The Impact of a "Well-Timed" Nudge on Individual Physical Activity and Health Outcomes Via Wearable Devices

Keegan Brown, Elena Spielmann, Nate Spilka

May 08, 2023

1. Introduction

The prevalence of obesity and sedentary lifestyles has increased significantly in recent years, leading to numerous health problems, particularly in the United States. Studies have shown self-monitoring, via a wearable device, can modify health behaviors, including physical activity (PA) (Alturuki and Gay, 2016 and McDonough et al., 2021). Additional studies have shown that digital nudges are effective in influencing human behavior (Sobolev, 2022 and Mele et al., 2020). With wearable devices such as the Apple Watch and Oura Ring gaining popularity, and being advertised as effective tools to improve physical activity, our goal is to investigate if these devices are being used to their full potential.

This study aims to examine the effectiveness of a well-timed nudge that appears on a wearable device to promote physical activity and, as a result, improve health outcomes. Specifically, the study will focus on the impact of the nudge on resting heart rate, daily number of steps, minutes of increased daily physical activity, and Body Mass Index (BMI). We will use a randomized control trial in the student population of the State University of New York (SUNY) and City University of NewYork (CUNY) system. Specifically, we will conduct this study at three campuses (i.e., one urban, rural, and suburban) in an attempt to make the study nationally representative. There is currently little research on how a nudge via a wearable device impacts PA and health outcomes. We will randomize an intervention of a digital nudge focused on meeting a minimum standard of daily physical activity. We will study the direct impact of these interventions on baseline resting heart rate, baseline daily number of steps, baseline weekly active minutes, and baseline BMI.

2. Motivation and Conceptual Framework

2.1 Motivation

Americans continue to experience an array of obesity-related health issues such as diabetes, and heart disease (CDC, 2020), in part due to factors such as poor nutrition and low physical activity (Roglic, 2016). Health ailments negatively impact not only individuals and their families but society at large, carrying financial burdens for agents such as employers and health insurance companies (Carlson et al., 2015; Jardim et al., 2019). Over one third of American adults are obese and less than half do not meet weekly recommended levels of moderate- to

vigorous-intensity PA (MVPA) (150 minutes). With this in mind, it is a public health priority to develop low-cost interventions to increase PA levels.

2.2 Conceptual Framework

The conceptual framework for this study is based on the Transtheoretical Model of Behavioral Change, which posits that behavior change is a process that involves multiple stages, from precontemplation to maintenance. The model suggests that individuals progress through different stages of change, and that interventions may be more effective depending on the individual's stage of change (Mastellos et al., 2014).

In this study, a digital nudge delivered via a wearable device is used as the intervention to promote physical activity and improve health outcomes. The framework identifies three key constructs - self-efficacy, motivation, and outcome expectations - that are believed to influence the effectiveness of the digital nudge (Johnson et al., 2008). Reference Figure 1 below for a visualization of the conceptual framework.

We hypothesizes that the digital nudge will have the greatest impact on individuals who are in the precontemplation, contemplation or preparation stages of behavior change. These individuals may have lower levels of self-efficacy, motivation, and outcome expectations, and may benefit most from a well-timed nudge delivered via a wearable device (Forward, 2014). The aim is to increase these key constructs and ultimately lead to improved physical activity and health outcomes.

The conceptual framework provides a structured approach to understanding how the digital nudge fits into the broader context of behavior change, and highlights the importance of identifying key constructs and tailoring interventions to the individual's stage of change (Zey, 2022).

