Daily Exercise Regimens and Quality Sleep?

Howie, Daniel, Max

Research Question:

Does undertaking a daily exercise regimen improve overall sleep quality among college students?

Hypotheses:

College students who engage in a daily exercise regimen will show an increase in total sleep time (TST) after two weeks, compared to a control group. Students who exercise will demonstrate lower sleep latency, fewer nighttime awakenings, lower wake after sleep onset (WASO), and higher sleep efficiency. The effect of exercise on sleep quality will be moderated by individual differences in chronotype, sleep environment, and alcohol/drug use.

Study Design:

We will recruit approximately 1,000 undergraduate students from universities in Georgia. Participants will be screened to exclude those who regularly take melatonin or prescription medications that significantly affect sleep. A power calculator will be used to confirm that the sample size is sufficient to detect expected effects. To reduce self-reporting bias, our primary outcome (TST) and secondary outcomes (sleep latency, awakenings, WASO, and sleep efficiency) will be measured using wearable sleep-tracking devices. These devices will also be capable of monitoring physical activity, including heart rate elevation and movement patterns, to confirm exercise compliance.

Treatment & Control Conditions:

The treatment group will be asked to complete at least 30 minutes of moderate-intensity aerobic exercise (e.g., jogging, cycling, brisk walking) daily for 14 consecutive days. Participants will be encouraged to exercise at approximately the same time each day to promote consistency. The control group will not be asked to modify their activity levels. Baseline & Compliance: All participants will wear the sleep-tracking device for one week before the intervention begins to establish baseline sleep metrics. Compliance will be supported through daily check-ins (via app or text message) and small incentives (e.g., gift cards, entry into a raffle) for participants who complete their assigned activities.

Data Analysis:

We will use OLS regression to estimate the effect of daily exercise on sleep quality, controlling for gender, chronotype, sleep environment quality, and self-reported alcohol/drug use. While we acknowledge the limitations of self-reported substance use, we anticipate the effect of this covariate may be small. To address the risk of false positives from multiple hypothesis testing, we will apply corrections such as the Bonferroni method.

Section One: Introduction

Researcher Names: Max, Howie, Daniel

Research project Title: Daily Exercise Regimens and Quality Sleep

One Sentence Summary of Your Specific Research Question:

How do daily exercise regimens impact the quality of sleep among college students?

General Motivation

Why should someone who is not an academic care about the results of this research?

- Someone could use this information for their own lives to hopefully implement this to see how to optimize their sleep.

What policy decisions will your research help inform - Should the government implement required physical health/education courses for universities in the United States? - It is well known that adequate sleep maximizes academic and work potential amongst students, so by implementing exercise routines, are students more inclined to get more sleep because of physical exertion in addition to the mental energy required for school?

Theoretical Motivation

What theoretical questions can this research shed light on Key debates/literature that will be informed by the answer to your research question

Primary Hypotheses:

What are the key parameter/estimands the research design seeks to estimate? What sign and/or magnitude is predicted by primary hypotheses for each parameter/estimand? [1-2 paragraphs]

The primary estimand in this study is the average treatment effect (ATE) of a daily exercise regimen on total sleep time (TST) among college students. Secondary estimands include the effects of exercise on sleep latency, number of awakenings, wake after sleep onset (WASO), and sleep efficiency. We predict a positive effect of exercise on TST and sleep efficiency, and negative effects on sleep latency, awakenings, and WASO—indicating improvements in sleep quality. In terms of magnitude, prior literature suggests moderate effects, particularly for previously sedentary individuals, with TST increasing by 20–40 minutes and improvements in sleep efficiency by 5–10 percentage points after consistent physical activity.

What is the logic or theory of change behind the primary hypotheses [1-2 paragraphs]

The theory of change behind these hypotheses is grounded in the physiological and psychological impacts of exercise. Physical activity promotes sleep by regulating circadian rhythms, reducing arousal and stress levels, and increasing homeostatic sleep drive, especially through the depletion of energy stores. Regular aerobic exercise, in particular, has been linked to deeper and more restorative slow-wave sleep. Among college students, who often experience sleep disruption due to irregular schedules, stress, and sedentary behavior, introducing a consistent exercise routine may counteract these effects by creating biological regularity and improving mood, which in turn enhances sleep initiation and maintenance.

