Digital and Physical Lifestyles: Exploring Smartphone Use and Physical Activity's Impact on Young Adults' Mental Health

*, *, *, Utrecht University

Abstract In recent years, excessive smartphone use and physical inactivity among young adults have raised concerns about their impact on mental health. This study explores the relationship between screen time, physical activity, and their effects on stress and depression. Using objective measures and self-reported data from the K-emoPhone dataset, we employed fixed effects and multiple linear regression models to analyze these interactions. Contrasting to most research, findings indicate that higher screen time is significantly correlated with lower stress levels across participants. Though, it shows no within-person effects over time. Calorie expenditure consistently correlates with reduced stress, underscoring the importance of physical activity. However, the categorical physical activity feature lacked significance, suggesting the need for refined measures. The analysis of depression data revealed invalid model findings, leading to its removal from the process. The study highlights the complexities of these relationships and suggests future research to incorporate more objective mental health measures, enhanced data collection methods, and broader variable exploration. These insights are crucial for developing strategies to better understand the effects of digital and inactive lifestyles on mental health.

1. Introduction

In today's digital era, smartphones have become central to daily life, holding the power to impact well-being when used excessively. While smartphones offer vast connectivity, concerns are growing about their potential addictive use and the subsequent mental health implications, like stress and depressive symptoms [1]. Simultaneously, there has been a noticeable decline in activity levels, possibly due to increased digitization among young adults over the past six years, which can lead to mild or serious health problems [2]. This research proposal seeks to explore the connection between smartphone screen time and physical activity, and their effects on stress and depression among young adults. By examining these variables, this study aims to address a gap in the literature: the dual influence of digital and physical activities on psychological well-being.

Research has explored the individual and combined effects of screen time and physical activity on mental health. Numerous studies indicate that excessive screen time correlates with negative mental health outcomes, including stress and depression, particularly due to its association with sedentary behavior and addictive usage patterns [3], [4]. Concurrently, physical activity is universally acknowledged for its positive impact on mental health, acting as a protective factor against stress and depression [5], [6]. Despite this understanding, the intersection of these two factors—how they might counterbalance or exacerbate each other—remains underexplored in natural settings and daily life of young adults.

The limitation of prior research primarily lies in its reliance on self-reported data, which can be subjective and prone to biases such as mood-congruent memory or rosy retrospection [7]. This proposal aims to address these gaps by utilizing objective measures of screen time and physical activity, collected for 7 days respectively, to provide a more accurate and reliable analysis.

The dataset at our disposal offers a highly detailed recording of screen time (ST), physical activity (PA), self-reported stress (SS), and depression scores (DS) in a highly naturalistic setting. Most importantly, this research hopes to capture the complex dynamics of the interactions between these variables, and K-EmoPhone's temporal dimension provides an unique opportunity for this. Additionally, the datasets' elaborate recordings of biological markers allows to bypass certain limitations seen in previous studies, often relying heavily on self-reported data of screen time and physical activity. Moreover, by leveraging real-world, objective data combined with self-reported measures, we hope to provide a clearer picture of the impact of screen time and physical activity on mental health, and ultimately closing a gap in understanding.

To conclude the introduction, the upcoming sections will delve into the related work section, reviewing previous studies pertinent to each aspect of our research question. The methodology used for data collection and preprocessing, detailing the objective measures and statistical models employed. Following this, the results section will present findings on the relationship between screen time, physical activity, and mental health outcomes, with a focus on stress and depression. We will discuss the implications of these findings, highlighting both significant and non-significant effects, and conclude with the limitations of the study and directions for future research. This structure ensures a comprehensive understanding of the complex interactions between digital, psychological habits well-being of young adults.

2. Related work

Device usage is typically denoted by screen time, a term referring to the amount of time spent and the diverse activities performed online using digital devices [8]. Excessive screen time, defined as daily screen time of more than 2 hours outside of school or work [9], has been associated with various negative effects. Previous studies have focused on the effects screen time could have on sleep, academic performance, mental health, social development, and physical activity [1], [8], [10], [11]. A substantial amount of the research however, has relied on the use of self-report measures. One could question whether participants are able to accurately assess their own behaviour. There is a significant inconsistency between self-report measures and objective measures such as screen tracking [11]–[15]. To address these limitations of previous studies, this study will be using objective data on screen time obtained through software installed on participants' smartphones [16].

2.1 Screen time and physiological stress

In the context of stress and excessive screen time, studies tend to find a strong link. For example, higher self-reported mobile phone use is associated with current stress among young adults [17]–[19]. Stress is a physiological and psychological response to challenges or demands, characterized by a state of heightened arousal and tension [20]. Disturbances in sleep, avoidance of healthy coping and constant availability are examples of how excessive screen time can contribute to elevated stress levels [17]. Additionally, extensive phone use is associated with its dependency, which in turn can lead to more stress [21].

Not only does excessive screen time likely contribute to stress, but young adults may also engage in overindulgence, seeking sedation or emotional support as a response to stress and unpleasant feelings [3], [4], underlining the importance of further investigation of the relationships between screen time and mental health.

Though stress has clear biological markers, such as cortisol levels, much research on stress relies heavily on self-reported measurements [22]. Though this is not inherently problematic, emotional and retrospective biases can be mitigated by frequently assessing stress levels. One example for this is the Experience Sampling Method (ESM)[23], [24]. ESM allows for real-time data on participants' mental state in their natural environment [25]. Often, self-report data is combined with automated sensor data from participants' smartphones and wearables, as this has been proven to be valuable and capable of providing previously unattainable insights [26]. It should be noted that this type of self-reported data differs significantly from the aforementioned self-reported screen time data, which is typically not collected in real-time. Recent studies that have utilized this state-of-the-art method have been able to uncover nuanced patterns of people in their natural environment [27], [28]. However, most previous research that has combined self-reports with sensor data has either relied on pre-/post-study surveys, or daily surveys [28]–[30].

2.2 Stress, depression and screen time

Numerous studies have demonstrated a clear association between clinical depression and negative mood with elevated cortisol levels, with stress often being a significant contributing factor to the development or exacerbation of depressive symptoms [31]–[33]. Closely related to the current study, a study by Zhu et al (2021) found (academic) stress to be positively associated with anxiety and depression among Chinese adolescents [6]. Depression, a complex and multifaceted condition within psychiatry and clinical psychology, includes a range of symptoms from temporary feelings akin to sadness to more severe and prolonged physical, psychological, and behavioral changes. While there is no universally accepted definition, depression is recognized as a form of mental illness that often includes feelings of loneliness and an array of negative emotions [34].

The measurement of depression can be complex and over the past decades, a variety of available tools have been developed and utilized for screening and diagnosis [7].

In effort of defining depression, The American Psychiatric Association's DSM-5 [35] outlines criteria for diagnosing major depressive disorder. This includes experiencing at least five of a symptom list, over a two-week period, which must represent a change from previous functioning. The symptoms must include either a depressed mood or a loss of interest or pleasure [36]. Among multiple tests, the Patient Health Questionnaire (PHQ-9) is a widely used tool that assesses the severity of depression through nine questions based on the DSM-IV criteria, each scored from 0 (not at all) to 3 (nearly every day). In the current study, a Korean version of PHQ-9 has been used as a measure for depression, which is shown to be reliable and valid for screening depressive symptoms [37].

The directionality of the relationship between depressive symptoms, stress and screen time is complex, though screen time seems to predict future depressive symptoms more consistently than depressive symptoms predict future screen time [38]. Supporting this, a research by Anderl et al, (2023) found participants who spent more time on their smartphones in the hour preceding an assessment reported lower psychological well-being [39]. A longitudinal study by Hobky et al, 2023 found that hourly digital screen time increases adolescents' depressive symptoms and emotional disregulation [38]. In contrast, Orben and Przybylski (2019) did not find such significant effects. Additionally, associations between screen time and depressive symptoms generally have small effect sizes [38].

2.3 Physical activity

Physical inactivity has been identified as a growing health concern globally, contributing to various health problems [2]. The physiological benefits of physical activity hold a wide range of body systems, substantially impacting energy balance and body composition [40]. Physical activity can be regarded as any bodily movement produced by skeletal muscles that results in caloric expenditure, highlighting the energy aspect inherent to these movements [41]. Illustrating this, Zhu et al. (2021) found a negative correlation between academic stress and physical activity [6]. Similarly, another study among healthy adults found that those who engaged in regular exercise reported experiencing fewer negative emotional consequences of stress [5]. Such findings are widely supported by research [18], [42]–[44]. Additionally, studies found that screen time was moderated by healthy habits like physical activity and sufficient sleep, might have less of a negative impact on stress levels. This suggests that the relationship between screen time and stress can be mitigated by other healthy behaviors [45]. In a similar vein, Wu et al. (2015) highlighted that low physical activity (PA) coupled with high screen time correlated with heightened risks of mental health problems.

Several methods exist to measure physical activity, each with its strengths and considerations. Self-reported questionnaires, activity logs, and direct observation provide subjective data, while devices like pedometers, heart rate monitors, and accelerometer-based systems offer more objective measures [46].

The measure of calories are discussed in more detail because caloric expenditure is intricately linked to physical activity and is crucial for understanding individual energy requirements. Research suggests that the daily calorie intake is an indicative measure of physical activity levels. Observations indicate that more physically active groups, identified through leisure and occupational activity measures, generally exhibit higher caloric intake. The underlying principle is based on the energy balance equation (energy balance = energy intake + energy expenditure), implying that physical activity can be inferred from energy intake, assuming no change in energy balance occurs [47].

3. Research question

While there are notable associations between screen time, stress and depressive symptoms, the causal pathways are complex and influenced by multiple factors. For example, healthier lifestyle choices like regular physical activity, suggest that managing screen time along with other aspects of health could mitigate its negative impacts on stress. By examining these interactions, we can better understand the potential mental health implications of digital and physical habits. Therefore, we aim to investigate the following research question:

How is the amount of smartphone screen time and amount of physical activity (measured by burned calories and mobile measured activity events) related to prolonged self-reported stress and depression scores among young adults?