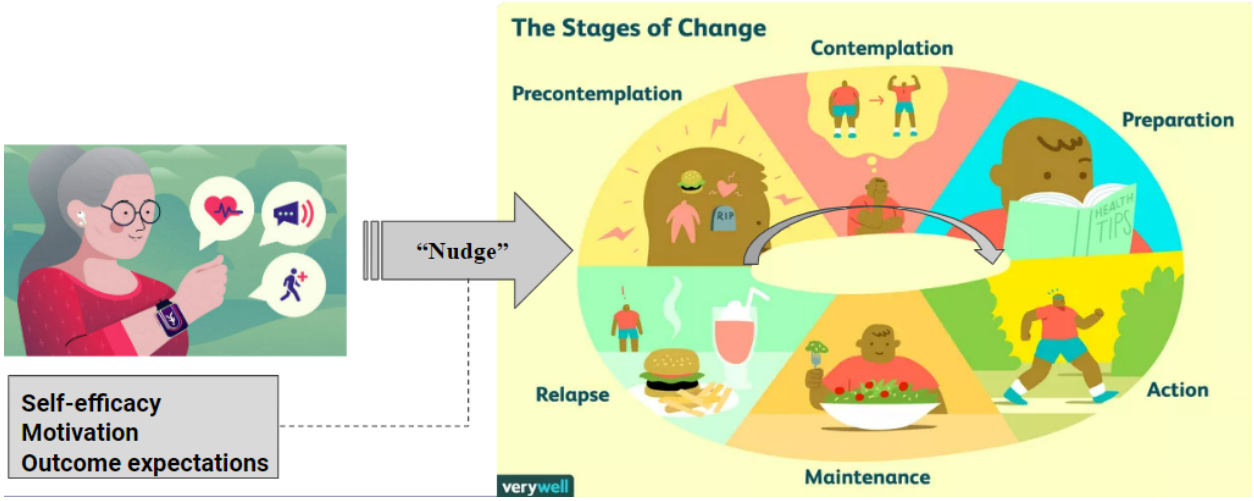

Figure 1. Conceptual Framework

3. Samples and Intervention

3.1 Samples

Our experiment will take place within the student population of the SUNY and CUNY system (i.e., one rural, urban, and suburban campus). College students are a convenient and accessible population for researchers to study. This can help to increase the generalizability of the findings, as college students are often seen as a representative sample of young adults in general. To recruit participation in our study from this population, we will place paper advertisements in common areas throughout the campuses, send digital flyers via student emails and advertise on the school website and through social media. Students will be compensated at $20/hr for travel time and while students are with a medical professional during the baseline visit and over the phone during the midline survey. Furthermore, participants will be compensated at $40/hr for the endpoint visit in an effort to minimize attrition. Compensation will be delivered at the endpoint visit, and non-compliance will result in no compensation. Additionally, students will be able to keep their wearable, or receive a cash payout for the equivalent value of the used wearable.

To be eligible to participate, students must be between the ages of 18 and 25. Limiting the age of participants has several advantages. By limiting the age range of participants, we can create a more homogeneous sample. This can help to reduce variability in the data and increase the statistical power of the study. Additionally, young adults between the ages of 18 and 25 are in a unique developmental stage, characterized by significant changes in their social, emotional, and cognitive development. Focusing on this age range can help us to better understand how behaviors and habits are formed and how they can be modified. Focusing on young adults in this age range can have significant public health implications. Theortically, this age group is at risk for developing sedentary lifestyles and poor health habits, which can have long-term consequences for their health and well-being. Finally, limiting the age range of participants to adults over 18 helps to ensure that participants are legally able to provide informed consent and make their own decisions about participating in the study (Wang et al., 2015).

A power analysis was conducted to assess our hypothesis that those in the ML condition will have a higher average daily step count compared to those in the control condition. Using a standardized mean difference of 17%, a power of 80%, and an alpha level of 0.05 (two-sided, two-sample t test), a sample size of 300 to 540 is required. This effect size and sample size range is consistent with prior literature if we consider conducting campus and demographic specific subanalyses (Wang et al., 2015). We selected a sample size of 450 to allow for attrition and such subanalyses. Finally, to give the effect size context, given that the average American walks 3,000-4,000 steps a day (Mayo Clinic, 2010), a 17% increase in the average number of daily steps would result in a 510-680 average step increase a day (i.e., roughly 2 extra miles a week). Individuals that are noncompliant with the study will be dropped from the dataset, and will not

receive compensation. Noncompliance is measured on a daily basis for all study participants. Participants not wearing the device for longer than 90% of their waking hours will be excluded from the daily sample. Participants with noncompliant days greater than 90% of the study period will be excluded from the study.

3.2 Intervention

In our experiment, we will evaluate the impacts of customized timed nudges based on a Thompson Sampling model (Russo, et al. 2018) stemming from daily data collection. Individuals assigned to this intervention will receive a notification customized to each individual's personal physical activity patterns and the time of the nudge will be optomized to induce physical activity. For example, the wearable device tracks an individual's daily activity levels, such as the number of steps taken, or the amount of time spent exercising. Based on this data, the sampling algorithm will isolate the best times to send personalized nudges to encourage the participant to engage in physical activity.

Thompson sampling is particularly useful for nudge based experiments as it balances exploiting prior data to maximize return with testing alternative inputs (e.g. timeframes) to produce positive results. In short, the sampling method takes the probability of a successful intervention at a given time, and weights the frequency of sending a nudge at a given time based on the probability of successful intervention. This functions similar to the Beyesian control rule, and is one of the strongest algoroithms to maximize the chance of treatment success (Russo, et al. 2018).

