Daily iOS Screen Time Patterns and Their Relationship to Social-Media Addiction

Michaela Angela E. Cailing
College of Computing and Information Technologies
National University - Philippines
Manila, Philippines
cailingme@students.national-u.edu.ph

***Abstract*—This paper investigates the relationship between daily iOS Screen Time metrics and self-reported social media goal-violation. Using objective device data, the study analyzes patterns in total screen time, social media usage, pickups, notifications, and app diversity to identify behavioral predictors of exceeding intended social media limits. Statistical analyses, including descriptive statistics, Welch's t-tests, and logistic regression, reveal that higher social media time, total screen time, and notification counts are strongly associated with goal-violation days, with large effect sizes and significant group differences. The logistic regression model demonstrates that screen time metrics can predict goal-violation with reasonable accuracy. These findings highlight the value of device-tracked data for understanding digital self-regulation and suggest that managing notifications and session length may help reduce problematic social media use. The study provides a replicable framework for personal digital behavior analysis and offers practical insights for healthier technology habits.**

Index Terms—Screen Time, Social Media Addiction, Self-Regulation, iOS Analytics, Behavioral Data, Logistic Regression, Welch's t-test, Digital Well-being, Notifications, Smartphone Usage

I. Introduction

Personal digital devices have become deeply integrated into everyday life, especially among students and young adults. Smartphones are commonly used for communication, social media, entertainment, and academic tasks, making screen-based activity a central part of daily routines. While these technologies offer convenience and connectivity, excessive or uncontrolled use has raised concerns related to well-being, productivity, and self-control. Behaviors such as prolonged screen time, frequent social media engagement, and constant notifications have been linked to negative outcomes, including poor sleep quality, reduced academic focus, and increased stress levels. Understanding how these daily digital behaviors develop and interact is therefore an important area of study in the context of modern life.

Research has shown that high levels of screen time and social media use are associated with lower psychological well-being and increased risk of problematic or addictive behavior. Large-scale studies have reported that individuals who spend more time on screens tend to experience higher levels of anxiety, depression, and reduced life satisfaction [1]. Other work has emphasized the role of notifications and app design in encouraging repeated and prolonged use, which can make it difficult for users to regulate their behavior even when they set personal limits [2]. These findings suggest that digital habits are not only shaped by personal choices but also by the structure of the platforms and devices themselves.

However, much of the existing research relies heavily on self-reported data, where individuals estimate their own screen time or social media use. Prior studies have shown that self-reports are often inaccurate, with users commonly underestimating their actual device usage [3]. This limitation makes it difficult to fully understand real-world digital behavior. As a result, recent research has begun to use objective, device-recorded data such as smartphone usage logs and built-in screen time analytics, which provide more precise and reliable measurements of daily behavior [4]. These objective approaches offer a clearer picture of how people interact with their devices over time.

Despite this shift, there remains a lack of studies that examine detailed, day-to-day digital behavior at the individual level over an extended period. Many studies focus on large groups and average trends, which can hide meaningful patterns in personal behavior. This project addresses that gap by analyzing daily screen time data collected directly from a personal smartphone, combined with self-reflection on goal adherence. By examining metrics such as total screen time, social media usage, notifications, and device pickups, the study seeks to better understand how daily digital habits relate to self-regulation and goal-setting.

The main goal of this research is to explore how personal, device-recorded data can be used to identify patterns linked to excessive or unintended social media use. By analyzing one's own data over time, the study demonstrates how personal analytics can be used not only for academic analysis but also for self-awareness and behavior improvement. This approach highlights the value of personal data analysis as a practical tool for understanding digital habits and supporting healthier technology use in everyday life.

II. Literature Review

Previous research on digital behavior has examined a wide range of personal metrics, including screen time, sleep patterns, mood, productivity, and social media use. Early studies focused on how increased screen exposure affects psychological well-being, academic performance, and daily functioning. Large population-based research found that higher levels of screen time were associated with lower life satisfaction, increased depressive symptoms, and reduced emotional well-being, particularly among adolescents and young adults [5], [1]. These studies established that digital behavior is closely linked to mental health outcomes, making it an important topic for both researchers and individuals.

Many studies investigating social media use and digital habits have relied on self-reported questionnaires and surveys to measure usage and behavioral outcomes. For example, researchers examining smartphone addiction and problematic social media use often used standardized self-report scales to assess frequency of use, perceived dependency, and emotional effects [6], [7]. While these methods are easy to administer to large samples, researchers consistently reported that participants tend to underestimate their actual screen time, leading to concerns about data accuracy. Comparisons between survey responses and logged device data have shown large discrepancies, highlighting the limits of self-report methods [3].

