

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

Mohammed, S. and Angell, L. C. (2003). Personality heterogeneity in teams: Which differences make a difference for team performance? *Small group research*, 34(6):651–677.

Mohammed, S. and Harrison, D. A. (2013). The clocks that time us are not the same: A theory of temporal diversity, task characteristics, and performance in teams. *Organizational behavior and human decision processes*, 122(2):244–256.

Mohammed, S. and Nadkarni, S. (2011). Temporal diversity and team performance: The moderating role of team temporal leadership. *Academy of Management Journal*, 54(3):489–508.

Moon, S. M. and Illingworth, A. J. (2005). Exploring the dynamic nature of procrastination: A latent growth curve analysis of academic procrastination. *Personality and Individual differences*, 38(2):297–309.

Muilenburg, L. Y. and Berge, Z. L. (2005). Student barriers to online learning: A factor analytic study. *Distance education*, 26(1):29–48.

Nath, R. K. and Thapliyal, H. (2021). Smart wristband-based stress detection framework for older adults with cortisol as stress biomarker. *IEEE Transactions on Consumer Electronics*, 67(1):30–39.

Negra, A., Mzoughi, N., and Bouhlel, O. (2008). E-procrastination: a netnographic approach. *Journal of Customer Behaviour*, 7(2):103–119.

Nguyen, B., Steel, P., and Ferrari, J. R. (2013). Procrastination's impact in the workplace and the workplace's impact on procrastination. *International Journal of Selection and Assessment*, 21(4):388–399.

Nicholson, L. and Scharff, L. F. (2007). The effects of procrastination and self-awareness on emotional responses. *Psi Chi Journal of Undergraduate Research*, 12(4).

Noguchi, K., Gel, Y. R., Brunner, E., and Konietschke, F. (2012). nparld: an r software package for the nonparametric analysis of longitudinal data in factorial experiments. *Journal of Statistical software*, 50(12).

OSHA (2022). Workplace stress. `https://www.osha.gov/workplace-stress#ref5`. Accessed: 2023-05-25.

Owens, A. M. and Newbegin, I. (1997). Procrastination in high school achievement: A causal structural model. *Journal of Social Behavior and Personality*, 12(4):869.

Owens, S. G., Bowman, C. G., and Dill, C. A. (2008). Overcoming procrastination: the effect of implementation intentions. *Journal of Applied Social Psychology*, 38(2):366–384.

Özer, B. U., Demir, A., and Ferrari, J. R. (2009). Exploring academic procrastination among turkish students: Possible gender differences in prevalence and reasons. *The Journal of social psychology*, 149(2):241–257.

Özer, B. U. and Sarıcaoğlu, H. (2014). An experimental study on academic procrastination: Effects of weekly homework. *Psycho-Educational Research Reviews*, 3(1):1–7.

Paden, N. and Stell, R. (1997). Reducing procrastination through assignment and course design. *Marketing Education Review*, 7(2):17–25.

Parkinson, C. N. and Osborn, R. C. (1957). *Parkinson's law, and other studies in administration*, volume 24. Houghton Mifflin Boston.

Patalano, A. L., Juhasz, B. J., and Dicke, J. (2010). The relationship between indecisiveness and eye movement patterns in a decision making informational search task. *Journal of Behavioral Decision Making*, 23(4):353–368.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830.

Peixoto, E. M., Pallini, A. C., Vallerand, R. J., Rahimi, S., and Silva, M. V. (2021). The role of passion for studies on academic procrastination and mental health during the covid-19 pandemic. *Social Psychology of Education*, 24(3):877–893.

Pychyl, T. A., Lee, J. M., Thibodeau, R., and Blunt, A. (2000). Five days of emotion: An experience sampling study of undergraduate student procrastination. *Journal of social Behavior and personality*, 15(5):239.

Qi, Y. (2012). Random forest for bioinformatics. In *Ensemble machine learning: Methods and applications*, pages 307–323. Springer.