3.3 Randomization

The study will recruit a total of 450 participants, with 50% (around 225 individuals) assigned to the control group and 50% (around 225 individuals) assigned to the treatment group. The sample will be stratified by race and ethnicity, with n=150 individuals per campus (urban, suburban, and rural). The random assignment of participants will be double-blind, meaning neither the participants nor the researchers will know which group each participant is in. Stratified random sampling will help account for endogenous factors and improve the precision of the research, making the results more generalizable to the population (Vajravelu and Arslanian, 2021). The sample size is supported by a power analysis and previous studies focused on digital nudging via text (Friedenreich et al 2010; Napolitano, et al., 2010; Rogers et al., 2009; Wang et al., 2015).

Reference Figure 2 below for a visualization of the study randomization. Following the two-week baseline assessment, participants in this study will be randomly assigned at the individual level to either a treatment group receiving an ML nudge intervention or a control group with no nudge intervention. Stratified random sampling will be used to ensure proportional representation of different racial and ethnic groups based on the most recent U.S. Census Bureau estimates. The latest Census Bureau estimates are that the population was: 60.1 percent non-Hispanic White; 18.5 percent Hispanic; 12.5 percent non-Hispanic Black; 5.8 percent

non-Hispanic Asian; 2.2 percent non-Hispanic of two or more races; and .9 percent Native Peoples (U.S. Census Bureau, 2022).

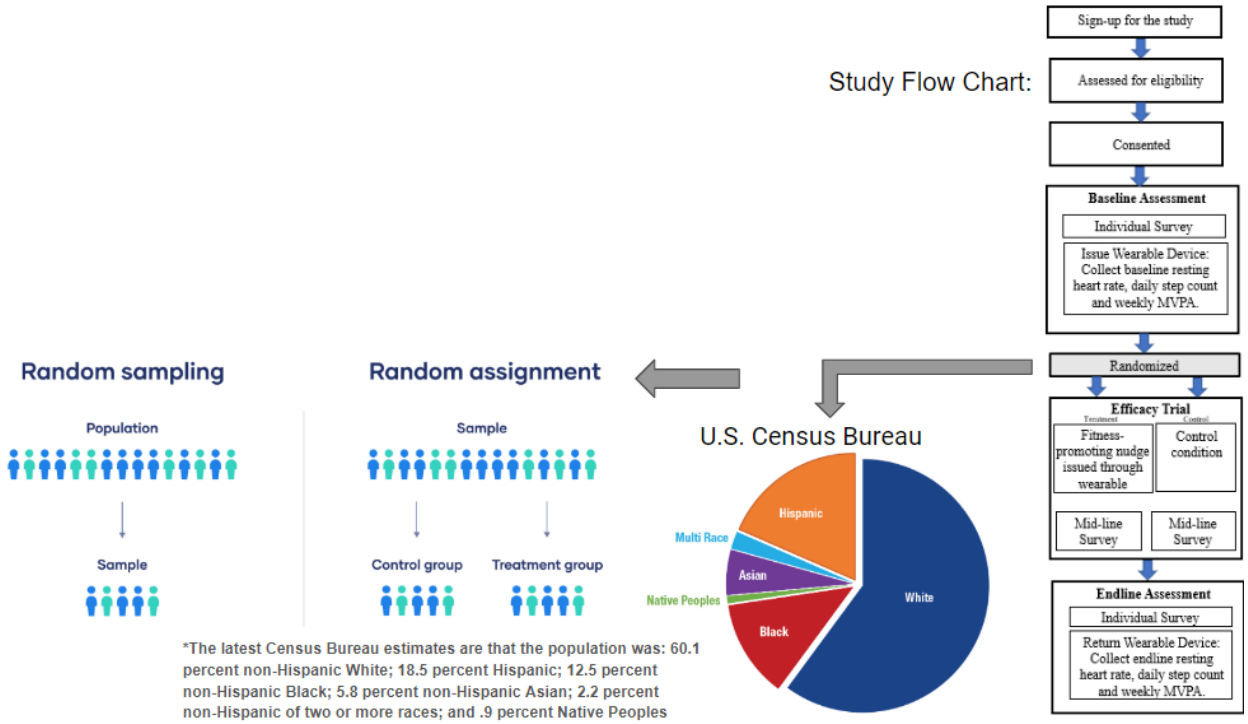

Figure 2. Study Randomization

3.4 Risks

The risks associated with this study are minimal, as the intervention is a well-timed nudge on a wearable device. Our project will be reviewed by the American Public Health Association (APHA), the Association for Behavior Analysis International (ABAI), and the National Institutes of Health (NIH). The data collection involves no more risk than a participant is already used to as a regular wearable user (i.e., potential discomfort or irritation from prolonged periods of wearing the device). Some participants will be randomly selected to receive a fitness-focused nudge from their wearable device, but all participants will wear the wearable device.