What are the key pieces in the relevant academic literature that inform your hypotheses? [2-3 pieces]

Relevant Academic Literature:

- Kredlow et al. (2015) conducted a meta-analysis showing that both acute and regular exercise were associated with small-to-moderate improvements in sleep quality, particularly in sleep onset latency and sleep efficiency. (**kredlow2015effects?**)
- Baron et al. (2013) found that individuals with insomnia who participated in a 16-week exercise intervention showed significantly improved sleep outcomes, including decreased WASO and increased TST. (**baron2013exercise?**)
- Driver & Taylor (2000) provided theoretical grounding on how exercise influences sleep architecture, including increases in slow-wave sleep and decreases in REM latency, emphasizing the role of energy expenditure and thermoregulation.

(**driver2000exercise?**)

Secondary Hypotheses:

What are the secondary paramater/estimands the research design seeks to estimate? What sign and/or magnitude is predicted by the secondary hypotheses for each parameter/estimand [These may be conditional effects for subgroups or hypotheses about additional outcomes or cross- randomized treatments.]

The secondary parameters of this study focus on conditional effects—specifically, how the impact of daily exercise on sleep quality may vary by chronotype, sleep environment quality, and alcohol/drug use. These are moderating variables, and the estimands of interest are interaction effects between the treatment (exercise) and each of these moderators. We predict that evening chronotypes (those who naturally fall asleep and wake later) will benefit more from the exercise intervention than morning types, as physical activity may help regulate their circadian rhythms and promote earlier sleep onset. For students with poor sleep environments (e.g., noisy or uncomfortable rooms), we anticipate a weaker positive effect, since environmental disruptions may offset the benefits of exercise. For students who report frequent alcohol or drug use, we expect the exercise effect to be smaller or negligible, as substances can interfere with sleep architecture regardless of physical activity. The magnitude of these moderating effects is likely to be modest but directionally consistent with the predictions.

What is the logic or theory of change behind each secondary hypothesis? [Explain what effects we should expect if the theory behind your primary hypothesis is correct.]

If the theory behind the primary hypothesis is correct—that exercise improves sleep quality by enhancing circadian stability, reducing stress, and promoting physical fatigue—then we should expect its effectiveness to be amplified or dampened by individual and contextual factors that interact with these mechanisms. For example, chronotype reflects internal circadian timing, and exercise may serve as a behavioral cue (a "zeitgeber") that helps shift biological rhythms earlier in the day. Thus, evening types may show greater improvements. Likewise, the sleep environment can moderate the pathway from exercise to better sleep: even if someone is more physiologically primed for sleep due to exercise, an uncomfortable or loud environment could still prevent restorative rest. Finally, alcohol and drug use can chemically disrupt sleep stages and override the beneficial effects of physical exertion, suggesting that among frequent users, exercise alone may not significantly change sleep outcomes. These subgroup analyses help us understand for whom and under what conditions exercise is most effective as a behavioral sleep intervention.

Alternative explanations if results are consistent with hypotheses:

What alternative theories could explain the results? Hypothesis for an alternative outcome (or other subgroups) that would be consistent only with the alternative explanation and not the logic behind your primary hypothesis

If the results show that students who exercised daily experienced improved sleep quality, several alternative explanations—distinct from the physiological mechanisms in our primary theory—could account for these findings. One possibility is the placebo or expectation effect: students in the exercise group might believe that exercising should improve their sleep and therefore perceive their sleep as better, even if objective improvements are minimal. While the use of objective sleep-tracking devices reduces this bias, participants' behavior (e.g., going to bed earlier, feeling more relaxed) might still be influenced by their expectations. Another plausible explanation is that participating in a structured routine—not the exercise itself—improves sleep. The intervention could increase general self-discipline, reduce time spent on screens before bed, or promote better time management. These behavior changes might improve sleep regardless of the physical activity component.

