36-627 Project Proposal

Due: April 15, 2026 at 5:00 PM

Dominick Robinson

Executive Summary

We propose a controlled study to evaluate the impact of 8AM course scheduling on student academic performance and sleep wellness, specifically focusing on whether these early start times create systemic inequities across different student demographics. To ensure the results are robust and actionable, we recommend a study size of 384 students. The total projected cost is $\$115,200$, which covers participation rewards and Fitbit costs for each student. By randomly assigning students and instructors across morning and afternoon sections, we can isolate the course time effect from other factors like teaching quality or student preference.

Our analysis is designed to be highly successful in detecting meaningful shifts in student outcomes. We anticipate a near-certain ability to identify even a 0.2 change in term GPA or a 30-minute change in sleep caused by class timing. Furthermore, the study is powered to identify if these early start times disproportionately affect students based on gender or race. While detecting these specific demographic differences is statistically more challenging than measuring broad trends, our proposed sample size provides an estimated 72 success rate in identifying GPA inequities and an estimated 82 success rate in identifying sleep inequities.

There are minor limitations to consider. Because we cannot access students' cumulative GPAs until after they are assigned to sections, we will utilize cumulative GPA as a statistical covariate. This will ensure that we adjust for students' academic baselines, providing a clear picture of the course time impact regardless of the initial section assignment. Additionally, while we have maximized our sample size within reasonable budget constraints, there remains a slight risk that small inequities in GPA may be difficult to confirm.

Design

To estimate the effect of 8AM courses on student performance and sleep, we will implement a Completely Balanced Factorial 2^5 Design. This design allows us to isolate the effect of time while controlling for variance resulting from demographic qualities and instructor.

We will block on race, gender, year, and instructor so that our treatment is balanced across demographic qualities and instructor.

While cumulative GPA is expected to be a major predictor of success, it cannot be used as a blocking variable because student GPA data is not accessible prior to section assignment. Therefore, we will include Cumulative GPA as a covariate in our final model to control for possible correlation between time and cumulative GPA. If student assignments are random, then cumulative GPA should not be correlated with time like it might typically be when allowing students to register for times they want (e.g. higher performing students may be more likely to sign up for morning classes, or students who don't get proper sleep may be more likely to sign up for later class times).

To ensure that instructor and time are not aliased, we will randomly assign each of the instructors to teach exactly one morning and one afternoon section.

Students will be randomly recruited and assigned such that there will be an equal amount of students with each race, gender, and year in each section (and therefore for each instructor) combination. Based on our power analysis, we would like to experiment on 384 total students.

See Table 2 for the instructor layout and Table 3 for the demographic distribution per section.

Power Analysis

To determine the required sample size, we performed a power analysis using residual variance estimates from pilot study data ($N=$630). We specifically analyzed the power to detect a 0.20 point drop in term GPA and a 30 minute reduction in total sleep time, which we felt were reasonable minimal effect sizes that are important to test for.

We used the worst-case noncentral F distribution logic from *Theorem 8.2*, and we found that the experiment has high power (> 0.99) for main effects even at small sample sizes.

However, detecting disproportionate harm to certain demographics is more difficult. Even when recruiting 384 students, which is near the maximum of 400, with 12 per cell, we achieve a power of 0.717 for GPA interactions and 0.823 for Sleep interactions, the former of which falls short of the generally agreed upon 80% power to aim for. Because identifying inequity is so important to the university, although the total cost of $\$115,200$, we recommend recruiting the 384 students for the study to maximize our ability to identify possible inequity in early morning classes on GPA.

See Table 1 for a complete breakdown of our power analysis.

Data Analysis

To test whether class time inequitablty affects term GPAs and amount of sleep, we will fit the following models:

$$
\begin{aligned}
Y_{\text{Term GPA}} &= \beta_0 + \beta_1 X_{\text{Cum. GPA}} + \tau_{\text{time}} + \theta_{\text{race}} + \theta_{\text{gender}} + \theta_{\text{year}} + \theta_{\text{instructor}} + (\tau\theta)_{\text{time}\times\text{race}} + (\tau\theta)_{\text{time}\times\text{gender}} + \epsilon \\
Y_{\text{Sleep}} &= \beta_0 + \beta_1 X_{\text{Cum. GPA}} + \tau_{\text{time}} + \theta_{\text{race}} + \theta_{\text{gender}} + \theta_{\text{year}} + \theta_{\text{instructor}} + (\tau\theta)_{\text{time}\times\text{race}} + (\tau\theta)_{\text{time}\times\text{gender}} + \epsilon
\end{aligned}
$$

where race, gender, year, and instructor are fixed blocking factors, and time is a treatment factor. Cumulative GPA is included as a covariate to control for additional variance, although we were unable to block on it.