4. Methodology

4.1 Data

The current research utilizes the K-EmoPhone dataset, a comprehensive and recently gathered dataset by Soowon Kang et al (2023) [16]. In their study, 77 Korean students participated in a 7 day study. All collected data broadly falls into the following three categories 1.) Demographic and Psychological Profiles, 2.) In-Situ Affective and Cognitive States and 3.) Multimodal Sensor Data. The demographic and psychological profiles were measured preceding the data tracking process. This data includes basic demographic information (age, gender) and psychological traits assessed via the Big Five Inventory (openness, conscientiousness, neuroticism, extraversion, agreeableness). Mental health scales such as the Perceived Stress Scale (PSS), Patient Health Questionnaire (PHQ) for depression, and General Health Questionnaire (GHQ) for psychiatric disorders are included, providing a baseline and post-study snapshot of mental health. The In-Situ Affective and Cognitive States utilize the Experience Sampling Method (ESM), the dataset captures real-time emotions, stress levels, attention, and task disturbance through mobile app prompts, allowing up to 16 self-reports per day from each participant. Variables such as emotional valence, arousal, attention levels, and perceived stress are reported using a 7-point Likert scale, offering a granular view of emotional dynamics. Lastly, the dataset includes sensor data from smartphones and MS Band 2 smartwatches, reflecting physical activity (step count, calories burned, activity type), physiological responses (heart rate, skin temperature, electrodermal activity), and environmental context (light exposure, ultraviolet radiation). Smartphone usage data details screen time, app usage, call and message logs, and system events like screen on/off and charging status.

This dataset serves our research question in multiple ways. First, the smartphone sensor data provide direct measures of screen time, useful for correlating with self-reported mental health states. Compared to self-report measures, a direct measure reflects screen time more accurately than self-report. However, note that we only have data from mobile devices which confines us to accounting for mobile screen time only. Additionally, this dataset entails tracking of biophysical markers measured by the MS Band smartwatch. Which among others, entails information on calorie expenditure and heart rate, useful for estimating physical activity. Additionally, the mobile provides data on whether a participants is engaging in an activity, for example walking or running.

Further, the real-time aspect of the ESM data captures fluctuations in stress and in ecologically valid contexts, allowing for dynamic assessment of how screen time and physical activity impact feelings of stress over time. Lastly, the dataset allows analysis of degree of depression severity (PHQ) at the end of the week.

Table 1: Variables and their descriptions

Variable	**Role**	**Type**	**Description**	**Source**
Screen Time (ST)	IV	Continuous	Changes in smartphone's screen state	ScreenEvent.csv (Smartphone; Device; Event)
Physical Activity (PA)	IV	Discrete	Combination of calories burned and detected activity with associated confidence	Calorie.csv (Microsoft Band 2; Physiology; Event) and ActivityEvent.csv (Smartphone; Mobility; Adaptive)
Self-reported Stress (SS)	DV	Continuous	In-situation emotion, stress, and attention labels	EsmResponse.csv (Smartphone; ESM; 45min)
Depression Score (DS)	DV	Continuous	Depression severity measured by PHQ questionnaire	UserInfo.csv (Pre/post survey; User; One-time)

4.2 Pre-processing

In the context of investigating the relationships between screen time, physical activity, self-reported stress, and depression scores preparing the data by pre-processing and extracting features from it was essential. The following steps outline this process, ensuring integrity and usability of the K-emoPhone dataset.

First, we loaded the data to restructure according to our needs and only include relevant files. The initial dataset is structured to contain measurement data as subfolders per person. As we want to access the data per topic for all individuals instead of the other way around, we wrote a script which traversed through the directory structure to access and collect individual data for each file type. This resulted in four lists containing 77 dataframes, each dataframe corresponding to a person. This included a list with all calorie.csv, HR.csv, ScreenEvent.CSV and AcitivtyEvent.csv. Additionally, we loaded ESMResponse.CSV for self-reported stress, UserInfo.csv for depression scores which were both single files containing all participants. To include participants' unique identifier, we made sure to extract the corresponding 'pcode' from the participant folder names for later use. At this point, it was useful to extract date and times during loading, using the `pd.to_datetime` module for a safe conversion from raw Unix timestamps. Sporadic out-of-bound datetimes were dropped. Importantly, datetime had to be adjusted to Asian/Tokyo time zones due to the fact we were dealing with Korean data.

The dataset contained a considerable amount of missing values due to the erratic collection methods associated with high-volume, wearable sensor data. For example, participants did not always wear their smartwatches when required, or filled out ESM responses infrequently. However, as we were aggregating data over days, the impact of missing values was minimized. Additionally, a balanced panel is not a strict requirement for fixed effects models and are partially built to handle missing data. However, to eliminate extreme cases we dropped participants 22 and 61, who had less than 3 recorded rows in the variables we were interested in.

4.3 Feature extraction

The next step is to extract features, simplifying the input data by reducing its dimensionality, distilling the large sets of raw data into what is meaningful to the performing the analyses. Here, the objective is to obtain a panel dataset containing all variables of interest for each day of the week, for each participant.

4.3.1 In-Situ Stress

Extracting daily stress levels of the participants was relatively straightforward. In-situ stress scores were recorded on a 7-point scale, ranging from -3 to 3. As there were up to 16 ESM entries per day, we calculated the daily mean stress scores by dividing by the amount of entries that day. Lastly, daily mean stress values were added to a new dataframe where it was grouped by date and person (pcode).

4.3.2 Screen Time

The next feature that needed to be extracted was participants' screen time. For this the screenEvent dataframes were accessed which contained recordings on every change in screen status, it being either 'ON', 'UNLOCK', and 'OFF'. We measured the duration in milliseconds that was spent between an when the screen was actually unlocked (UNLOCK event) until it was turned off again (OFF event), as the 'ON' event is no proof of the user engaging with their phone. These periods were converted to seconds and were summed for each participant, for each day. The result was appended to a new dataframe, storing pcode, date and total daily screen time.

4.3.3 Calorie Expenditure

The next feature we extracted was based on the calorie expenditure of a participant in one day that is measured by the MS smartwatch. Aimed at revealing participants' physical acitivity levels, Calorie.csv contains information on the number of calories a participant is burning during a given day and the total number of calories burned since the beginning of the experiment. Most likely, the MS calculates this based on its internal heart rate monitor, accelerate, gyroscope and GPS. Therefore, we estimate this measure to be a decent reflection of physical activity. As the original CSV files record calories cumulatively, the value in the last entry provides the total number of calories burned for that day. For the total burned calories (caloriesTotal), we directly used the data from the original dataframe as the columns already contained the required cumulative values. The limitation of this feature is that calorie expenditure alone is not the best reflection of physical activity, in the sense that it does not account for the intensity or type of activity performed. Additionally, we have no information on participant weight or metabolic rate, which prohibits a more comprehensive view on physical activity.

4.3.4 Activity Event

Another way of determining participant's physical activity levels, we constructed a feature based on data from the activityEvent.csv file. In this file, approximately every 15 milliseconds, the participants' mobile device records an estimate of the participants' activity state. This is in the form of a confidence level (0-1). A script running through each participants' activity data detected periods of activity by comparing the confidence level in either one of the columns `confidenceRunning, confidenceOnBicycle, confidenceOnFoot, confidenceWalking` to our (heuristically) set threshold of 0.8. If the one of these cell values exceeded this threshold, the individual was considered active. Again using the timestamps, we calculated the time spent in active periods and converted this to minutes. Again, these were grouped by day, per participant and appended to a new dataframe containing columns `pcode, date, total active minutes`. Our criterion for determining activity was a confidence level above 80 percent. Specifically, we assumed that if the mobile device was more than 80 percent confident that the participant was walking, running, or biking, the participant was likely engaged in one of these activities. Missing entries were interpreted as the participant being inactive.

One limitation of this approach is that it only records activity when the participant is carrying their phone, which is not guaranteed. Despite this, it may still provide valuable insights into time participants spent being active in a day.

4.3.5 Combined physical activity

To create a comprehensive measure of physical activity, we combined the Calories and Activity Event (AE) features, leveraging the strengths of each.

First, we split the data from Activity Event into categories. According to the World Health Organization (WHO) guidelines, we categorized row as 'inactive' if the total daily minutes of activity were less than 21.42 minutes, 'moderate' if the minutes ranged between 21.42 and 42.8, and 'highly active' if the minutes exceeded 42.8.

Next, we split the data from Calorie Expenditure into categories. Based on average resting expenditure, also known as Basal Metabolic Rate (BMR), we estimated whether someone was active relative to their own baseline. The Harris-Benedict equations, a widely used method for estimating BMR, were utilized for this purpose ([48]). The equations for men and women are as follows:

$$\text{BMR (male)} = (13.7516 \times \text{weight in kg}) + (5.0033 \times \text{height in cm}) - (6.755 \times \text{age in years}) + 66.473$$
$$\text{BMR (female)} = (9.5634 \times \text{weight in kg}) + (1.8496 \times \text{height in cm}) - (4.6756 \times \text{age in years}) + 655.0955$$

Since we did not have the weight and height information of the participants, we estimated the BMR for men and women based on the average weight and height of Korean young adults. According to recent studies, the average height and weight for male Korean young adults are 174 cm and 76.5 kg, and for female Korean young adults 161 cm and 57 kg respectively ([49]–[51]. Given the participants' gender and age information, we inserted these values into the respective formulas:

$$\text{BMR (male)} = (13.7516 \times 76.5) + (5.0033 \times 174) - (6.755 \times \text{age in years}) + 66.473$$
$$\text{BMR (female)} = (9.5634 \times 57) + (1.8496 \times 161) - (4.6756 \times \text{age in years}) + 655.0955$$

Lastly, we defined a simple set of logic rules to attribute someone to one of the categories. If a participant was 'highly active' in both Activity Event and Calorie expenditure, they were classified as 'highly active' under combined activity level. Cases where the activity levels assigned by the Calories and AE features differ by only one level, we adopted the lower activity level. For example, if the Calories feature classifies a participant's day as 'highly active' while the AE feature classifies it as 'moderately active', we assigned them 'moderately active' as the combined activity level. When the activity levels differ by more than one level (e.g., one feature classifies the day as 'inactive' and the other as 'highly active'), we chose the middle ground, 'moderately active'.

By combining these features and implementing a system to reconcile differing activity levels, we aimed to create a more comprehensive measure of physical activity. This combined feature is expected to enhance our ability to interpret and analyze participants' physical activity more accurately by taking into account potential limitations and inconsistencies in the data.