Hypothesis for an Alternative Outcome (or Subgroup Evidence): If improved sleep outcomes are observed equally in participants who engage in structured, non-physical daily activities (e.g., meditation, reading at the same time each evening), that would support the routine/structure hypothesis rather than the physiological theory behind the exercise-specific intervention. Similarly, if improvements occur regardless of exercise intensity or time of day, that would undermine the physiological mechanisms tied to circadian regulation or energy expenditure, and suggest that behavioral predictability is the key driver. In addition, if participants

with poor sleep environments or frequent substance use benefit equally from the intervention, this would challenge our primary logic that environmental and behavioral factors moderate the physiological benefits of exercise. It would suggest the improvements stem from non-specific effects—like attention from researchers or general increases in health awareness.

Alternative explanations if results are inconsistent with hypotheses:

What alternative theories could explain the results?

If the results show no significant improvement in sleep quality for students who exercised daily, or if the control group performs similarly or better, several alternative explanations could account for the findings. One possibility is the timing and intensity of exercise were not optimized to impact sleep. For instance, exercising too late in the evening may increase arousal and core body temperature, delaying sleep onset rather than promoting it. Similarly, if the intensity was too low (e.g., walking slowly), it may not have triggered the physiological fatigue needed to enhance sleep. Another explanation is that college students' sleep quality is more strongly influenced by psychosocial stressors, academic workload, or irregular schedules than by physical activity. In that case, the marginal benefit of adding exercise may be too small to detect against the background noise of student life. Additionally, the duration of the intervention (14 days) might not have been long enough to create lasting physiological or behavioral changes. It's also possible that non-compliance diluted the treatment effect—if participants did not consistently follow the exercise regimen, the average treatment effect would be underestimated. Finally, baseline activity levels may play a role. If many participants in the sample were already moderately active, then adding more exercise might have had little additional impact, suggesting a ceiling effect. This would mean the intervention is effective only for previously sedentary individuals—a subgroup distinction the primary hypothesis did not account for.

Section Two: Population and Sample

Population of interest:

Where and when will your study take place?

Does this match up to your population of interest, or are there conditions that make this study context different? The study will take place across multiple universities in Georgia during the fall academic semester, when students are on campus and following regular class schedules. This timing aligns with our target population—undergraduate college students—allowing us to observe sleep patterns in a typical academic context. However, factors like midterm stress or varying campus cultures may introduce contextual differences that slightly

limit generalizability to all college populations or to students during less structured periods like summer break.

Sample size:

How is this sample selected? Be specific about the procedure.

The sample will be selected through stratified random sampling across several universities in Georgia to ensure diversity in gender, academic year, and baseline activity levels. Recruitment will occur via campus-wide emails, student organization listservs, flyers in common areas, and classroom announcements. Interested students will complete an online screening survey to assess eligibility, including questions about current physical activity habits, sleep medication use, and availability to participate for the full study period. Eligible participants will be randomly assigned to either the exercise or control group. To ensure balance, randomization will be stratified by chronotype and gender. Participants will provide informed consent before beginning the study.

Consent:

How will you obtain informed consent? If you will not, what is the justification? Is this population vulnerable to being coerced into participating in the study? Informed consent will be obtained through a digital form provided at the beginning of the screening survey. The form will clearly explain the study's purpose, procedures, potential risks and benefits, data privacy protections, and the voluntary nature of participation. Students will have the opportunity to ask questions before consenting. While college students are not legally considered a vulnerable population, care will be taken to minimize any sense of coercion—participation will not be tied to course credit or grades, and it will be emphasized that they may withdraw at any time without penalty.

Ethics:

Is the sample size large enough that you have sufficient power for your research conclusions to be credible and useful? Is the sample size no larger than necessary for the research? Can the research (results) be used to target people or make people more vulnerable? The planned sample size of 1,000 participants is likely sufficient to ensure adequate statistical power to detect small to moderate treatment effects on sleep quality, especially given the use of objective measurements and multiple covariates. A power analysis will be conducted to confirm that this number is not larger than necessary, ensuring efficiency and ethical use of participant time and resources. The research does not involve sensitive personal data or predictive algorithms that could be used to target individuals. Results will be reported in aggregate, and findings

are intended to inform general sleep and wellness interventions—not to stigmatize or make any group more vulnerable.