R Core Team (2020). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Rapanta, C., Botturi, L., Goodyear, P., Guàrdia, L., and Koole, M. (2020). Online university teaching during and after the covid-19 crisis: Refocusing teacher presence and learning activity. *Postdigital Science and Education*, 2(3):923–945.

Ren, P., Barreto, A., Huang, J., Gao, Y., Ortega, F. R., and Adjouadi, M. (2014). Off-line and on-line stress detection through processing of the pupil diameter signal. *Annals of biomedical engineering*, 42:162–176.

Robb, H. (1998). Reducing procrastination in the workplace. *Handbook of organizational health psychology: Programs to make the workplace healthier*, pages 55–77.

Roberts, B. W., Jackson, J. J., Fayard, J. V., Edmonds, G., and Meints, J. (2009). Conscientiousness. *Handbook of individual differences in social behavior*, pages 369–381.

Rose, D., Seidler, A., Nübling, M., Latza, U., Brähler, E., Klein, E., Wiltink, J., Michal, M., Nickels, S., Wild, P., et al. (2017). Associations of fatigue to work-related stress, mental and physical health in an employed community sample. *BMC psychiatry*, 17(1):1–8.

Rosset, S. (2004). Model selection via the auc. In *Proceedings of the twenty-first international conference on Machine learning*, page 89.

Ruiz-Gallardo, J.-R., Castaño, S., Gómez-Alday, J. J., and Valdés, A. (2011). Assessing student workload in problem based learning: Relationships among teaching method, student workload and achievement. a case study in natural sciences. *Teaching and teacher education*, 27(3):619–627.

Scher, S. J. and Osterman, N. M. (2002). Procrastination, conscientiousness, anxiety, and goals: Exploring the measurement and correlates of procrastination among school-aged children. *Psychology in the Schools*, 39(4):385–398.

Scheunemann, A., Schnettler, T., Bobe, J., Fries, S., and Grunschel, C. (2022). A longitudinal analysis of the reciprocal relationship between academic procrastination, study satisfaction, and dropout intentions in higher education. *European Journal of Psychology of Education*, 37(4):1141–1164.

Schouwenburg, H. C. and Lay, C. H. (1995). Trait procrastination and the big-five factors of personality. *Personality and Individual differences*, 18(4):481–490.

Schreiber, J. B., Nora, A., Stage, F. K., Barlow, E. A., and King, J. (2006). Reporting structural equation modeling and confirmatory factor analysis results: A review. *The Journal of educational research*, 99(6):323–338.

Sebastião, R. (2021). Classification of anxiety based on eda and hr. In *IoT Technologies for HealthCare: 7th EAI International Conference, HealthyIoT 2020, Viana do Castelo, Portugal, December 3, 2020, Proceedings*, pages 112–123. Springer.

Senécal, C., Julien, E., and Guay, F. (2003). Role conflict and academic procrastination: A self-determination perspective. *European journal of social psychology*, 33(1):135–145.

Setz, C., Arnrich, B., Schumm, J., La Marca, R., Tröster, G., and Ehlert, U. (2009). Discriminating stress from cognitive load using a wearable eda device. *IEEE Transactions on information technology in biomedicine*, 14(2):410–417.

Shaukat-Jali, R., van Zalk, N., Boyle, D. E., et al. (2021). Detecting subclinical social anxiety using physiological data from a wrist-worn wearable: Small-scale feasibility study. *JMIR Formative Research*, 5(10):e32656.

Shukla, J., Barreda-Angeles, M., Oliver, J., Nandi, G. C., and Puig, D. (2019). Feature extraction and selection for emotion recognition from electrodermal activity. *IEEE Transactions on Affective Computing*, 12(4):857–869.

Smyth, J., Birrell, S., Woodman, R., and Jennings, P. (2021). Exploring the utility of eda and skin temperature as individual physiological correlates of motion sickness. *Applied Ergonomics*, 92:103315.

Solomon, L. J. and Rothblum, E. D. (1984). Academic procrastination: frequency and cognitive-behavioral correlates. *Journal of counseling psychology*, 31(4):503.