4.4 Statistical analysis

The data will be analysed using statistical regression models. Firstly, a fixed effects model will be employed to capture the potential relationship between ST, PA and SS over time. This model allows us to analyse within-person changes in SS over time while accounting for individual-specific factors that remain constant over time. It does so by including individual-specific fixed effects, which control for variables that have not or cannot be measured [52]. Essentially each participant is used as their own control. The fixed effects capture individual-specific deviations from the overall mean of SS, accounting for individual heterogeneity. In other words, we ensure that the estimated effects are not confounded by unobserved individual characteristics. These factors could include personality traits, living situation, or mental health history, which may influence smartphone usage patterns, physical activity and stress levels. In order to employ the fixed effects model, the data will be calculated as means for each participant per day, where ST and PA serve as the independent variables, and SS as the dependent variable. This would result in the following model:

$$\text{StressScore}_{it} = \alpha_i + \beta_1 \text{ScreenTime}_{it} + \beta_2 \text{PhysicalActivity}_{it} + \varepsilon_{it} \tag{1}$$

Here, StressScore_{it} is the stress score for participant i at time t. α_i represents the fixed effects for each individual, capturing unobserved individual-specific characteristics that are constant over time. β_1 and β_2 are the coefficients for Screen Time and Physical Activity, respectively, measuring their impact on the stress score. ε_{it} is the error term for individual i at time t, accounting for random fluctuations in stress score that are not explained by the predictors. A limitation of this approach is that fixed effects estimates could have rather large standard errors, leading to wider confidence intervals [52]. This is due to the fact that only within-subject differences are included, while disregarding between-subject information. In order to mitigate such shortcomings, a separate regression analysis will follow.

A multiple linear regression model with a between-subjects design will be utilised to examine the relationships between the independent variables (ST, PA) and one dependent variable (SS or DS). This model is suitable here as it allows us to analyze the relationships between multiple independent variables and a single dependent variable in each analysis [53]. In this case we have two independent variables (ST and PA) and two dependent variables (SS and DS). By examining the relationships between ST and PA with each dependent variable individually, we can gain a more detailed understanding of how these independent variables affect SS and DS separately, rather than combining all variables in a single analysis.

The multiple linear regression model can be specified as follows, for stress and depression scores respectively:

$$\text{StressScore}_i = \beta_0 + \beta_1 \text{ScreenTime}_i + \beta_2 \text{PhysicalActivity}_i + \epsilon_i \tag{2}$$

$$\text{DepressionScore}_i = \alpha_0 + \alpha_1 \text{ScreenTime}_i + \alpha_2 \text{PhysicalActivity}_i + \zeta_i \tag{3}$$

Here, $StressScore_i$ is the median stress score for participant i and $DepressionScore_i$ is the singular depression score for participant i. β_0 and α_0 are the intercepts for the stress and depression models, respectively. β_1, β_2 are the coefficients representing the impact of screen time and physical activity on stress scores, and α_1, α_2 are the coefficients on depression scores. ϵ_i and ζ_i are the error terms for stress and depression scores, respectively. Note here that these models look very similar to the fixed effects model. However, these models do not include time dimension t. Instead, measurements are taken at a single point in time.

Moreover, the combination of these two methods allows us to answer our research question effectively and provide a thorough investigation into the dynamic interplay between all variables, combining a temporal within-subjects analysis, with a between-subjects regression analysis.

4.5 Assumptions

To perform the panel data regression analysis correctly and to ensure reliability and validity, it is important to ensure the panel data fixed effects assumptions are fulfilled, which are in essence extensions of the least squares assumptions for linear regression [54]

The first assumption dictates that the zero conditional mean assumption requires that the expected value of the error term, given the explanatory variables, is zero, also referred to as strict exogeneity [54]. Additionally, it is assumed that observations across entities are independent and identically distributed. The model also relies on the premise that large outliers are unlikely, and that there is no perfect collinearity among the predictors. These assumptions are to ensuring the model's validity and accuracy in estimating the relationships of the variables [54]. The Variance Inflation Factor (VIF) was computed to check for multicollinearity among the predictors. Lastly, the Durbin-Watson statistic was calculated to assess cross-sectional dependence, ensuring independence of observations.

The Multiple Linear Regression Model (MLR) also operates under several assumptions. The normality assumption requires that the variables are normally distributed. It is also essential for the model to assume a linear relationship between the independent and dependent variables. Another assumption is that the variables are measured without error. Finally, homoscedasticity is assumed, meaning the variance of errors is constant across all levels of the independent variables [55]. For the MLR, the normality of residuals can be checked using Q-Q plots, and the Shapiro-Wilk test. Scatterplots of the independent variables against the dependent variable are used to confirm linear relationships. The Breusch-Pagan test was conducted to check for homoscedasticity. Additionally, residuals vs. fitted plots were examined.

In the event that assumptions are not met for one of the dependent variables, the extremity of the violations will be assessed first. Depending on this assessment, the analysis may still be conducted. However, the results will be interpreted with caution.

5. Results

First, analyses were conducted only on participants that contained sufficient data for all variables for at least 3 days (N = 75). Participants were, on average, active for 33.84 minutes per day (μ = 33.84, σ = 26.84) and burned 1606.05 kCal per day (μ = 1606.05, σ = 343.10). The combined physical activity feature leans towards 'inactive' (μ = 0.34, σ = 0.47, range = 0-2). The average stress scores per day leans towards 'not stressed' (μ = -0.27, σ = 1.02, range = -3-3). Lastly, participants reported a mild depression score (PHQ) on average (μ = 4.92, σ = 4.61, range = 0-27). Figure 1 shows the distribution of the data that proved to be significant for this research. Figure 2 shows the data distribution of depression scores.

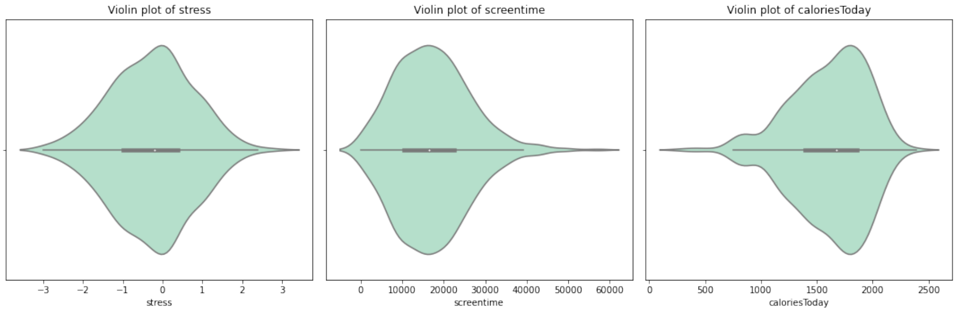

Figure 1: Data Distribution for Stress, Screen time, and Calories

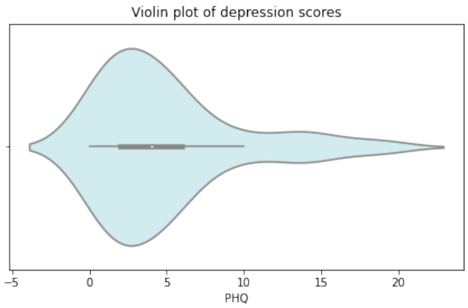

Figure 2: Data Distribution for Depression Scores

5.1 Fixed Effects Regression Model

Firstly, the inclusion of the combined physical activity feature did not produce a significant effect on stress. Specifically, none of the levels were found to have statistically significant coefficients. The overall R-squared value was very low (R2 = -0.0176) indicates a poor fit. The high p-value of the F-statistic (F = 0.0310, P = 0.9927) suggest that these predictors are not adequate for modeling stress within this dataset. The significant F-test for poolability (F = 5.3010, P = 0.0000) indicates the predictors do not capture what causes variations in stress effectively. See Appendix B.1 for the full regression table.

As a result, the model was re-ran with both predictors (calorie expenditure and activity event), separately. Running the model with Activity Event as substitute for physical activity was not significant either β_0 = x, (P = x). However, running the model with calorie expenditure instead yielded a significant effect at a level of 5% of calorie expenditure on stress, with β = -0.0004 (SE = 0.0002, t = -2.086, P = 0.0376). This indicates that higher calorie expenditure is associated with a slight decrease in stress. This relationship, though statistically significant, had an extremely low effect size (R2 = 0.0112), implying that while calorie intake does impact stress, the magnitude of this impact is small, explaining only 1.2% of the variance. From here onwards, we continued the analyses using calorie expenditure as estimate for physical activity.

In the model using calorie expenditure, screen time was found to have no significant effect on stress, with β_0 = -6.185e-07 (SE = 5.942e-06, t = -0.1041, P = 0.9172), indicating the fixed effects model did not identify effect of screen time on stress levels within participants.

The entity-specific intercepts (fixed effects) are together significantly different from zero (F(74,384) = 4.888, P $<$ 0.001). The model may not effectively explain variations in stress, across all participants and days, though calorie expenditure did appear to be a significant predictor for stress. Overall, this confirms the necessity of accounting for individual differences via fixed effects, as evidenced by the significant F-test for poolability. See Appendix B.3 for the full regression table.

In figure 3, calorie expenditure and screen time are plotted against the stress scores. Upon visual inspection, neither plots show a distinct correlation. In stress against calorie expenditure, a very slight negative relationship can be observed, supporting the notion that lower stress is associated with higher calorie expenditure within participants.

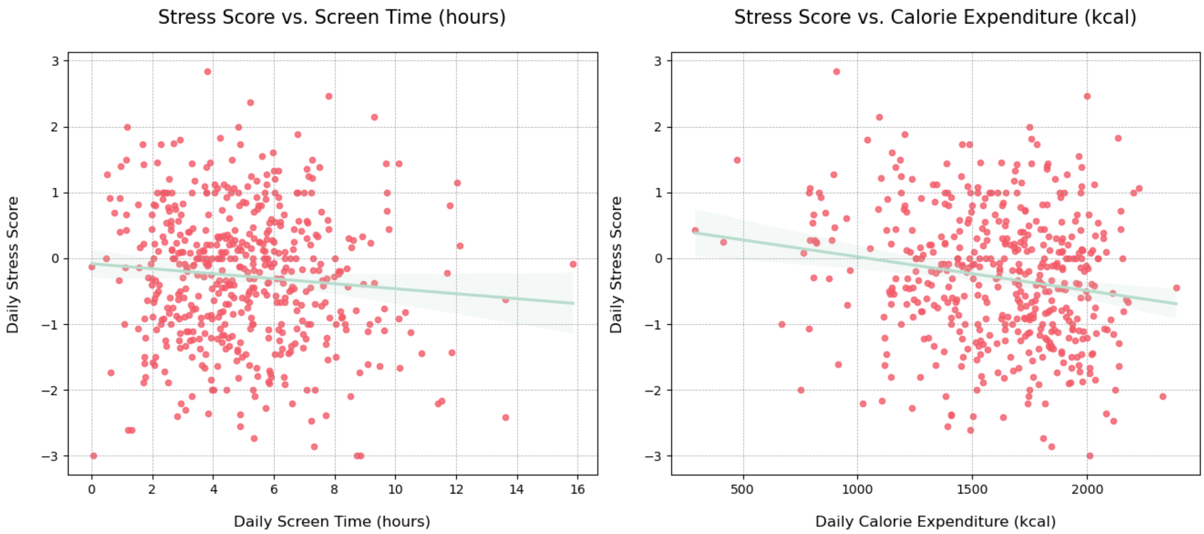

Figure 3: Scatterplots of daily screen time and calorie expenditure against stress scores

In figure 4, a sample of participants (10) is plotted over time against all variables.