Section Three: Intervention

Status Quo

Describe the status quo–what are the current conditions in terms of the outcomes you hope to change? What aspects of the intervention already exist, if any?

In the current status quo, many college students experience poor sleep quality, characterized by short sleep duration, irregular sleep schedules, and frequent awakenings. Research indicates that a large proportion of students get less than the recommended 7–9 hours of sleep per night, often due to academic stress, screen use, inconsistent routines, and substance use. Although exercise is widely recommended as a healthy habit, there is no consistent effort on most campuses to promote daily physical activity specifically as a tool for improving sleep. Some students may engage in exercise casually or for fitness goals, but structured, routine-based exercise programs tied to sleep outcomes are rarely implemented or evaluated. Thus, the intervention—daily, moderate-intensity exercise with monitored compliance—builds on existing behaviors but introduces a targeted, consistent approach intended to improve sleep quality in a measurable way.

Intervention

Describe your intervention(s) What is already known about the effect of the proposed intervention relative to the status quo? Is there credible evidence on the question?

The intervention consists of a structured daily exercise regimen, requiring participants in the treatment group to engage in at least 30 minutes of moderate-intensity aerobic activity (e.g., brisk walking, jogging, or cycling) at a consistent time each day for 14 days. Compliance will be monitored through wearable devices and daily check-ins. This differs from the status quo, where exercise is irregular, unstructured, or pursued for reasons unrelated to sleep. There is credible evidence suggesting that regular physical activity improves various aspects of sleep quality, including total sleep time, sleep onset latency, and sleep efficiency. Meta-analyses and controlled trials—such as those by Kredlow et al. (2015) and Kline et al. (2013)—have shown that both acute and chronic exercise can positively affect sleep outcomes, particularly in populations with poor sleep quality or insomnia. However, much of the existing research has focused on adults or clinical populations, so this study contributes valuable evidence specific to college students—a group often overlooked despite their high rates of sleep disruption.

Control

Describe the control condition Is the control condition a pure control (no intervention at all) or a placebo? What is the placebo condition designed to control for?

The control condition in this study is a pure control, meaning participants in this group will not receive any intervention or be asked to change their daily routines. They will continue with their usual activities, including their typical exercise and sleep habits, without any added structure or guidance. This design allows for a clear comparison between those who adopt a consistent exercise regimen and those who do not, isolating the effect of the intervention itself. Since the control is not a placebo (e.g., a non-exercise routine or low-intensity task), it does not control for potential expectation effects or participant engagement, which are limitations we acknowledge. However, the objective nature of our outcome measures (collected through wearable devices) helps reduce the risk of bias due to these factors.

Units

To what units (level) will the intervention be applied? Individual, classroom, school, village, municipality, etc. Is this the same level at which outcomes will be measured? If not, how will you address the different levels if they do not perfectly overlap?

The intervention will be applied at the individual level, with each participant randomly assigned to either the exercise or control group. Outcomes—such as total sleep time, sleep latency, and sleep efficiency—will also be measured at the individual level using wearable sleep-tracking devices. Because both the treatment and the outcomes are implemented and assessed at the same unit of analysis, there is no issue of level mismatch. This alignment ensures that the estimated effects can be directly attributed to the intervention applied to each individual, improving the internal validity of the study.

Compliance

What does it mean to "take" (comply with) the the intervention? If the intervention is a program, how much someone need to attend (showing up once? finishing the program?) in order to count as having attended?

To "take" or comply with the intervention means that a participant in the treatment group completes at least 30 minutes of moderate-intensity aerobic exercise each day for the full 14-day intervention period. Compliance will be verified using wearable devices that track physical activity and heart rate, as well as daily check-ins. To be considered as having "attended" or complied with the intervention, a participant must complete at least 12 out of the 14 sessions (about 85% adherence). This threshold ensures a meaningful dose of the intervention while

allowing for minor lapses. Participants who fall below this threshold will be considered non-compliant in per-protocol analyses, though they will still be included in intention-to-treat analyses.

Non- Compliance

Is there any concern with non-compliance (either taking the intervention if assigned to control/placebo or failing to take the intervention if assigned to treatment)?