Sriramprakash, S., Prasanna, V. D., and Murthy, O. R. (2017). Stress detection in working people. *Procedia computer science*, 115:359–366.

Stadtfeld, C., Vörös, A., Elmer, T., Boda, Z., and Raabe, I. J. (2019). Integration in emerging social networks explains academic failure and success. *Proceedings of the National Academy of Sciences*, 116(3):792–797.

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

Steel, P. and König, C. J. (2006). Integrating theories of motivation. *Academy of management review*, 31(4):889–913.

Stewart, G. L. (2006). A meta-analytic review of relationships between team design features and team performance. *Journal of management*, 32(1):29–55.

Sun, T. and Kim, J.-E. (2022a). The effect of procrastination heterogeneity on team performance. *International Journal of Industrial Ergonomics*, 87:103231.

Sun, T. and Kim, J.-E. (2022b). The effects of online learning and task complexity on students' procrastination and academic performance. *International Journal of Human–Computer Interaction*, pages 1–7.

Sun, T., Tjuandi, G., and Kim, J.-E. (2022). Dynamic nature of procrastination in the online learning environment. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 66, pages 597–600. SAGE Publications Sage CA: Los Angeles, CA.

Taelman, J., Vandeput, S., Spaepen, A., and Van Huffel, S. (2009). Influence of mental stress on heart rate and heart rate variability. In *4th European Conference of the International Federation for Medical and Biological Engineering: ECIFMBE 2008 23–27 November 2008 Antwerp, Belgium*, pages 1366–1369. Springer.

Teo, T., Tsai, L. T., and Yang, C.-C. (2013). Applying structural equation modeling (sem) in educational research: An introduction. In *Application of structural equation modeling in educational research and practice*, pages 1–21. Brill Sense.

The MathWorks Inc. (2022). Matlab version: 9.13.0 (r2022b).

Tice, D. M. and Baumeister, R. F. (1997). Longitudinal study of procrastination, performance, stress, and health: The costs and benefits of dawdling. *Psychological science*, 8(6):454–458.

Timmerman, T. A. (2000). Racial diversity, age diversity, interdependence, and team performance. *Small group research*, 31(5):592–606.

Tuckman, B. W. (1991). The development and concurrent validity of the procrastination scale. *Educational and psychological measurement*, 51(2):473–480.

Val-Calvo, M., Álvarez-Sánchez, J. R., Ferrández-Vicente, J. M., Díaz-Morcillo, A., and Fernández-Jover, E. (2020). Real-time multi-modal estimation of dynamically evoked emotions using eeg, heart rate and galvanic skin response. *International Journal of Neural Systems*, 30(04):2050013.

Van Eerde, W. (2000). Procrastination: Self-regulation in initiating aversive goals. *Applied Psychology*, 49(3):372–389.

van Eerde, W. and Klingsieck, K. B. (2018). Overcoming procrastination? a meta-analysis of intervention studies. *Educational Research Review*, 25:73–85.

Van Hooft, E. A. and Van Mierlo, H. (2018). When teams fail to self-regulate: Predictors and outcomes of team procrastination among debating teams. *Frontiers in psychology*, 9:464.

Van Knippenberg, D. and Schippers, M. C. (2007). Work group diversity. *Annu. Rev. Psychol.*, 58:515–541.

van Offenbeek, M. (2001). Processes and outcomes of team learning. *European journal of work and organizational psychology*, 10(3):303–317.

Vinodh, S. and Joy, D. (2012). Structural equation modelling of lean manufacturing practices. *International Journal of Production Research*, 50(6):1598–1607.

Volkema Gorman, Roger J, R. H. (1998). The influence of cognitive-based group composition on decision-making process and outcome. *Journal of Management Studies*, 35(1):105–121.

Waller, M. J., Zellmer-Bruhn, M. E., and Giambatista, R. C. (2002). Watching the clock: Group pacing behavior under dynamic deadlines. *Academy of Management Journal*, 45(5):1046–1055.