We will evaluate the significance of the $(\tau\theta)_{\text{time}\times\text{race}}$ and $(\tau\theta)_{\text{time}\times\text{gender}}$ terms using ANCOVA tests.

The hypothesis tests will be conducted for each model:

$$
\begin{aligned}
H_0 &: (\tau\theta)_{\text{time}\times\text{race}} = 0 \\
H_\alpha &: (\tau\theta)_{\text{time}\times\text{race}} \neq 0
\end{aligned}
$$

Following this, to test whether class time impacts each of term GPA and total sleep, we will fit the same ANCOVA models as before, dropping the insignificant interaction terms to maximize our degrees of freedom, and run an ANCOVA test to measure the significance of τ_{time}. The following hypothesis tests will be conducted for each model:

$$
\begin{aligned}
H_0 &: \tau_{\text{time}} = 0 \\
H_\alpha &: \tau_{\text{time}} \neq 0
\end{aligned}
$$

Each of the hypothesis tests will be conducted at the 95% confidence level.

Appendix

Code and results of power analysis

```
library(tidyverse)

fit_gpa <- lm(term_gpa ~ cum_gpa + factor(demo_race) + factor(demo_gender), data = sleep)
sig2_gpa <- sigma(fit_gpa)^2

fit_sleep <- lm(TotalSleepTime ~ factor(demo_race) + factor(demo_gender), data = sleep)
sig2_sleep <- sigma(fit_sleep)^2

calc_main_power <- function(n_cell, delta, sigma2) {
  v <- 2          # treatment levels: morning or afternoon
  n <- n_cell * (2 * 2 * 2 * 2 * 2)   # n_cell x (# cells: race x gender x year x instructor
  r <- n / v   # replicates per blocking combination
  alpha <- 0.05

  df1 <- v - 1
  df2 <- n - 7   # interaction + 6 mains (race x gender x year x instructor x time x cum_gpa)

  f_crit <- qf(1 - alpha, df1, df2)

  delta2 <- (r * delta^2) / (2 * sigma2)

  power <- pf(f_crit, df1, df2, ncp = delta2, lower.tail = FALSE)
  return(power)
}

calc_interaction_power <- function(n_cell, delta, sigma2) {
  alpha <- 0.05
  n_total <- n_cell * (2 * 2 * 2 * 2 * 2)   # n_cell x (race x gender x year x instructor x
  r <- n_total / (2 * 2)   # replicates per subgroup (race/gender and time combination)

  df1 <- 1   # testing one interaction
  df2 <- n_total - 8   # intercept + 6 main + 1 interaction

  f_crit <- qf(1 - alpha, df1, df2)

  delta2 <- (r * delta^2) / (4 * sigma2)
```

```
  power <- pf(f_crit, df1, df2, ncp = delta2, lower.tail = FALSE)
  return(power)
}

n_cell_range <- seq(2, 12, by = 2)

power_analysis_table <- data.frame(n_cell = seq(2, 12, by = 2)) |>
  mutate(
    `Total Students`     = n_cell * 32,
    `Total Cost`  = scales::dollar(`Total Students` * 300),
    `GPA vs Time`    = map_dbl(n_cell, ~calc_main_power(.x, 0.2, sig2_gpa)),
    `Sleep vs Time`  = map_dbl(n_cell, ~calc_main_power(.x, 30, sig2_sleep)),
    `GPA vs Time and Race`  = map_dbl(n_cell, ~calc_interaction_power(.x, 0.2, sig2_gpa)),
    `GPA vs Time and Gender`  = map_dbl(n_cell, ~calc_interaction_power(.x, 0.2, sig2_gpa)),
    `Sleep vs Time and Race`  = map_dbl(n_cell, ~calc_interaction_power(.x, 30, sig2_sleep))
    `Sleep vs Time and Gender`  = map_dbl(n_cell, ~calc_interaction_power(.x, 30, sig2_sleep
  )

library(knitr)

power_analysis_table |>
  mutate(across(where(is.numeric), ~round(.x, 3))) |>
  kable(
    caption = paste0(
      "Power analysis for 8am class study. ",
      "Calculated to detect a 0.2 GPA difference and 30 minute sleep difference. ",
      "Estimates based on pilot residual variances (GPA sigma^2 = ", round(sig2_gpa, 4),
      ", Sleep sigma^2 = ", round(sig2_sleep, 4), ")."
    ),
    col.names = c(
      "Replicates Per Cell",
      "Total Students",
      "Total Cost",
      "GPA Main (Power)",
      "Sleep Main (Power)",
      "GPA x Race (Power)",
      "GPA x Gender (Power)",
      "Sleep x Race (Power)",
      "Sleep x Gender (Power)"
    )
  )
```

Table 1: Power analysis for 8am class study. Calculated to detect a 0.2 GPA difference and 30 minute sleep difference. Estimates based on pilot residual variances (GPA sigma^2 = 0.1488, Sleep sigma^2 = 2580.0254).