One concern is that the individual may potentially sell the device that they are issued. This will be mitigated through each participant signing a contract that holds them financially liable for the full cost of the wearable device should they fail to complete the exit procedure for the study. Further, compensation will be tied to completion of the exit survey as well. Only data from those who were issued and returned the device at the endline assessment will be used in the study.

Additional concerns include biases associated with voluntary participation. Self-selection bias is where individuals who choose to participate in the study may have different characteristics or motivations than those who do not. This can lead to biased results, where the study participants are not representative of the larger population (Tang et al., 2020). Response bias may also be an issue. Even among those who choose to participate, some may be more likely to respond to certain questions or in certain ways than others. This can lead to inaccurate or incomplete data (Jin et al., 2020).

These bias limitations will be mitigated in several ways. First, our study will have a large sample size (n=450). This can help to increase the representativeness of the sample and improve the statistical power of the study. Second, we will conduct stratified random sampling in an effort to construct similar control and treatment groups, mitigating sources of bias that may arise from inter-group characteristics. Finally, we will use multiple recruitment methods to reach a diverse group of participants. For example, we will recruit participants from different locations, use social media and online platforms to reach a wider audience, and partner with community and campus organizations to increase participation among underrepresented groups.

All adults will consent to participate via written agreement outlining all relevant study details prior to the start of the research. All data will be anonymized to avoid potential risks of data privacy breaches. Individuals will also be free to withdraw from the study at any point or not to answer any questions during the interviews.

3.5 Data Sources

We plan to collect data from individuals continuously over the academic year (i.e., a 9-month period) and will administer surveys at three points in time: at baseline (upon issuance of the wearable device), four months after baseline (midline), and 9 months after baseline upon return of the wearable device (endline), which coincides with the end of the academic year. The continuous data will be passively collected and obtained through the wearable. This data will include resting heart rate, daily number of steps, and weekly MVPA. Additionally, BMI will be recorded at baseline and endline visits. These data will help determine if there are significant changes in health outcomes. The baseline survey will collect information about the individual's previous wearable use, eating habits, exercise habits, sleeping habits, alcohol and tobacco use, stress levels, and demographic information. The midline survey, six to seven months after baseline, will be a shorter survey focused on any changes (improvements or deteriorations) to baseline the individual noticed. The endline will measure the same outcomes as baseline. The continuous data collected over the 9 month period will be plotted over time and the baseline data from the first month will also be compared to the endline data from the 9th month to determine if there was significant change. The baseline and endline surveys will be face-to-face, tablet-based interviews. The midline will be administered over the phone.

3.6 Timeline

The study will take place over a 9 month period, with participants wearing the device throughout the study period. Data will be collected continuously throughout the study, with three surveys supporting measured data from the wearable. The final analysis will be conducted at the end of the study.

Baseline surveys of participants (Start of survey)

- Baseline survey will be administered
- Health measures (BMI, Resting HR, etc.) will be taken by medical staff
- Control and treatment groups will be automatically assigned by a computer program to maintain a double blind design
- Wearables will be administered

Implementation of intervention (Two weeks)

- After two weeks of collecting data on activity and health measures individuals in the treatment group will begin recieving treatment messages

Midline survey of participants (Month 4)

- Survey will be administered
- Schedule updates will be recorded by staff and plugged into administration algorithm to ensure nudges are not being sent during the new class schedule
- Midpoint compensation administered

Endline survey of participants (Month 9)

- Full survey and medical evaluation administered
- Wearables are returned and full compensation is provided

4. Outcomes and Hypotheses

The wearable device, baseline and endpoint assessments, and survey will be used to measure the impact of wearables and the nudge intervention on participant health. Specifically, four primary outcomes will be measured to understand the effect of the two conditions on participant behavior:

a. The degree to which one engages in MVPA.
b. Resting heart rate.
c. Average daily number of steps.
d. BMI.

Each of these dimensions in addition to relevant control variables are discussed in greater detail below.

Wearables:

- Wearables assigned to both the ML and control groups will record the following three variables: (a) Moderate to Vigorous Physical Activity (MVPA), (b) Resting Heart Rate, and (c) Average daily number of steps. These real-time measures will be recorded every day for each participant.
    - Moderate to Vigorous Physical Activity (MVPA): This will be a continuous variable indicating the number of minutes a participant engages in moderate to vigorous physical activity each day. "Moderate" and "vigorous" intensity will be measured at the following thresholds suggested by the CDC: a heart rate between 64%-76% and 77%-93% of one's maximum heart rate, respectively (CDC, 2011). Maximum age-related heart rate is measured by subtracting one's age from 220. Participants in the ML condition will receive messages encouraging them to engage in MVPA at optimal moments throughout each day. When data are analyzed, the amount of time following each message that a participant engages in MVPA (if they engage in MVPA) will be studied to understand the efficacy of the intervention. The control group will receive no such messaging.
    - Resting Heart rate: This will be measured as a continuous variable indicating a participant's average daily resting heart rate. Generally, a resting heart rate within a healthy range (60-100 bpm) suggests that one's heart is more efficiently pumping blood around the body. When comparing groups, the percentage of participants within the healthy range during baseline/endpoint will be assesed.
    - Steps: The average daily number of steps will be recorded as another method to understand a participant's engagement with PA (in addition to studying MVPA). This will be recorded as a continuous variable indicating the average number of steps a participant takes on a given day. Participants in the ML condition will be prompted with well-timed messages, encouraging them to engage in lower levels of activity throughout each day (in addition to the higher levels of PA mentioned above). For example, they will be encouraged to take the stairs instead of the elevator whenever possible, or to walk to class instead of using a motorized scooter. The control group will receive no such messaging.