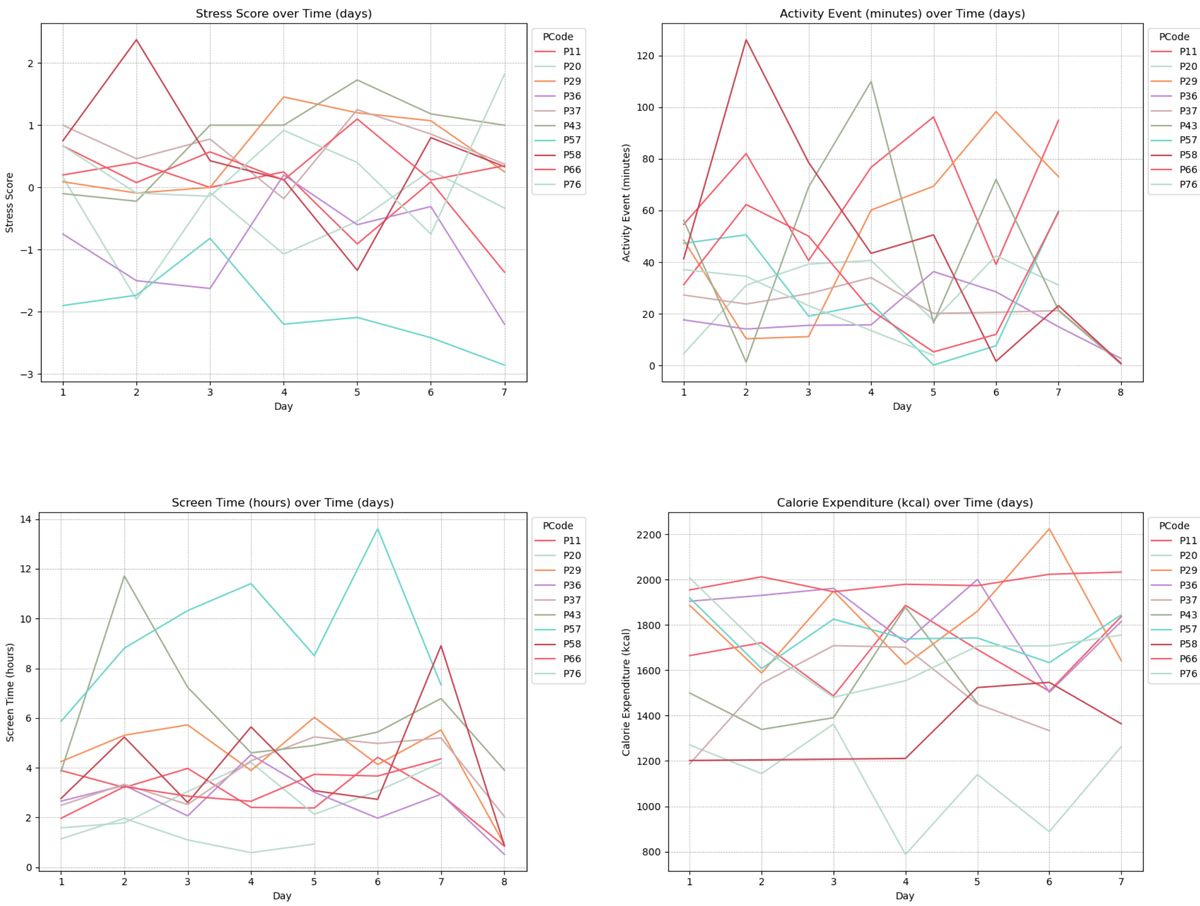

Figure 4: Random sample of 10 participants plotted over time against all variables

5.2 Multiple Linear Regression Model

When running the multiple regression for depression, neither screen time nor daily calorie expenditure are significant predictors, with β = -3.661e-06 (SE = 9.1e-05, t = -0.040, P = 0.968) for screen time and β = 0.0015 (SE = 0.002, t = 0.805, P = 0.423) for calorie expenditure. The intercept is not statistically significant either, with β = 2.517, (SE = 3.576, t = 0.704, P = 0.484), indicating that the expected depression score is not significantly different from zero.

With a negative adjusted R-squared (-0.018), the model seems poorly fitted. Overall, the model does not significantly predict depression (F = 0.3296, 0.720). Most likely, the problem of non-normal residuals and potential numerical issues indicated by the high condition number (1.2e+05) suggests issues with multi-collinearity, making the depression data unreliable for this analysis. See Appendix B.5 for the regression table.

Conversely, the multiple regression model for stress show significant effects of both screen time and daily calorie expenditure: With a coefficient β = = -3.317e-05 for screen time (SE = 1.43e-05, t = -2.325, P = 0.023), and for calorie expenditure β = -0.0006 (SE = 0.000, t = -2.158, P = 0.034). The intercept is also significant with β = 1.3420 (SE = 0.561, t = 2.393, P = 0.019).

The entire model explains about 9% of the variance in stress levels between people (adj. R-squared = 0.090). Overall, calorie expenditure and screen time collectively provide a better fit than an intercept-only model (F = 4.657, P = 0.0125). The model indicates that both screen time and calorie expenditure negatively correlate with stress levels between participants, and both predictors are statistically significant. The overall fit of the model is moderate, and shows some predictive power, though the condition number remains high (1.20e+05).

Below, scatter plots of the variables against each other are displayed. The visualizations are accompanied by confidence bands indicating the uncertainty around the trend lines. The stress score plots both show a slight negative relation between screen time and stress and calorie expenditure and stress, supporting that between people, higher screen time is associated with lower stress levels, and higher calorie expenditure is associated with lower stress as well.

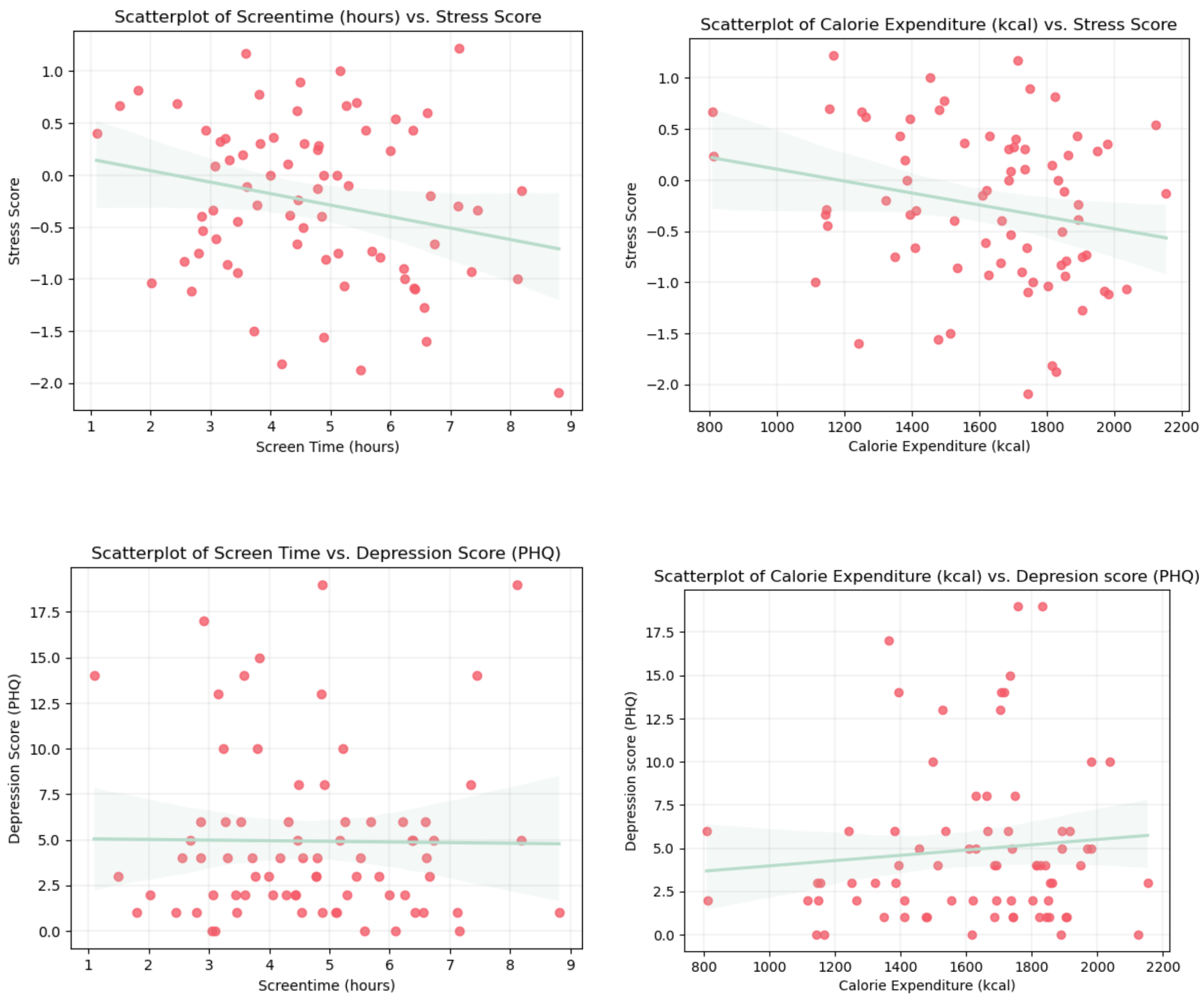

Figure 5: Scatterplots Multiple Regression

5.3 Assumption tests

The residuals of the fixed effects model and both independent variables are upon visual inspection randomly scattered, showing no systematic bias in the models prediction [54], verifying the assumption of zero conditional mean (see Appendix A.1). Next, a Durbin-Watson test resulted in a value close to 2 (statistic = 1.896), indicating no autocorrelation, satisfying the assumption of cross-sectional dependence. With visual inspection and z-score tests, we conclude that the residuals of the fixed effects model and both independent variables generally fall within a reasonable range, and likely do not affect the estimation significantly [54]. Therefore, Satisfying the third assumption (see Appendix A.2). Lastly, we assessed the models' multi-collinearity with the variance inflation factor (VIF = 4.0752), which is generally acceptable. The predictors may have some level of linear dependency, which is somewhat expected in fixed effects models [54]. Though some outliers were visible, the fixed effects model overall satisfies the required assumptions.