Yes, non-compliance is a potential concern in both directions. Participants assigned to the treatment group may fail to complete the daily exercise regimen, either due to time constraints, lack of motivation, illness, or competing academic obligations. Despite daily check-ins and wearable tracking, maintaining consistent engagement over 14 days can be challenging for college students. Conversely, participants in the control group may independently engage in regular exercise, especially if they are already health-conscious or become more motivated due to study participation. This form of "treatment contamination" could dilute the observed treatment effect. To address these issues, we will conduct both intention-to-treat (ITT) and per-protocol analyses. The ITT analysis will estimate the average effect of being assigned to the treatment, regardless of compliance, while the per-protocol analysis will estimate the effect among those who adhered to the exercise regimen. Additionally, we will monitor control group activity levels via the same wearable devices to account for any unexpected behavior shifts.

Ethics

Is the control condition no worse than the status quo, according to the best evidence available? Are there concerns that participants may be forced to comply with the intervention? What are the risks and magnitude of potentially negative effects of the treatment? Are such risks concentrated on a particular subset of your population?

Yes, the control condition reflects the status quo and is no worse than current conditions, based on the best available evidence. Control participants will not be asked to change their routines and will continue their normal behavior, which aligns with how most college students currently approach sleep and exercise. There is no evidence suggesting that this would put them at any additional risk. There are no concerns that participants will be forced to comply with the intervention. Participation is entirely voluntary, and students will be reminded that they may withdraw from the study or skip sessions at any time without penalty. Incentives are small and designed to encourage participation, not coerce it. The risks of the treatment are minimal, especially since the intervention consists of moderate-intensity aerobic exercise, which is generally safe for healthy college-aged adults. Potential risks include mild muscle soreness, fatigue, or, in rare cases, injury due to improper form or overexertion. These risks are not expected to be severe and are likely concentrated among students who are completely sedentary at baseline or have underlying health issues that were not disclosed. To mitigate

this, participants will be asked to confirm they are medically able to engage in exercise before enrollment, and safety guidelines will be provided at the start of the study.

Section Four: Outcome and Covariates

Primary Outcome

What is your primary outcome?

The primary outcome of the study is total sleep time (TST), measured in minutes per night. This outcome will be objectively recorded using wearable sleep-tracking devices worn by participants throughout the study period. TST provides a direct, quantifiable measure of overall sleep duration and is a widely accepted indicator of sleep quality and health.

Measurement

How will it be measured? (Give the actual text of the survey question and response options, if using a survey measure. Is the outcome continuous, binary, etc.?)

The primary outcome, total sleep time (TST), will be measured objectively using wearable sleep-tracking devices (e.g., Fitbit, WHOOP, or similar validated consumer-grade trackers). These devices record minute-by-minute data on sleep duration based on movement, heart rate, and other physiological indicators. TST will be captured as a continuous variable, recorded in minutes of sleep per night. Since this measure is not based on a survey, there is no survey question associated with it. However, if a backup self-report measure is used for triangulation or in case of device failure, the survey question would read: "On average, how many total hours of sleep did you get per night over the past week?"

Response options: Less than 4 hours 4–5 hours 6–7 hours 7–8 hours 8–9 hours More than 9 hours Still, the primary measure remains continuous and device-recorded, providing higher accuracy and avoiding self-report bias.

Priors

What is the expected distribution of the primary outcome? (This may come from a prior study on a similar population or you may have to make an educated guess).

The expected distribution of the primary outcome, total sleep time (TST), is likely to be approximately normal but slightly left-skewed, based on prior studies of college student sleep patterns. Most college students tend to average between 5 to 7 hours of sleep per night, with fewer students reaching the recommended 7 to 9 hours. A study by Lund et al. (2010) on sleep habits in college students found an average TST of around 6.5 hours with a standard

deviation of about 1 to 1.5 hours. We therefore anticipate a mean TST of approximately 390–420 minutes, with a standard deviation of 60–90 minutes. Some outliers may report very low sleep durations, particularly during high-stress periods like midterms, which could contribute to a slight left skew.