Baseline and endpoint measures:

- Participant vitals (pulse rate and blood pressure), demographics, and clinical characteristics (diabetes, hypertension, dyslipidemia, heart issues, and substance usage), will be recorded by a medical professional during baseline and endpoint. These measures will be used to better understand the potential impact of the wearables through cross

validating variables, subgroup analyses, and while controlling for various relevant characteristics.

- **Vitals**:
  - Heart rate: This will be a continuous variable indicating a participant's heart rate. This will be cross validated with the participant's wearable to ensure the device is operating correctly. Additionally, this measure will be used as part of the primary outcome variables.
  - Blood pressure: this will be measured as a continuous variable and used to understand how one's blood pressure changes over the course of the study. Additionally, this variable will be used as a control variable in the final analysis, as it is a potential risk factor for various relevant health conditions (e.g., hypotension/hypertension and Addison's disease).
- **Medical characteristics:**
  - BMI: This will be a continuous variable calculated using a participant's height and weight. This variable will be used as a primary outcome variable with attention to changes between baseline and endpoint as well as what the CDC considers to be a healthy range (18.5-24.9) (CDC, 2022).
  - Sex (male/female): A dichotomous variable indicating the sex of the participant. This will help us both control for endogeneity related to sex as well as conduct sub-analyses.
  - Age: This will be measured as a continuous variable indicating the participant's age.
  - Health conditions: This will be a set of dichotomous variables for each health condition, indicating whether a participant experiences any relevant health conditions (e.g., diabetes, hypertension, dyslipidemia, heart issues). Health conditions will be controlled for in the analysis of this study as they may bias the estimates of each condition on the primary outcome measures of interest.
  - Relevant substance use: This will be a set of dichotomous variables for specific substances used by participants (e.g., cigarettes and vapes). Controlling for participant substance use will help us address endogeneity concerns and better isolate the impact of wearables and nudges on various health outcomes. For example, given that nicotine increases blood pressure and heart rate, a participant who quits smoking cigarettes mid-study may bias results such that the intervention (whether that be the control or ML condition) may appear to be more effective than in reality.
  - Race/ethnicity: This will be a categorical variable indicating a participant's race/ethnicity. This will help better understand how wearables may affect individuals of different backgrounds differently, motivating important future research.

Survey measures:

- The survey, which will be administered at three times throughout the study, will be used to (a) help identify mechanisms for behavioral and physiological outcomes, and (b) control for factors in our analysis that may bias our results. Many of the questions below will ask about one's experiences in the current week compared to the week prior. This approach aims to understand general trends in one's thoughts, behaviors, and experiences, while minimizing day-to-day variability across participants.
    - Question: "Overall, do you feel that the fitness-related messaging is encouraging you to engage in physical activity?"
        - This question will be measured as a dichotomous variable and help us understand the reception of the messaging (i.e., better isolate whether it is a mechanism for changes in exercise rates). Only participants in the treatment condition will receive this message.
    - The following questions explore the effectiveness of the digital nudge related to the Transtheoretical Model:
        - Precontemplation: Do you feel that increasing physical activity is important for your health? (Yes/No)
        - Contemplation: Have you considered increasing your physical activity in the next few weeks? (Yes/No)
        - Preparation: Are you actively making plans to increase your physical activity in the next week? (Yes/No)
        - Action: Have you increased your physical activity since receiving the digital nudge? (Yes/No)
        - Maintenance: How long have you sustained an increase in physical activity since receiving the digital nudge? (Less than a week/1-3 weeks/1-3 months/More than 3 months)
        - Self-efficacy: How confident are you that you can increase your physical activity level? (Not at all confident/Slightly confident/Moderately confident/Very confident)
        - Motivation: How motivated are you to increase your physical activity level? (Not at all motivated/Slightly motivated/Moderately motivated/Very motivated)
        - Outcome expectations: Do you think increasing your physical activity level will have a positive impact on your health? (Strongly disagree/Disagree/Agree/Strongly agree)
    - The following questions will be measured as categorical variables and used as controls to help isolate the effect of one's PA on their health outcomes (e.g., BMI).