Replicates Per Cell	Total Students	Total Cost	GPA Main (Power)	Sleep Main (Power)	GPA x Race (Power)	GPA x Gender (Power)	Sleep x Race (Power)	Sleep x Gender (Power)
2	64	$19,200	0.531	0.642	0.175	0.175	0.213	0.213
4	128	$38,400	0.829	0.912	0.307	0.307	0.381	0.381
6	192	$57,600	0.947	0.983	0.431	0.431	0.530	0.530
8	256	$76,800	0.985	0.997	0.542	0.542	0.653	0.653
10	320	$96,000	0.996	1.000	0.637	0.637	0.750	0.750
12	384	$115,200	0.999	1.000	0.717	0.717	0.823	0.823

Code to generate design CSV

```r
library(tidyverse)

design <- expand.grid(
  year = c("freshman", "sophomore"),
  demo_race = c(0, 1),
  demo_gender = c(0, 1),
  section = 1:4
) |> slice(rep(1:n(), each = 12))

design <- design |>
  mutate(
    section = case_when(
      section == 1 ~ "8am-1",
      section == 2 ~ "8am-2",
      section == 3 ~ "130pm-1",
      section == 4 ~ "130pm-2"
    ),
    instructor = case_when(
      section == "8am-1" ~ "A",
      section == "8am-2" ~ "B",
      section == "130pm-1" ~ "B",
      section == "130pm-2" ~ "A"
    )
  )
```

```r
design_file <- "dominicr-design.csv"
write.csv(design, design_file, row.names = FALSE)

source("check-project-design.R")
check_design(design_file)
```

Our design passes the design check.

Useful counts to illustrate design

Table 2: Student distribution by section and instructor.

section	instructor	n
130pm-1	B	96
130pm-2	A	96
8am-1	A	96
8am-2	B	96

Table 3: Demographic breakdown per individual section.

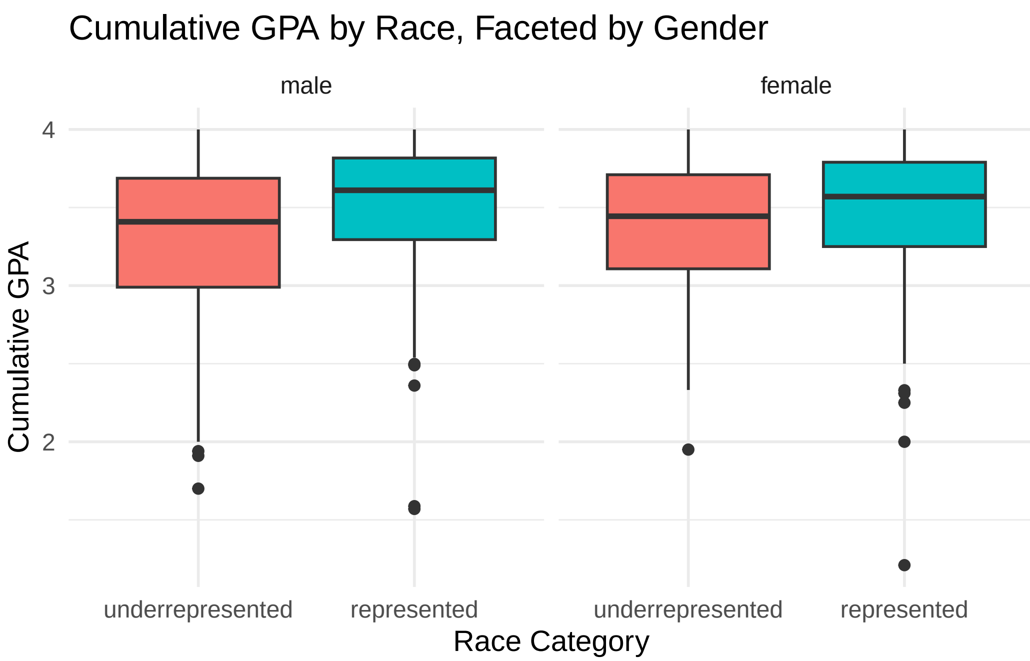
demo_race	demo_gender	year	n
Underrepresented	Male	freshman	12
Underrepresented	Male	sophomore	12
Underrepresented	Female	freshman	12
Underrepresented	Female	sophomore	12
Represented	Male	freshman	12
Represented	Male	sophomore	12
Represented	Female	freshman	12
Represented	Female	sophomore	12