Regarding the multiple regression, the residuals for screen time and calories and the fitted values were evenly distributed, suggesting no systematic error in predictions. No apparent pattern or curvature showed, supporting linearity for both stress and depression scores (See Appendix A.3. Next, the Durbin-Watson test was close to 2 (statistic = 2.324), suggesting there was likely no significant autocorrelation, therefore the assumption of independence holds [54]. Next, the Breusch-Pagan test for the fitted values for stress yielded a p-value larger than 0.05 (p = 0.608) and for depression the same was observed (p = 0.768). This indicates that there is also no evidence of heteroscedasticity in this model [54]. Next, the Q-Q plots show that stress scores are normally distributed. For depression, the central portion of data aligns well, suggesting normal behavior in the middle range of the data. However, there are significant deviations from the line at both tails, especially the upper tail, suggesting normality of errors might be violated (See Appendix A.4[54]. The shapiro-Wilk test for stress was non-significant, suggesting the residuals are normally distributed (P = 0.349). However, for depression the result was significant, indicating that the results are not normally distributed (P = 1.62..5e-07), violating the assumption of normality of residuals. Moving on, the VIF value was close to 1 (1.0065), indicating no multi-collinearity among variables [54]. The final assumption, exogeneity, was not accounted for as it is not directly testable with basic diagnostics, and due to time constraints unfeasible to the current research [54]. To summarize, while most assumptions for the multiple regression are met, there are concerns regarding the normality of residuals, particularly for depression scores, which might impact the validity of the regression model for predicting depression. Since all other assumptions for depression do hold, the analysis was conducted anyway. However, it should be emphasized that the results of this analysis may not reflect accurate relationships.

6. Discussion

The current study was aimed at gaining an understanding how the amount of smartphone screen time and amount of physical activity related to self-reported stress and depression scores among young adults. Our fixed effects analysis showed a significant effect of calorie expenditure on stress levels, and failed to find any significant results of screen time on stress. For the multiple regression, a significant effect of screen time on stress was found and calorie expenditure on stress. No significant results were found for any of the predictors on depression scores.

The lack of significant effect of screen time on stress within the FEM suggests that, within individuals, variations in screen time do not substantially correlate with variations in stress levels over time. This could indicate that individual-specific factors, such as personality traits or mental health history, might moderate the relationship between screen time and stress, making it difficult to detect significant changes using only temporal within-subject data. Unsurprisingly, many factors such as sleep disturbances, other health-related habits, negative mood and acute environmental factors can also impact stress, and considering the simplicity of our model, this is not entirely unexpected [17], [31]–[33], [45].

Interestingly, the multiple linear regression model (MLR) showed that higher screen time is negatively related to stress. Thus, we can conclude that differences in screen time across participants slightly contribute to differences in stress levels. This finding was rather unexpected in light of our scope, as we had enough support from literature that generally, higher self-reported mobile phone use is linked to stress among young adults [17]–[19]. Additionally, extensive phone use is associated with dependence, which is expected to further increase stress [21]. Possibly, this contrasting finding may be attributed to the speculation that participants who are more occupied with university or their work, have less time to spend on their phone in general, which would explain why higher stress is related to lower screen time. Such factors lifestyle or environmental factors which are not captured in the multiple regression model. Another possible explanation might be that screen time is in fact, a soothing behavior. Spending time on ones' smartphone may possibly lead to stress reduction, as it can be considered a means of distracting oneself, or socialising with friends, explore interests and discover content.

Either way, this further emphasizes our argument that combining both fixed effects and multiple regression allows us to provide a more comprehensive understanding of the different relationships between all variables.

Calorie expenditure was a significant predictor in both models, confirming a relationship between physical activity and stress levels. Thus, increasing one's physical activity seems a promising strategy to managing stress. It was apparent that our categorical physical activity feature did not perform well in neither models. Instead, daily calorie expenditure showed to be better estimation of physical activity. Most likely, our categorization of participants based on different aspects of physical activity was an oversimplification of the data.

The relationship between calorie expenditure and stress levels across both models underscores the potential of increasing physical activity as a stress management strategy. Similarly, another study among healthy adults found that those who engaged in regular exercise reported experiencing fewer negative emotional consequences of stress [5]. Such findings are widely supported by research [18], [42]–[44].

Overall, all the models' effect sizes were very small, suggesting that the significant results should not be considered very meaningful and should be interpreted carefully. However, this aligns with research by Orben and Przybylski (2019), who showed associations between screen time and mental health symptoms generally have small effect sizes.

6.1 Limitations and future works

A first limitation of the current study is the fact that relying on self-reported stress and depression scores may have introduced bias and inaccuracy. Inconsistent use of wearable devices led to missing data, as participants did not always follow the specified time frame for wearing the smartwatch, resulting in the exclusion of participants with insufficient data (see Pre-processing 4.2). To ensure data completeness and consistency, better adherence protocols for data collection are needed. This would yield more accurate data sets. Implementing clear rules for data collection, possibly involving more participants but only using complete datasets, would be beneficial. For continued research, generating specific data for the study or collating existing data that aligns with the study's objectives would be useful.

Additionally, future studies should aim to incorporate more objective measures of mental health by combining self-reported and objective data. Although mental health is difficult to measure objectively, using physiological indicators like heart rate and cortisol levels can be beneficial. However, it is challenging to separate these indicators from other similar physiological responses, especially in natural settings.

The depression data violated the assumption of normal distribution, with most participants having normal scores and some outliers being extremely depressed (Figure 2). This limits the interpretation of the models, possibly also explaining insignificant results. The PHQ scores may not have been reliable, as depression may occur in phases, a PHQ score measured in the current study reflects a one-time measurement. Comparing a one-time score with variables reported over one week may not be precise. When dealing with outliers, studies should consider whether it is logical to exclude extreme scores, as such extreme cases can also provide useful insights. Otherwise, future works could include different measurements for more reliable results.

A next limitation was that the dataset did not provide a straightforward measure of physical activity, to which attempted at estimating this using various data sources. We intended to include heart rate as a predictor, but inconsistencies prevented its use. However, we combined calorie and physical features, which are likely oversimplified information and misclassified individuals. For example, the combined feature showed no one as highly active, while some were highly active in each category separately, leading to information loss. This issue may arise from calculating BMR using average values across genders and ages. If the dataset included weight and height, the measure could have been more accurate. Instead, calorie expenditure was used as estimator for physical activity, which may not be perfectly representative as this can also be a result from stress or cognitive load. Future work should focus on using data proven to accurately measure physical activity. Self-reported evidence or formulas based on various physiological values could also be beneficial.

The screen time variable requires a more precise definition, as its impact varies depending on the activities participants engage in on their devices. For example, screen time effects differ between work-related and leisure activities. Additionally, the current measure may overlook other devices, such as computers or televisions. Future studies should analyze the "screen time" variable in greater depth, including the use of other mobile devices. Preliminary studies or existing research on these aspects would help refine the variables, improving the accuracy and reliability of findings. On that note, as our results do not correlate with the main findings of existing research it is clear that this relationship is more complex than it seems. More intricate dynamics may be at play here. On that note, we didn't take different types of screen activities into account in this study. Texting friends may have a different effect on stress than scrolling through Instagram. This underscores the notion that additional research could be done on the impact these different screen activities could have on stress levels.

7. Conclusion

In conclusion, this study underscores the relationship between smartphone use, physical activity, and mental health among young adults. Our findings indicate that while increased screen time is associated with lower stress levels, it does not significantly impact stress within individuals over time. Additionally, calorie expenditure emerged as a consistent predictor of reduced stress, highlighting the critical role of physical activity in managing stress.

The lack of significant within-person effects for screen time suggests that individual traits or environmental factors may moderate this relationship, pointing to the complexity of the screen time impact on mental health. This complexity is emphasized by the contrasting findings that higher screen time reduces stress levels between-subjects. The categorical measure of physical activity, is proven ineffective, likely due to its oversimplification. This calls for more precise measures of physical activity in future research.

The depression data did not yield valid results in our models, emphasizing the need for more sophisticated and perhaps longitudinal measures of depression that can capture its fluctuating nature. Future studies should integrate objective mental health assessments alongside self-reported data to enhance the reliability and comprehensiveness of findings.

Our research highlights the importance of developing strategies to better understand the effects of excessive screen time and promote physical activity among young adults. Enhanced data collection methods and broader variable exploration will be crucial for future research to unravel the complex interplay between digital habits, physical activity, and mental health.

This study provides valuable insights into the dual influence of digital and physical activities on stress. However, it also emphasizes the need for refined measures and more comprehensive data to fully understand these relationships and develop effective interventions.

References

[1] J. M. S. d. Lucena, L. A. Cheng, T. L. M. Cavalcante, V. A. d. Silva and J. C. d. Farias Júnior, 'Prevalence of excessive screen time and associated factors in adolescents,' *Revista Paulista de Pediatria*, vol. 33, pp. 407–414, 2015.

[2] L. Vanhees, J. Lefevre, R. Philippaerts *et al.*, 'How to assess physical activity? how to assess physical fitness?' *European Journal of Cardiovascular Prevention & Rehabilitation*, vol. 12, no. 2, pp. 102–114, 2005.

[3] G. Mangkuliguna, V. M. Sidharta, L. Hananta *et al.*, 'Sedentary screen time as a coping strategy of distance learning-induced distress during covid-19 pandemic,' *Journal of Educational, Health and Community Psychology*, vol. 10, no. 2, p. 175, 2021.