        - "Would you say you're eating habits are currently "healthy?"" (rate on a scale of 0-5 with zero being extremely unhealthy and 5 being extremely healthy).
        - "Would you say your eating habits this week have been healthier, less healthy, or the same degree of healthy compared to last week?"
        - "In the past week, on average, how many hours of sleep a night do you get?"
        - "Would you say your sleeping habits are healthier (e.g., 8-hours a night), less healthy, or the same degree of healthy compared to last week?"
        - "In the past week, how many days have you consumed alcoholic beverages?"
        - "On average, how many drinks do you consume in one sitting?"
        - "Would you say you're drinking more, less, or the same amount of alcohol this week compared to last?"
        - "Based on last week, have you smoked any cigarettes or used electronic smoking devices?"
        - "Would you say you smoked more, less, or the same number of cigarettes this week compared to last?"
        - "How would you rate your current stress level?" (rate on a scale of 0-5 with zero being completely unstressed and 5 being extremely stressed).
        - "How stressed are you this week compared to last? More, less, or the same?"

Primary hypotheses:

- The primary hypotheses in the current study compare outcomes between the ML and control conditions of the intervention. Specifically, given the outcome measures described above, we hypothesize that compared to the control, the ML group will have a:
    - Greater increase in the number of minutes spent engaging in MVPA per week compared to baseline.
    - Greater percentage of participants that migrated to the healthy heart rate range.
    - Greater increase in the number of daily steps compared to baseline.
    - Increased percentage of participant in the healthy BMI range compared to baseline.

5. Estimation Methodology

5.1 Measurement and Model Design

The study will use a multi-arm experimental design, with single-blind assignment to treatment and control groups, with surveyors blind to treatment allocation. Broadly, the sample will be divided into two categories:

- 50% of the sample will be assigned to the control group, and will not receive nudges from the wearable.
- 50% of the sample will be assigned to the treatment group, where participants will receive nudges based on a Thompson Sampling model.

Stratified sampling will account for demographic representations, with an effort to align race, gender, and ethnicity variables in a representative manner.

The primary analysis of the data will be conducted through multiple fixed-effect regression models leveraging collected activity and health data as the dependent variable. Multiple regressions for the study are critical to determining the effectiveness of the intervention because measured outcomes (e.g., MPVA, resting heart rate, step count) provide sufficiently valuable health information independently. For example, if treatment is associated with an increased step-count but no change in MPVA or resting heart rate, it would suggest that the nudges were effective in inducing only low-difficulty exercise. A two-way (time and entity) fixed effect model is the best method for eliminating omitted-variable bias as it can control for time-varying factors associated with the messaging and our various outcome measures, and unchanging variables (e.g. demographics). Interaction terms will be used to evaluate subgroup analyses. The following model is used to estimate the effect of the nudge:

$$Y_{it} = \beta_0 + \beta_1 T_{it} + \alpha_i + \gamma_t + \eta_{it} + \epsilon_{it}$$

where *i* represents a participant index, *t* represents a time index, Y represents the different outcome measures detailed above, T represents whether one is in the treatment group, α represents entity fixed effects, γ represents year fixed effects, η represents all control variables, and ϵ is the error term.

Controls for substance use, race, gender, prior health status, age and prior wearable use will be included in the model.

5.2 Treatment Effect Heterogeneity

Beyond standard demographic subanalysis, there are a few sources of treatment effect heterogeneity that will be accounted for in the interpretation of the study results as well as the study design.

1. Prior use of wearable: We will measure the effect of the prior use of the wearable by comparing the behaviors of prior wearable users across the control and the treatment

group. Our standard analysis concerning the efficacy of experimental nudges will be used to evaluate differential effects that may be induced by prior exposure to wearable devices. The experimental design will stratify those that used a wearable device prior to the experiment, and the entry survey will collect this data so it can be controlled for in our final analysis.