Distributions of variables in pilot study

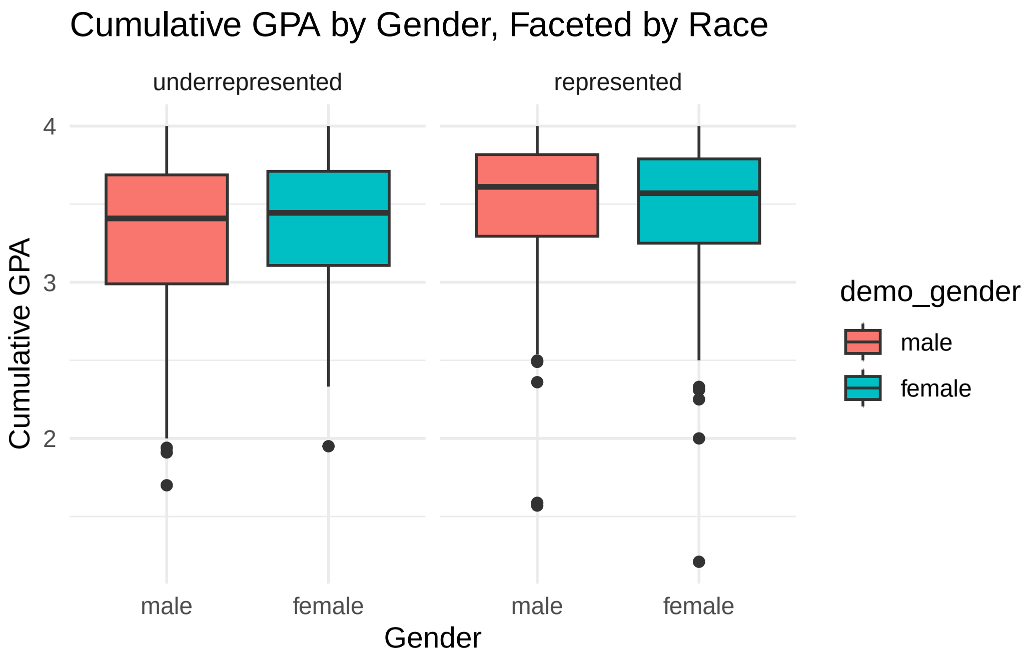
```r
library(ggplot2)

ggplot(sleep, aes(x = demo_race, y = cum_gpa, fill = demo_race)) +
  geom_boxplot() +
```

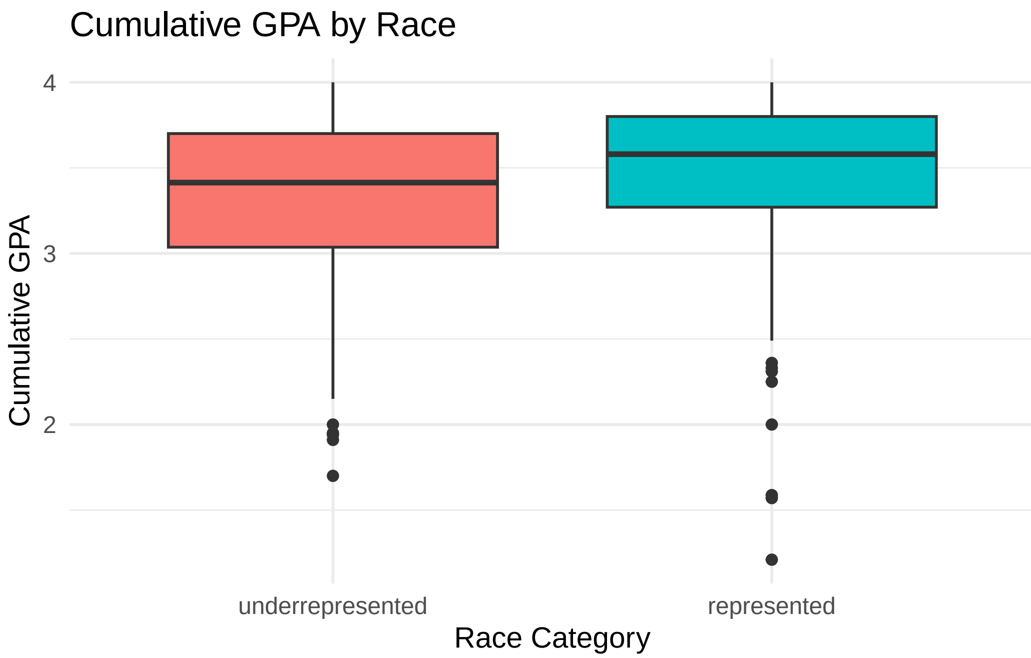
```
  facet_wrap(~demo_gender) +
  labs(title = "Cumulative GPA by Race, Faceted by Gender",
       x = "Race Category",
       y = "Cumulative GPA") +
  theme_minimal() +
  theme(legend.position = "none")
```