[4] N. Khalili-Mahani, A. Smyrnova and L. Kakinami, 'To each stress its own screen: A cross-sectional survey of the patterns of stress and various screen uses in relation to self-admitted screen addiction,' *Journal of medical Internet research*, vol. 21, no. 4, e11485, 2019.

[5] E. Childs and H. De Wit, 'Regular exercise is associated with emotional resilience to acute stress in healthy adults,' *Frontiers in physiology*, vol. 5, p. 87 290, 2014.

[6] X. Zhu, J. A. Haegele, H. Liu and F. Yu, 'Academic stress, physical activity, sleep, and mental health among chinese adolescents,' *International Journal of Environmental Research and Public Health*, vol. 18, no. 14, p. 7257, 2021.

[7] A. J. Mitchell, J. C. Coyne *et al.*, *Screening for depression in clinical practice: an evidence-based guide*. Oxford University Press, 2009.

[8] A. Pandya and P. Lodha, 'Social connectedness, excessive screen time during covid-19 and mental health: A review of current evidence,' *Frontiers in Human Dynamics*, vol. 3, p. 684 137, 2021.

[9] C. O. COMMUNICATIONS *et al.*, 'Children, adolescents, and the media,' *Pediatrics*, vol. 132, no. 5, pp. 958–961, 2013.

[10] P. Baiden, S. K. Tadeo and K. E. Peters, 'The association between excessive screen-time behaviors and insufficient sleep among adolescents: Findings from the 2017 youth risk behavior surveillance system,' *Psychiatry research*, vol. 281, p. 112 586, 2019.

[11] D. D. Felisoni and A. S. Godoi, 'Cell phone usage and academic performance: An experiment,' *Computers & Education*, vol. 117, pp. 175–187, 2018.

[12] R. Junco, 'Comparing actual and self-reported measures of facebook use,' *Computers in Human Behavior*, vol. 29, no. 3, pp. 626–631, 2013.

[13] J. Boase and R. Ling, 'Measuring mobile phone use: Self-report versus log data,' *Journal of Computer-Mediated Communication*, vol. 18, no. 4, pp. 508–519, 2013.

[14] Y.-H. Lin, Y.-C. Lin, Y.-H. Lee *et al.*, 'Time distortion associated with smartphone addiction: Identifying smartphone addiction via a mobile application (app),' *Journal of psychiatric research*, vol. 65, pp. 139–145, 2015.

[15] L. N. Hodes and K. G. Thomas, 'Smartphone screen time: Inaccuracy of self-reports and influence of psychological and contextual factors,' *Computers in Human Behavior*, vol. 115, p. 106 616, 2021.

[16] S. Kang, W. Choi, C. Y. Park *et al.*, 'K-emophone: A mobile and wearable dataset with in-situ emotion, stress, and attention labels,' *Scientific data*, vol. 10, no. 1, p. 351, 2023.

[17] S. Thomée, A. Härenstam and M. Hagberg, 'Mobile phone use and stress, sleep disturbances, and symptoms of depression among young adults-a prospective cohort study,' *BMC public health*, vol. 11, pp. 1–11, 2011.

[18] Y. Ge, S. Xin, D. Luan *et al.*, 'Independent and combined associations between screen time and physical activity and perceived stress among college students,' *Addictive behaviors*, vol. 103, p. 106 224, 2020.

[19] S. Thomée, M. Eklöf, E. Gustafsson, R. Nilsson and M. Hagberg, 'Prevalence of perceived stress, symptoms of depression and sleep disturbances in relation to information and communication technology (ict) use among young adults–an explorative prospective study,' *Computers in human behavior*, vol. 23, no. 3, pp. 1300–1321, 2007.

[20] S. J. Lupien, F. Maheu, M. Tu, A. Fiocco and T. E. Schramek, 'The effects of stress and stress hormones on human cognition: Implications for the field of brain and cognition,' *Brain and cognition*, vol. 65, no. 3, pp. 209–237, 2007.

[21] M. Beranuy, U. Oberst, X. Carbonell and A. Chamarro, 'Problematic internet and mobile phone use and clinical symptoms in college students: The role of emotional intelligence,' *Computers in human behavior*, vol. 25, no. 5, pp. 1182–1187, 2009.

[22] T. Razavi, 'Self-report measures: An overview of concerns and limitations of questionnaire use in occupational stress research,' 2001.

[23] J. F. Kihlstrom, E. Eich, D. Sandbrand and B. A. Tobias, 'Emotion and memory: Implications for self-report,' in *The science of self-report*, Psychology Press, 1999, pp. 93–112.

[24] T. R. Mitchell, L. Thompson, E. Peterson and R. Cronk, 'Temporal adjustments in the evaluation of events: The "rosy view",' *Journal of experimental social psychology*, vol. 33, no. 4, pp. 421–448, 1997.

[25] J. M. Hektner, J. A. Schmidt and M. Csikszentmihalyi, *Experience sampling method: Measuring the quality of everyday life.* Sage, 2007.

[26] N. Van Berkel, D. Ferreira and V. Kostakos, 'The experience sampling method on mobile devices,' *ACM Computing Surveys (CSUR)*, vol. 50, no. 6, pp. 1–40, 2017.

[27] R. Wang, F. Chen, Z. Chen *et al.*, 'Studentlife: Assessing mental health, academic performance and behavioral trends of college students using smartphones,' in *Proceedings of the 2014 ACM international joint conference on pervasive and ubiquitous computing*, 2014, pp. 3–14.

[28] S. M. Mattingly, J. M. Gregg, P. Audia *et al.*, 'The tesserae project: Large-scale, longitudinal, in situ, multimodal sensing of information workers,' in *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, 2019, pp. 1–8.

[29] A. Sano, S. Taylor, A. W. McHill *et al.*, 'Identifying objective physiological markers and modifiable behaviors for self-reported stress and mental health status using wearable sensors and mobile phones: Observational study,' *Journal of medical Internet research*, vol. 20, no. 6, e210, 2018.

[30] S. Saeb, M. Zhang, C. J. Karr *et al.*, 'Mobile phone sensor correlates of depressive symptom severity in daily-life behavior: An exploratory study,' *Journal of medical Internet research*, vol. 17, no. 7, e4273, 2015.

[31] B. E. P. Murphy, 'Steroids and depression,' *The Journal of steroid biochemistry and molecular biology*, vol. 38, no. 5, pp. 537–559, 1991.

[32] P. M. Plotsky, M. J. Owens and C. B. Nemeroff, 'Psychoneuroendocrinology of depression: Hypothalamic-pituitary-adrenal axis,' *Psychiatric Clinics of North America*, vol. 21, no. 2, pp. 293–307, 1998.

[33] M. Pruessner, D. H. Hellhammer, J. C. Pruessner and S. J. Lupien, 'Self-reported depressive symptoms and stress levels in healthy young men: Associations with the cortisol response to awakening,' *Psychosomatic medicine*, vol. 65, no. 1, pp. 92–99, 2003.

[34] P. R. Shaver and L. S. Wrightsman, *Measures of personality and social psychological attitudes.* Elsevier, 1991, vol. 1.

[35] D. American Psychiatric Association, A. P. Association *et al.*, *Diagnostic and statistical manual of mental disorders: DSM-5.* American psychiatric association Washington, DC, 2013, vol. 5.

[36] J. W. Williams Jr, M. Pignone, G. Ramirez and C. P. Stellato, 'Identifying depression in primary care: A literature synthesis of case-finding instruments,' *General hospital psychiatry*, vol. 24, no. 4, pp. 225–237, 2002.

[37] S.-J. Park, H.-R. Choi, J.-H. Choi, K.-W. Kim and J.-P. Hong, 'Reliability and validity of the korean version of the patient health questionnaire-9 (phq-9),' *Anxiety and mood*, vol. 6, no. 2, pp. 119–124, 2010.

[38] S. Hökby, J. Westerlund, J. Alvarsson, V. Carli and G. Hadlaczky, 'Longitudinal effects of screen time on depressive symptoms among swedish adolescents: The moderating and mediating role of coping engagement behavior,' *International journal of environmental research and public health*, vol. 20, no. 4, p. 3771, 2023.

[39] C. Anderl, M. K. Hofer and F. S. Chen, 'Directly-measured smartphone screen time predicts well-being and feelings of social connectedness,' *Journal of Social and Personal Relationships*, p. 02 654 075 231 158 300, 2023.

[40] L. Miles, 'Physical activity and health,' *Nutrition bulletin*, vol. 32, no. 4, pp. 314–363, 2007.

[41] C. J. Caspersen, K. E. Powell and G. M. Christenson, 'Physical activity, exercise, and physical fitness: Definitions and distinctions for health-related research.,' *Public health reports*, vol. 100, no. 2, p. 126, 1985.

[42] P. J. Gianaros and T. D. Wager, 'Brain-body pathways linking psychological stress and physical health,' *Current directions in psychological science*, vol. 24, no. 4, pp. 313–321, 2015.

[43] F. Fleshner, 'Physical activity and stress resistance: Sympathetic nervous system adaptations prevent stress-induced immunosuppression,' *Exercise and sport sciences reviews*, vol. 33, no. 3, pp. 120–126, 2005.

[44] H. W. Bland, B. F. Melton, L. E. Bigham and P. D. Welle, 'Quantifying the impact of physical activity on stress tolerance in college students,' *College student journal*, vol. 48, no. 4, pp. 559–568, 2014.

[45] M. C. Chandler, O. K. Ellison, A. L. McGowan, K. M. Fenn and M. B. Pontifex, 'Physical activity and sleep moderate the relationship between stress and screen time in college-aged adults,' *Journal of American College Health*, pp. 1–11, 2022.

[46] L. G. Sylvia, E. E. Bernstein, J. L. Hubbard, L. Keating and E. J. Anderson, 'A practical guide to measuring physical activity,' *Journal of the Academy of Nutrition and Dietetics*, vol. 114, no. 2, p. 199, 2014.

[47] M. L. Slattery, D. R. Jacobs Jr and M. Z. Nichaman, 'An assessment of caloric intake as an indicator of physical activity,' *Preventive medicine*, vol. 18, no. 4, pp. 444–451, 1989.

[48] J. A. Harris and F. G. Benedict, 'A biometric study of human basal metabolism,' *Proceedings of the National Academy of Sciences*, vol. 4, no. 12, pp. 370–373, 1918.