2. Asymmetric free time: The frequency of regular exercise may be impacted by both educational course load and extracurricular activities. In addition, the response to individual nudges is likely strongly influenced by environmental factors. For example, treatment participants will not be able to respond to a nudge if they are required to be seated for their work. While the ML model solves for some of this variability by conforming to student's behavioral patterns to optimize nudge efficacy, reports of hours of work and school schedule (e.g. number of credits and the area of study) could be taken to account for differences in the expected activity level. Entry, midpoint, and exit surveys will feature questions on available time to control for this effect.
3. Gym accessibility: The proximity of workout facilities to participant's place of residence will be considered when evaluating treatment efficacy. Nudges sent to an individual with a gym in their apartment complex will likely relieve more positive responses than those sent to a participant that must travel several miles to engage in exercise. Entry, midpoint, and exit surveys will feature questions on accessibility so this effect can be controlled for in our results.
4. Weather: Finally, weather conditions will be considered when evaluating the efficacy of nudges. Nudges ignored after a significant snowstorm, for example, may more accurately be attributed to weather conditions than the effectiveness of the nudge. Weather status will be retrieved from the location of the wearable at the time of the nudge, and at random intervals for the control group. Days that are below freezing or feature rain or snow will be coded as such. Weather data will be pulled from the national weather service for the GPS location of the wearable at the time of any nudge. Imputed data from the nudged weather analysis will be used for the control group when applicable. For deeper analysis on the impact of weather, collected data can be aggregated across participants in sufficiently similar locations over time.

5.3 Threats

Attrition: Attrition is an inherent risk with any longitudinal study. Participants benefit from the wearable itself, which helps mitigate the chance of attrition. Additionally, the compensation plan described in the samples section (3.1) should further reduce attrition risk. Regular follow ups with participants will help reduce attrition as well as catch cases where it has occurred at the midpoint and endpoint of the study. Research principles will take every effort to work with participants' schedules to limit cases of attrition due to midpoint survey non-compliance whenever possible.

Resale and Transfer: Resale of an electronic device used in any study is an inherent risk. While the possibility of resale cannot be eliminated, researchers will inform study participants that the wearables are equipped with GPS tracking and that the exit survey must be conducted at the end of the testing period to receive test compensation. Combined, these safeguards should reduce the possibility that a study participant will sell the device. Transfer of the device is more plausible. Regular monitoring of location and health data can be used to determine participant compliance. Suspected non-compliant participants will be flagged for review by the research principals, and sufficient evidence supporting noncompliance will cause the participant to be dropped from the study.

Nonstandard Use: Use of the wearable in nonstandard manners may compromise passively collected data. Resting heart rate, for example, is often calculated as an individual wakes up or prepares to sleep. If a study participant leaves the wearable in the charger overnight, these measurements may include elevated activity from the individual moving to place the wearable on their wrist. Furthermore, if individuals are charging their device at different times during the day, wearables may fail to collect critical activity data. This threat to treatment data is difficult to control for, as wearables require sufficient battery life to collect comprehensive health data. Participants will be provided with 3 regularly scheduled time slots to recharge the wearable on a given day, but every effort to improve battery life and reduce the frequency of charging disruptions will be undertaken when selecting the wearable for study use.

Self-Selection Bias: Introduction into the sample requires students to respond to advertisements for the study. As a result, self-selection bias is an inherent concern for the study. Because we cannot require individuals to engage with the study, this bias should be considered when interpreting study results.

References

Alturki, Ryan M., and Valerie Gay. “A Systematic Review on What Features Should Be Supported by Fitness Apps and Wearables to Help Users Overcome Obesity.” OPUS at UTS | Open Publications of UTS Scholars. International Journal of Research in Engineering and Technology, September 25, 2016. https://opus.lib.uts.edu.au/handle/10453/119606.

Carlson, S. A., Fulton, J. E., Pratt, M., Yang, Z., & Adams, E. K. (2015). Inadequate physical activity and health care expenditures in the United States. Progress in cardiovascular diseases, 57(4), 315-323.

CDC (Center for Disease Control and Prevention) (2020) National Diabetes Statistics Report 2020. U.S. Department of Health and Human Services, Atlanta.

CDC (Centers for Disease Control and Prevention). (2011). Target heart rate and estimated maximum heart rate. Retrieved from: http://www.cdc.gov/physicalactivity/everyone/measuring/heartrate.html

CDC (Centers for Disease Control and Prevention). (2022). Assessing Your Weight. Retrived from: https://www.cdc.gov/healthyweight/assessing/index.html

Forward, S.E. (2014). "Exploring people's willingness to bike using a combination of the theory of planned behavioural and the transtheoretical model". European Review of Applied Psychology. 64 (3): 151–159. doi:10.1016/j.erap.2014.04.00

Friedenreich C, Woolcott C, McTiernan A, et al. Adiposity changes after a 1-year aerobic exercise intervention among postmenopausal women: A randomized controlled trial. Int J Obes 2010;35:427–435.