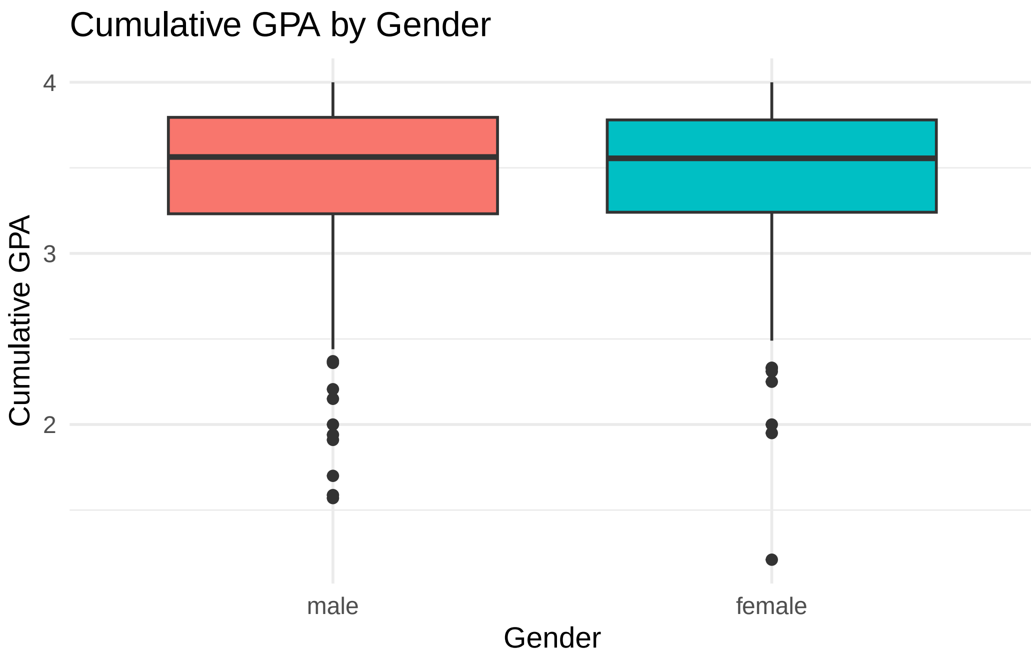

```
ggplot(sleep, aes(x = demo_gender, y = cum_gpa, fill = demo_gender)) +
  geom_boxplot() +
  facet_wrap(~demo_race) +
  labs(title = "Cumulative GPA by Gender, Faceted by Race",
       x = "Gender",
       y = "Cumulative GPA") +
  theme_minimal()
```




```
ggplot(sleep, aes(x = demo_race, y = cum_gpa, fill = demo_race)) +
  geom_boxplot() +
  labs(title = "Cumulative GPA by Race",
       x = "Race Category",
       y = "Cumulative GPA") +
  theme_minimal() +
  theme(legend.position = "none")
```




```r
ggplot(sleep, aes(x = demo_gender, y = cum_gpa, fill = demo_gender)) +
  geom_boxplot() +
  labs(title = "Cumulative GPA by Gender",
       x = "Gender",
       y = "Cumulative GPA") +
  theme_minimal() +
  theme(legend.position = "none")
```




```
table(sleep$demo_race, sleep$demo_gender)
```

```
                   male female
  underrepresented   59     60
  represented       204    307
```

```
prop.table(table(sleep$demo_race, sleep$demo_gender))
```

```
                         male     female
  underrepresented 0.09365079 0.09523810
  represented      0.32380952 0.48730159
```

```
summary(aov(cum_gpa ~ demo_gender + demo_race, data = sleep))
```

```
             Df Sum Sq Mean Sq F value   Pr(>F)    
demo_gender   1   0.04   0.042   0.228    0.633    
demo_race     1   4.12   4.123  22.130 3.14e-06 ***
Residuals   627 116.82   0.186                     
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```