[49] N. Y. Ryoo, H. Y. Shin, J. H. Kim, J. S. Moon and C. G. Lee, 'Change in the height of korean children and adolescents: Analysis from the korea national health and nutrition survey ii and v,' *Korean journal of pediatrics*, vol. 58, no. 9, p. 336, 2015.

[50] K. Tae-Young, 'Comparison of health level and physical activity between korean and other countries female college students,' *Iranian Journal of Public Health*, vol. 46, no. 11, p. 1586, 2017.

[51] P. by Statista Research Department and A. 9, *Average korean male weight by age group*, Apr. 2024. [Online]. Available: https://www.statista.com/statistics/940140/south-korea-average-weight-men-by-age-group/#statisticContainer.

[52] P. D. Allison, *Fixed effects regression models*. SAGE publications, 2009.

[53] M. R. Del Águila and N. Benítez-Parejo, 'Simple linear and multivariate regression models,' *Allergologia et immunopathologia*, vol. 39, no. 3, pp. 159–173, 2011.

[54] J. H. Stock and W. Watson, *Introduction to econometrics boston: Pearson addison wesley*, 2007.

[55] J. W. Osborne and E. Waters, 'Four assumptions of multiple regression that researchers should always test,' *Practical assessment, research, and evaluation*, vol. 8, no. 1, p. 2, 2019.