Validity and measurement error

Is there any concern with untruthful reporting? If so, how will you address it? There is minimal concern with untruthful reporting for the primary outcome of total sleep time, since it will be measured objectively using wearable devices. This reduces reliance on self-reported data and helps avoid bias from inaccurate or exaggerated responses. However, for secondary variables like alcohol or drug use, sleep environment quality, or self-reported chronotype, there is a greater risk of untruthful or socially desirable responses. To address this, participants will complete surveys anonymously using a secure online platform, and questions will be framed in a nonjudgmental way to reduce response bias. Additionally, these self-reported variables will be treated as covariates rather than primary outcomes, so any bias in them is less likely to compromise the study's central conclusions.

Stages

Will you collect a baseline? Will you collect a midline? Will you collect multiple waves of endline measurement? If you will collect a baseline or midline, how will you find the same respondents (minimize attrition?)

Yes, we will collect a baseline. All participants will wear the sleep-tracking device for one full week prior to the start of the intervention to establish their baseline sleep patterns. This will allow us to measure within-subject changes and improve the precision of treatment effect estimates. We do not plan to collect a midline, as the intervention is relatively short (14 days), and the key comparisons will be between pre- and post-intervention sleep data. We will collect one wave of endline data immediately following the 14-day intervention. Participants will continue wearing the device through the entire study period to ensure continuous data collection, eliminating the need for separate re-contact at endline. To minimize attrition, we will maintain regular communication with participants through daily check-ins, reminders, and small incentives. All study procedures, including surveys and check-ins, will be mobile-friendly and low-burden. We will also track compliance in real-time and follow up promptly with participants who miss scheduled activities or device usage.

Covariates

What covariate data do you need, including for subgroup analysis? How will covariates be measured? What additional covariates (if any) will you measure? What additional outcomes

or covariates will you collect to distinguish between your explanation and alternatives if your findings are consistent with your hypothesis?

The key covariate data needed includes chronotype, sleep environment quality, gender, alcohol and drug use, and baseline physical activity levels. These covariates will help control for individual differences in sleep patterns and support subgroup analyses. Chronotype will be measured using a brief version of the Morningness-Eveningness Questionnaire. Sleep environment quality will be assessed through survey questions about room noise, light, temperature, and comfort. Gender will be self-reported. Alcohol and drug use will be measured with frequency-based questions over the past two weeks. Baseline physical activity levels will be collected both through wearable data and a short survey asking about typical weekly activity. To distinguish between our primary explanation (exercise directly improves sleep) and alternative explanations (e.g., routine structure or placebo effects), we will collect additional covariates and outcomes. These include perceived stress (measured with a brief validated scale like the Perceived Stress Scale), screen time before bed, and self-reported expectations about whether participants believed the intervention would help their sleep. These additional measures will help us test whether observed sleep improvements are more strongly linked to psychological or behavioral changes unrelated to physical exertion.

Ethics

Will data collection be onerous (time, effort) or painful (physically, emotionally) for any respondents? Are these costs necessary? Have they been minimized? Are they outweighed by the potential benefits of the research to society?

Data collection is not expected to be particularly onerous or painful for respondents. The main requirements are wearing a sleep-tracking device continuously for three weeks (one week baseline, two weeks intervention) and completing brief daily check-ins, along with a few short surveys at the beginning and end of the study. The physical activity involved in the intervention is moderate and voluntary, with safety guidance provided to minimize discomfort or injury. These minimal costs are necessary to collect accurate, high-quality data and ensure compliance with the exercise regimen. Effort has been made to keep tasks low-burden—daily check-ins will be short, mobile-friendly, and automated where possible. Participants are also allowed to skip a small number of exercise sessions without being excluded from per-protocol analysis. The potential benefits of this research—improving understanding of how behavioral interventions like exercise can enhance sleep in young adults—are substantial. Poor sleep is a widespread issue among college students and is linked to worse academic performance, mental health, and long-term health outcomes. If the findings can inform low-cost, scalable interventions, the social value clearly outweighs the relatively minor participant burden.

Section Five: Randomization

Randomization strategy

Complete/simple, block, cluster, factorial etc.