Johnson, S. S.; Paiva, A. L.; Cummins, C. O.; Johnson, J. L.; Dyment, S. J.; Wright, J. A.; Prochaska, J. O.; Prochaska, J. M.; Sherman, K. (2008). "Transtheoretical model-based multiple behavior intervention for weight management: Effectiveness on a population basis". Preventive Medicine. 46 (3): 238–246. doi:10.1016/j.ypmed.2007.09.010. PMC 2327253. PMID 18055007

Jin, Daoyan, Hallegeir Halvari, Natalia Maehle, and Anja H. Olafsen. “Self-Tracking Behaviour in Physical Activity: a Systematic Review of Drivers and Outcomes of Fitness Tracking.” Behavior & Information Technology. Taylor & Francis Online, August 5, 2020.

Mastellos, Nikolaos; Gunn, Laura H.; Felix, Lambert M.; Car, Josip; Majeed, Azeem (2014-02-05). "Transtheoretical model stages of change for dietary and physical exercise modification in weight loss management for overweight and obese adults". Cochrane Database of Systematic Reviews (2): CD008066. doi:10.1002/14651858.cd008066.pub3. ISSN 1465-1858. PMID 24500864

Mayo Clinic (2010). 10,000 steps a day: Too low? Too high?. Retrieval from: https://www.mayoclinic.org/healthy-lifestyle/fitness/in-depth/10000-steps/art-20317391#:~:text=The%20average%20American%20walks%203%2C000,a%20day%20every%20two%20weeks.

McDonough, Daniel J, Xiwen Su, and Zan Gao. "Health Wearable Devices for Weight and BMI Reduction in Individuals with Overweight/Obesity and Chronic Comorbidities: Systematic Review and Network Meta-Analysis." British Journal of Sports Medicine. BMJ Publishing Group Ltd and British Association of Sport and Exercise Medicine, August 1, 2021. https://bjsm.bmj.com/content/55/16/917.

Mele, Christina, Tiziana Russo Spenna, Valtteri Kaartemo, and Maria Luisa Marzullo. "Smart Nudging: How Cognitive Technologies Enable Choice Architectures for Value Co-Creation." Journal of Business Research. Science Direct, September 21, 2020. https://www.sciencedirect.com/science/article/pii/S0148296320305786.

Napolitano MA, Borradaile KE, Lewis BA, et al. Accelerometer use in a physical activity intervention trial. Contemp Clin Trials 2010;31:514–523.

Rogers LQ, Hopkins-Price P, Vicari S, et al. Physical activity and health outcomes three months after completing a physical activity behavior change intervention: Persistent and delayed effects. Cancer Epidemiol Biomarkers Prev 2009;18:1410–1418.

Roglic, G. (2016). WHO Global report on diabetes: A summary. International Journal of Noncommunicable Diseases, 1(1), 3.

Russo, D. J., Van Roy, B., Kazerouni, A., Osband, I., & Wen, Z. (2018). A tutorial on Thompson sampling. *Foundations and Trends® in Machine Learning*, *11*(1), 1-96.

Sobolev, M. (2022). Digital Nudging: Using Technology to Nudge for Good. In D. Soman & N. Mazar (Ed.). *Behavioral Science in the Wild* (pp. 292-299). Toronto, Canada: University of Toronto Press.

Tang, Matilda Swee Sun, Katherine Moore, Andrew McGavigan, Robyn A Clark, Anand N Ganesan1, College of Medicine and Public Health, and Corresponding Author:Anand N Ganesan. "Effectiveness of Wearable Trackers on Physical Activity in Healthy Adults: Systematic Review and Meta-Analysis of Randomized Controlled Trials." JMIR mHealth and uHealth. JMIR Publications Inc., Toronto, Canada, July 2020. https://mhealth.jmir.org/2020/7/e15576.

U.S. Census Bureau. (2022). QuickFacts. Retrieved from https://www.census.gov/quickfacts/fact/table/US/PST045219

Vajravelu, Mary Ellen, and Silva Arslanian. "Mobile Health and Telehealth Interventions to Increase Physical Activity in Adolescents with Obesity: A Promising Approach to Engaging a

Hard-to-Reach Population - Current Obesity Reports." SpringerLink. Springer US, October 1, 2021. https://link.springer.com/article/10.1007/s13679-021-00456-8.

Wang, Julie B., Lisa A. Cadmus-Bertram, Loki Natarajan, Martha M. White, Hala Madanat, Jeanne F. Nichols, Guadalupe X. Ayala, and John P. Pierce. "Wearable Sensor/Device (Fitbit One) and SMS Text-Messaging Prompts to Increase Physical Activity in Overweight and Obese Adults: A Randomized Controlled Trial." *Telemedicine and e-Health* 21, no. 10 (2015): 782–92. https://doi.org/10.1089/tmj.2014.0176.

Zey, Ann-Sophie. "The Use of Digital Nudges on Wearable Devices for Improving Physical Activity." University of Twente, April 24, 2022. https://essay.utwente.nl/90510/1/Zey_MA_BMS.pdf.