A. Assumption plots

A.1 Residuals FE

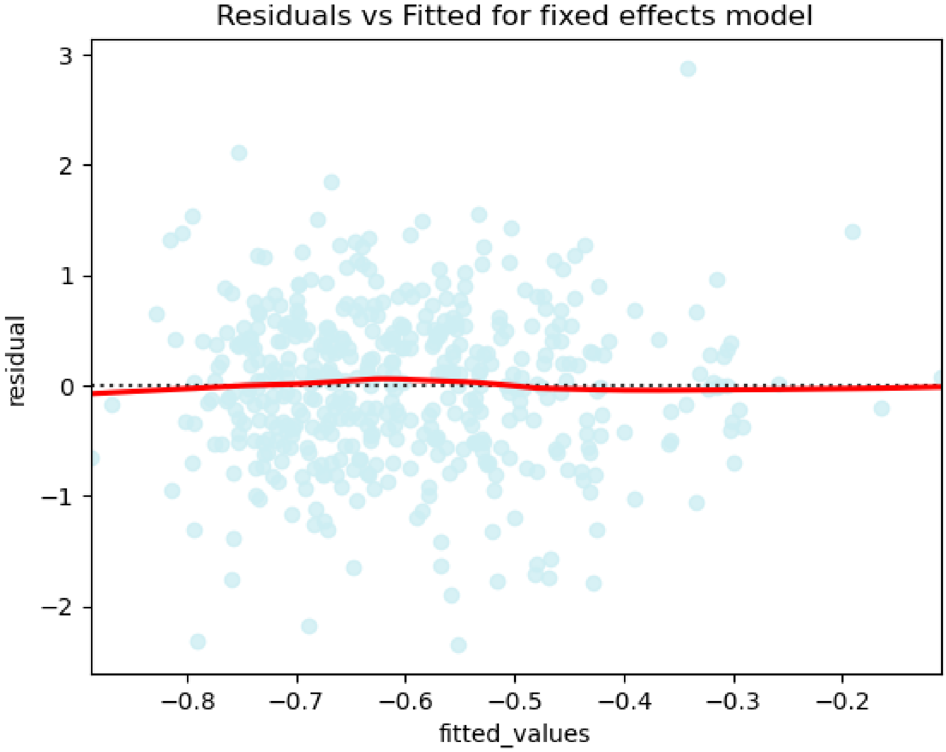

Figure 6: Residuals versus fitted model

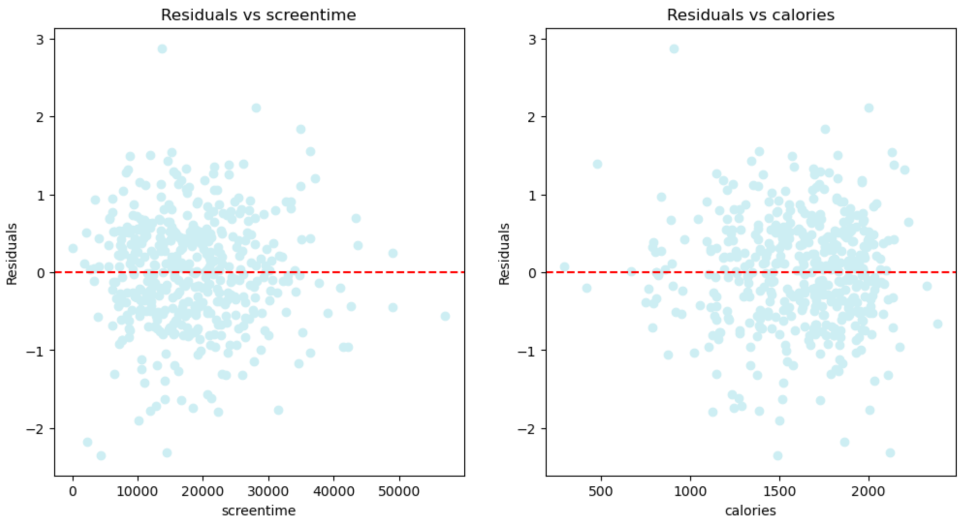

Figure 7: Residuals versus independent variables

A.2 Outliers FE

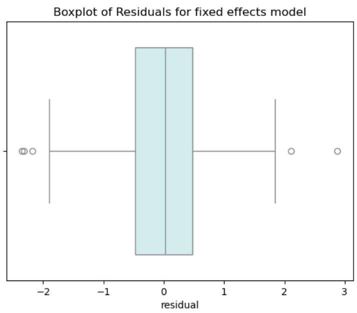

Figure 8: Residuals screentime

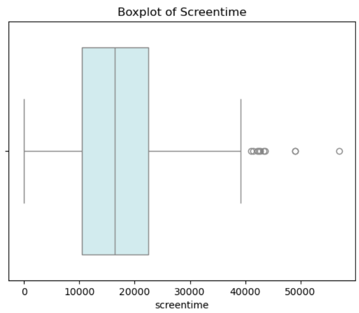

Figure 9: Residuals fixed effects

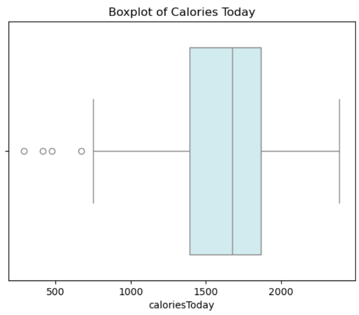

Figure 10: Residuals calories

A.3 Residuals vs. Predicted values

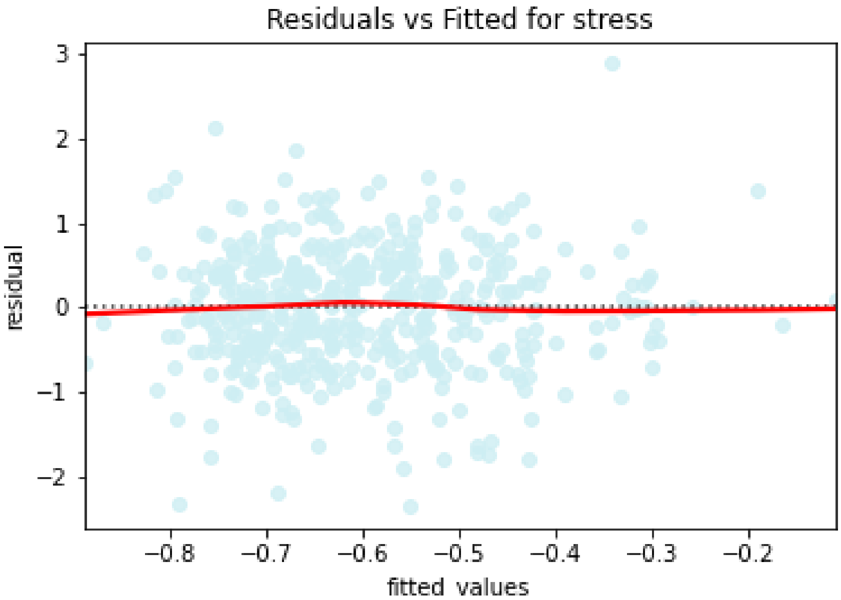

Figure 11: Scatterplot residuals vs. predicted values for stress

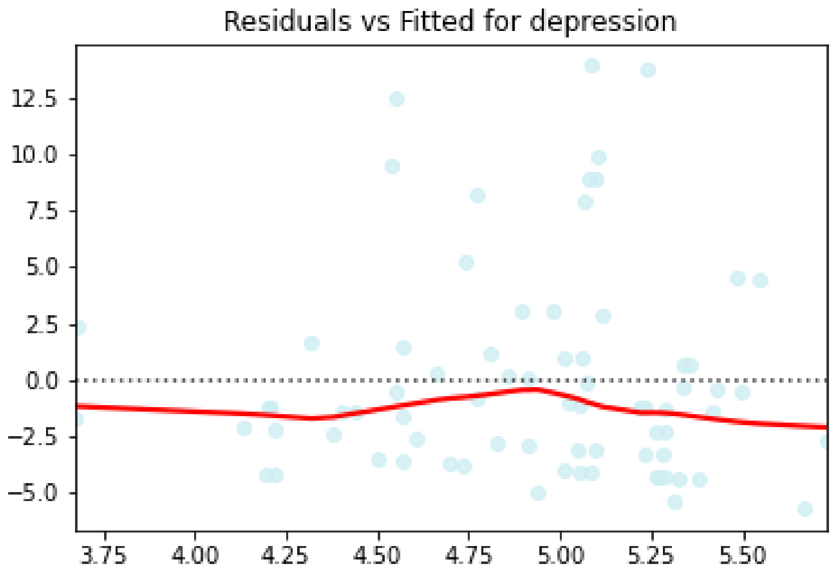

Figure 12: Scatterplot residuals vs. predicted values for depression

A.4 Normality: Q-Q Plots

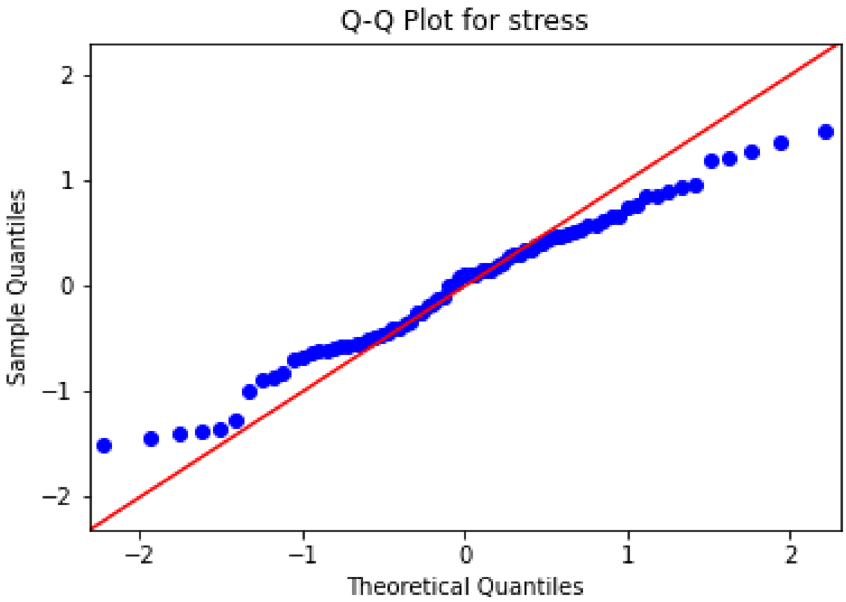

Figure 13: Q-Q Plot for Stress

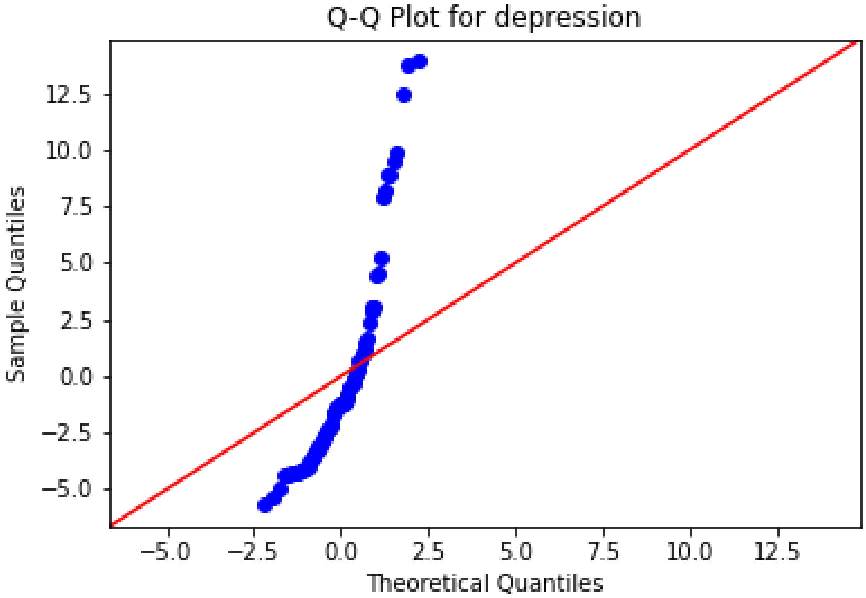

Figure 14: Q-Q Plot for Depression

B. Results Statistical Analyses

B.1 Fixed Effects Model: combined physical activity

```
                          PanelOLS Estimation Summary
================================================================================
Dep. Variable:                 stress   R-squared:                        0.0002
Estimator:                   PanelOLS   R-squared (Between):             -0.0340
No. Observations:                 509   R-squared (Within):               0.0002
Date:                Thu, Jun 06 2024   R-squared (Overall):             -0.0176
Time:                        19:46:05   Log-likelihood                   -568.40
Cov. Estimator:            Unadjusted
                                        F-statistic:                      0.0310
Entities:                          75   P-value                           0.9927
Avg Obs:                       6.7867   Distribution:                   F(3,431)
Min Obs:                       4.0000
Max Obs:                       8.0000   F-statistic (robust):             0.0310
                                        P-value                           0.9927
Time periods:                      23   Distribution:                   F(3,431)
Avg Obs:                       22.130
Min Obs:                       2.0000
Max Obs:                       26.000

                                    Parameter Estimates
==============================================================================================
                                    Parameter  Std. Err.     T-stat    P-value    Lower CI    Upper
----------------------------------------------------------------------------------------------
activity_level[T.inactive]             0.0367     0.1459     0.2518     0.8013     -0.2500      0.32
activity_level[T.moderately active]    0.0156     0.1499     0.1042     0.9170     -0.2789      0.31
screentime                          2.069e-07  5.705e-06     0.0363     0.9711  -1.101e-05   1.142e-
==============================================================================================

F-test for Poolability: 5.3010
P-value: 0.0000
Distribution: F(74,431)

Included effects: Entity
```

Figure 15: Fixed Effects Results using combined physical activity feature

B.2 Fixed Effects Model: Activity Event

```
                          PanelOLS Estimation Summary
================================================================================
Dep. Variable:                 stress   R-squared:                        0.0003
Estimator:                   PanelOLS   R-squared (Between):              0.0212
No. Observations:                 507   R-squared (Within):               0.0003
Date:                Thu, Jun 13 2024   R-squared (Overall):              0.0108
Time:                        21:45:09   Log-likelihood                   -564.08
Cov. Estimator:            Unadjusted
                                        F-statistic:                      0.0644
Entities:                          75   P-value                           0.9376
Avg Obs:                       6.7600   Distribution:                   F(2,430)
Min Obs:                       4.0000
Max Obs:                       8.0000   F-statistic (robust):             0.0644
                                        P-value                           0.9376
Time periods:                      23   Distribution:                   F(2,430)
Avg Obs:                       22.043
Min Obs:                       1.0000
Max Obs:                       26.000

                             Parameter Estimates
==============================================================================
            Parameter  Std. Err.     T-stat    P-value    Lower CI    Upper CI
------------------------------------------------------------------------------
active_min    -0.0006     0.0017    -0.3575     0.7209     -0.0040      0.0028
screentime -1.304e-07  5.632e-06    -0.0232     0.9815    -1.12e-05   1.094e-05
==============================================================================

F-test for Poolability: 5.2746
P-value: 0.0000
Distribution: F(74,430)

Included effects: Entity
```

Figure 16: Fixed Effects Results using activity event

B.3 Fixed Effects Model: Calorie Expenditure

```
                          PanelOLS Estimation Summary
================================================================================
Dep. Variable:                 stress   R-squared:                        0.0112
Estimator:                   PanelOLS   R-squared (Between):             -0.0259
No. Observations:                 461   R-squared (Within):               0.0112
Date:                Thu, Jun 06 2024   R-squared (Overall):              0.0168
Time:                        19:46:05   Log-likelihood                   -493.83
Cov. Estimator:            Unadjusted
                                        F-statistic:                      2.1783
Entities:                          75   P-value                           0.1146
Avg Obs:                       6.1467   Distribution:                    F(2,384)
Min Obs:                       2.0000
Max Obs:                       7.0000   F-statistic (robust):             2.1783
                                        P-value                           0.1146
Time periods:                      21   Distribution:                    F(2,384)
Avg Obs:                       21.952
Min Obs:                       17.000
Max Obs:                       25.000

                               Parameter Estimates
=================================================================================
               Parameter  Std. Err.     T-stat    P-value    Lower CI    Upper CI
---------------------------------------------------------------------------------
caloriesToday    -0.0004     0.0002    -2.0865     0.0376     -0.0007  -2.123e-05
screentime    -6.185e-07  5.942e-06    -0.1041     0.9172   -1.23e-05   1.106e-05
=================================================================================

F-test for Poolability: 4.8889
P-value: 0.0000
Distribution: F(74,384)

Included effects: Entity
```

Figure 17: Fixed Effects Results using daily calorie expenditure

B.4 Multiple Linear Regression Results

B.5 Multiple Linear Regression Results

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                 stress   R-squared:                       0.115
Model:                            OLS   Adj. R-squared:                  0.090
Method:                 Least Squares   F-statistic:                     4.657
Date:                Thu, 06 Jun 2024   Prob (F-statistic):             0.0125
Time:                        19:33:35   Log-Likelihood:                -81.270
No. Observations:                  75   AIC:                             168.5
Df Residuals:                      72   BIC:                             175.5
Df Model:                           2
Covariance Type:            nonrobust
=================================================================================
                    coef    std err          t      P>|t|      [0.025      0.975]
---------------------------------------------------------------------------------
Intercept         1.3420      0.561      2.393      0.019       0.224       2.460
screentime    -3.317e-05   1.43e-05     -2.325      0.023   -6.16e-05   -4.73e-06
caloriesToday    -0.0006      0.000     -2.158      0.034      -0.001   -4.89e-05
==============================================================================
Omnibus:                        1.630   Durbin-Watson:                   2.325
Prob(Omnibus):                  0.443   Jarque-Bera (JB):                1.325
Skew:                          -0.135   Prob(JB):                        0.516
Kurtosis:                       2.407   Cond. No.                     1.20e+05
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[2] The condition number is large, 1.2e+05. This might indicate that there are
strong multicollinearity or other numerical problems.
```

Figure 18: Multiple Linear Regression Results for Stress

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                    PHQ   R-squared:                       0.009
Model:                            OLS   Adj. R-squared:                 -0.018
Method:                 Least Squares   F-statistic:                    0.3296
Date:                Thu, 06 Jun 2024   Prob (F-statistic):              0.720
Time:                        19:42:15   Log-Likelihood:                -220.21
No. Observations:                  75   AIC:                             446.4
Df Residuals:                      72   BIC:                             453.4
Df Model:                           2
Covariance Type:            nonrobust
================================================================================
                   coef    std err          t      P>|t|      [0.025      0.975]
--------------------------------------------------------------------------------
Intercept        2.5174      3.576      0.704      0.484      -4.611       9.646
screentime   -3.661e-06    9.1e-05     -0.040      0.968      -0.000       0.000
caloriesToday    0.0015      0.002      0.805      0.423      -0.002       0.005
==============================================================================
Omnibus:                       23.985   Durbin-Watson:                   1.870
Prob(Omnibus):                  0.000   Jarque-Bera (JB):               32.815
Skew:                           1.443   Prob(JB):                     7.49e-08
Kurtosis:                       4.475   Cond. No.                     1.20e+05
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[2] The condition number is large, 1.2e+05. This might indicate that there are
strong multicollinearity or other numerical problems.
```

Figure 19: Multiple Linear Regression Results for Depression