The randomization strategy for this study will be block randomization, stratified by key variables such as gender and chronotype (morning vs. evening preference). This approach ensures balanced group sizes across important subgroups and reduces the risk of chance imbalances that could affect the results. Participants will be randomized individually, not in clusters, and assigned to either the treatment (exercise) or control (no change) group. The use of blocks will help maintain equal allocation as participants enroll on a rolling basis. This strategy improves internal validity by controlling for confounders and maximizing statistical power given the sample size.

Blocks

What are they, how many blocks, how many units per block?

The blocks will be defined based on two key stratification variables: gender (male, female, non-binary/other) and chronotype (morning-type, evening-type, neither). This creates a total of 6 blocks (3 gender categories $\times$ 2 chronotype categories). Within each block, participants will be randomly assigned in a 1:1 ratio to either the treatment or control group. Given a total sample size of approximately 1,000 participants, each block will include roughly 160–170 individuals, with about 80–85 participants per condition in each block. This structure ensures balanced representation across subgroups and increases the precision of subgroup analyses.

Clusters

What are they, how many clusters, how many units per cluster? If you have clusters, what is the intra-class correlation (ICC)? Is clustering strictly necessary, or could you randomize at the individual level?

This study does not use a clustered randomization design—randomization will occur at the individual level, not at the level of classrooms, dorms, or universities. Because there are no clusters, there is no intra-class correlation (ICC) to account for in the primary analysis. Clustering is not strictly necessary in this case since the intervention (daily exercise) and the outcomes (sleep metrics) are both applied and measured individually. Randomizing at the individual level maximizes statistical power and is more efficient given the structure of the study and the population being targeted.

Section Six: Analysis

Estimator

What is your estimator?

The primary estimator will be an ordinary least squares (OLS) regression model, used to estimate the average treatment effect (ATE) of the daily exercise intervention on total sleep time (TST), the primary outcome. The regression will control for key covariates including baseline TST, gender, chronotype, sleep environment quality, and alcohol/drug use to improve precision. Interaction terms will be included to explore heterogeneous treatment effects across subgroups. For robustness, both intention-to-treat (ITT) and per-protocol estimates will be reported. The ITT analysis will include all participants as randomized, regardless of compliance, while the per-protocol analysis will focus on those who met the adherence threshold for the intervention.

Standard Errors

What kind of standard errors will you use? I will use robust (heteroskedasticity-consistent) standard errors in the OLS regression to account for potential non-constant variance in the error terms across individuals. Since randomization is at the individual level and there's no clustering in the design, clustering of standard errors is not necessary. Robust standard errors will help ensure valid inference even if the residuals are not homoskedastic, which is a common concern in behavioral studies like this one.

Test

If you plan to report a p-value, what kind of test will you use? If reporting p-values, I will use two-tailed tests based on the t-distribution from the OLS regression model with robust standard errors. These tests will evaluate whether the estimated treatment effect is significantly different from zero. For secondary outcomes and subgroup analyses, multiple hypothesis testing will be addressed using methods such as the Bonferroni correction or the Benjamini-Hochberg procedure, to control the family-wise error rate or false discovery rate, respectively.

Missing Data

How will you handle missing data?

Missing data will be handled in two main ways, depending on the extent and nature of the missingness:

1. For missing covariate or survey data, we will use mean imputation or multiple imputation if the proportion of missing data is moderate and assumed to be missing at random. Indicator variables may also be included for missingness on key covariates to preserve sample size and avoid introducing bias.
2. For missing sleep tracking data, we will set a minimum threshold for inclusion—participants must have at least 5 out of 7 baseline nights and at least 10 out of 14 intervention nights of valid data. If data fall below this threshold, the participant will be excluded from the per-protocol analysis but retained in the intention-to-treat analysis where feasible.

We will also conduct sensitivity analyses to assess whether missing data may bias results and report rates and patterns of missingness transparently.

Effect size

What is the expected effect size? What is the minimum effect size that would make the study worth running? what effect sizes have similar studies found?