Westland, J. C. (2010). Lower bounds on sample size in structural equation modeling. *Electronic commerce research and applications*, 9(6):476–487.

WHO (2022). Mental health at work. `https://www.who.int/news-room/fact-sheets/detail/mental-health-at-work`. Accessed: 2023-05-25.

Widjaja, A. E. and Tarigan, R. E. (2017). Factors affecting procrastination in the virtual teams: An experimental research approach based on the theory of planned behavior. In *2017 4th International Conference on New Media Studies (CONMEDIA)*, pages 69–74. IEEE.

Williams, K. D. and Karau, S. J. (1991). Social loafing and social compensation: The effects of expectations of co-worker performance. *Journal of personality and social psychology*, 61(4):570.

Wood, R. E. (1986). Task complexity: Definition of the construct. *Organizational behavior and human decision processes*, 37(1):60–82.

Wu, C., Cha, J., Sulek, J., Sundaram, C. P., Wachs, J., Proctor, R. W., and Yu, D. (2021). Sensor-based indicators of performance changes between sessions during robotic surgery training. *Applied ergonomics*, 90:103251.

Wu, H., Gui, D., Lin, W., Gu, R., Zhu, X., and Liu, X. (2016). The procrastinators want it now: behavioral and event-related potential evidence of the procrastination of intertemporal choices. *Brain and Cognition*, 107:16–23.

Xie, Z., Li, L., and Xu, X. (2022). Real-time driving distraction recognition through a wrist-mounted accelerometer. *Human factors*, 64(8):1412–1428.

Yamakoshi, T., Yamakoshi, K.-i., Tanaka, S., Nogawa, M., Park, S.-B., Shibata, M., Sawada, Y., Rolfe, P., and Hirose, Y. (2008). Feasibility study on driver's stress detection from differential skin temperature measurement. In *2008 30th annual international conference of the ieee engineering in medicine and biology society*, pages 1076–1079. IEEE.

Yi, B., Cao, H., Song, X., Wang, J., Zhao, S., Guo, W., and Cao, D. (2023). How can the trust-change direction be measured and identified during takeover transitions in conditionally automated driving? using physiological responses and takeover-related factors. *Human factors*, page 00187208221143855.

Yilmaz, M. B. (2017). The relation between academic procrastination of university students and their assignment and exam performances: The situation in distance and face-to-face learning environments. *Journal of Education and Training Studies*, 5(9):146–157.

You, J. W. (2015). Examining the effect of academic procrastination on achievement using lms data in e-learning. *Journal of educational technology & society*, 18(3):64–74.

Zarrin, S. A., Gracia, E., and Paixão, M. P. (2020). Prediction of academic procrastination by fear of failure and self-regulation. *Educational Sciences: Theory & Practice*, 20(3):34–43.

Zhang, T., Yang, J., Liang, N., Pitts, B. J., Prakah-Asante, K. O., Curry, R., Duerstock, B. S., Wachs, J. P., and Yu, D. (2020). Physiological measurements of situation awareness: a systematic review. *Human factors*, page 0018720820969071.

Zhang, Y. and Kaber, D. (2016). Evaluation of strategies for integrated classification of visual-manual and cognitive distractions in driving. *Human factors*, 58(6):944–958.

Zhang, Y., Tiňo, P., Leonardis, A., and Tang, K. (2021). A survey on neural network interpretability. *IEEE Transactions on Emerging Topics in Computational Intelligence*, 5(5):726–742.

Zhu, X., Gedeon, T., Caldwell, S., Jones, R., and Gu, X. (2020). Deceit detection: identification of presenter's subjective doubt using affective observation neural network analysis. In *2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pages 3174–3181. IEEE.

Zontone, P., Affanni, A., Bernardini, R., Piras, A., and Rinaldo, R. (2019). Stress detection through electrodermal activity (eda) and electrocardiogram (ecg) analysis in car drivers. In *2019 27th European Signal Processing Conference (EUSIPCO)*, pages 1–5. IEEE.