The expected effect size for this study is moderate, based on prior literature. Meta-analyses and experimental studies have found that regular aerobic exercise can increase total sleep time (TST) by approximately 20 to 40 minutes per night, especially in populations with poor baseline sleep. Given this, we anticipate an average treatment effect in the range of 0.3 to 0.5 standard deviations, which is considered a meaningful behavioral change in sleep research. The minimum effect size that would make the study worth running is around 0.2 standard deviations, or roughly 15 minutes of additional sleep per night. Even a small, consistent increase in sleep duration can lead to improvements in mood, cognitive performance, and academic functioning for college students, making such findings valuable for campus health interventions. Similar studies, such as Kline et al. (2013) and Kredlow et al. (2015), have found small-to-moderate effects of exercise on sleep outcomes, with effect sizes ranging from 0.2 to 0.6 SD depending on the population and intensity of the intervention. These estimates provide a strong foundation for both powering the study and interpreting its practical significance.

What is your power?

The planned sample size of 1,000 participants (approximately 500 per group) is expected to provide around 80–90% power to detect a minimum effect size of 0.2 to 0.3 standard deviations in total sleep time (TST), assuming a two-tailed test with an alpha level of 0.05. This level of power is based on estimates from prior studies of sleep interventions among young adults and accounts for some anticipated attrition or noncompliance. Final power will be confirmed using a power analysis once baseline standard deviations and likely attrition rates are more precisely known. This level of power is sufficient to detect modest but meaningful improvements in sleep outcomes, making the study both statistically and practically worthwhile.

Section Seven: Implementation

Randomization

How will you conduct the randomization? (on a computer in advance, drawing from an urn in public, etc.) Randomization will be conducted in advance using a computer-based random number generator. After participants complete the eligibility screening and consent process, they will be assigned to either the treatment or control group using block randomization, stratified by gender and chronotype. The randomization procedure will be automated and implemented through a secure data platform to ensure transparency, reproducibility, and allocation concealment. This approach minimizes human bias and ensures balanced assignment across key subgroups.

Implementation

Who will implement the intervention?

Are there any dangers to your research team, including enumerators? How will you minimize them? How will you track the quality of the implementation of the intervention? The intervention will be implemented by the research team, with support from trained research assistants. Their role will include onboarding participants, distributing wearable devices, sending daily exercise reminders, monitoring compliance via device data, and conducting follow-up check-ins. There are minimal risks to the research team, as the study does not require physical interaction beyond orientation sessions and device support. To further minimize risk, all interactions can be conducted virtually or in open, low-risk campus settings. Standard safety protocols will be followed, especially in any in-person settings. To track the quality of implementation, the team will monitor real-time data from the wearable devices, ensuring participants are completing the prescribed daily exercise. Automated systems will flag missing or inconsistent data, prompting follow-up by the research team. Additionally, logs will be kept of daily check-in responses, and random spot-checks of device data will be used to verify that activity meets the intensity and duration requirements. All procedures and deviations will be documented to ensure fidelity to the intervention design.

Compliance

Who will measure compliance?

Compliance will be measured by the research team using data from the wearable devices issued to participants. These devices will track physical activity metrics such as duration, intensity (e.g., heart rate), and timing of exercise. The team will regularly review this data to determine whether participants in the treatment group are meeting the daily 30-minute

moderate-intensity exercise requirement. In addition to device data, participants will complete brief daily check-ins (via text or app) confirming whether they exercised and describing the activity. These self-reports will be used to cross-reference the device data and help flag discrepancies or technical issues. Compliance summaries will be generated weekly to track overall adherence and guide any follow-up communication or support needed.

Data management

How will you manage the data? (security, anonymity, etc.)

Data will be managed with strict attention to security and participant privacy. All personal identifying information (e.g., names, contact info) will be stored separately from study data in an encrypted, access-restricted database. Participants will be assigned unique ID codes, and all wearable device data, survey responses, and check-in logs will be linked to these codes rather than personal information.

Data will be stored on secure, Emory-approved servers with access limited to authorized members of the research team. All transmissions of data—such as uploads from wearable devices—will be encrypted. Survey platforms will be password-protected and compliant with relevant data privacy standards (e.g., FERPA or HIPAA, depending on Emory policy). To protect anonymity in reporting, only de-identified, aggregate results will be shared in publications or presentations. Participants will be informed of these protections in the consent process, and all data will be retained only as long as necessary for research purposes before being securely deleted.