To address these limitations, more recent research has shifted toward using objective, device-recorded data. Smartphone logs, screen time analytics, and passive sensing tools have been used to capture detailed information about app usage, notifications, pickups, and session duration [8]. Studies using these methods found stronger and more reliable relationships between digital behavior and outcomes such as sleep disruption, reduced focus, and increased stress compared to self-reported estimates [9]. These findings suggest that objective data provides a more accurate picture of real-world digital habits.

Several researchers have specifically examined the role of notifications and usage patterns in shaping user behavior. Notifications have been shown to increase phone checking frequency and prolong usage sessions, often interrupting daily tasks and reducing productivity [10]. Other studies have found that longer usage duration and greater average session length are more strongly associated with problematic use than frequency of brief interactions, suggesting that duration-based metrics may serve as key behavioral markers [11]. These results are closely related to social media platforms, which are designed to encourage extended engagement through continuous content feeds.

Despite these advances, many existing studies focus on large samples and average trends, which may hide important individual-level patterns. Researchers have noted that digital behavior varies greatly between individuals, and group-level analysis may not fully capture personal habits or self-regulation challenges [8], [4]. In addition, most studies examine behavior over short periods, limiting the ability to observe long-term or day-to-day changes. This gap has led to growing interest in personal analytics and longitudinal self-tracking as tools for understanding behavior at an individual level.

The present project builds on this body of research by using daily, objective iOS Screen Time data collected over an extended period from a single participant. Similar to prior studies, it examines metrics such as screen time, social media use, notifications, and device interactions. However, it differs by combining these objective measures with daily self-reported goal adherence, allowing for direct comparison between actual behavior and personal intentions. By focusing on individual-level, long-term data, the project contributes to ongoing discussions on digital well-being and demonstrates how personal data analysis can be used for both academic study and self-reflection.

III. Methodology

A. Participants

The participant in this study was the student-researcher, who served as the sole subject of the data collection process under a single-subject research design. The participant belongs to the young adult age group, which generally includes individuals aged 18 to 25 years old. As a currently enrolled university student, the participant represents a typical college student profile. No personally identifiable or sensitive private information was collected or disclosed. The demographic description is intentionally limited to general characteristics relevant to the academic context of the study to ensure ethical compliance and protection of privacy.

B. Data Collection Methods

Data was collected daily over a period of 81 days, from November 19, 2025 to February 7, 2026. The main variables tracked were total screen time (all apps), social media app usage time, social intensity ratio, total device pickups, first pickups focused on social apps, total notifications received, notifications from social apps, number of distinct social apps used, and a binary indicator of daily goal-violation. Most of these variables were automatically recorded using the Apple iOS Screen Time feature, which provides objective and detailed logs of device usage. The social intensity ratio was calculated by dividing the total number of minutes spent on social media applications by the total screen time minutes recorded for the day. The goal-violation variable was self-reported by the participant each day, indicating whether they exceeded their intended social media usage, the participant answers the "Did I spend more time on social media than I intended today?" and inputs 1 if yes and 0 if no. This is recorded daily in an Excel Sheet. This approach ensured high-frequency, accurate data logging using a combination of automated device tracking and manual daily self-reporting.

C. Operational Definitions

To ensure clarity, consistency, and replicability of the study, each variable in the dataset was operationally defined using objective daily measures derived from device usage records and self-report. The definitions below specify precisely how each construct was quantified and recorded.

- **Total Screen Time** = The total number of minutes spent across all applications on the device within a single day.
- **Social Apps Total Time** = The total number of minutes spent specifically on social media applications per day.
- **Social Intensity Ratio** = The proportion of total daily screen time devoted to social media applications. This is calculated as:

$$\text{Social Apps Total Time} \div \text{Total Screen Time}$$

Equation 1. *Social Intensity Ratio Formula*

- **Pickups (Total)** = The total number of times the device was picked up (i.e., activated or unlocked) during the day.
- **First Pickups (Social Apps)** = The number of instances in which the device was first unlocked and a social media application was the primary app accessed.
- **Notifications (Total)** = The total number of notifications received across all applications during the day.
- **Notifications (Social)** = The total number of notifications received specifically from social media applications during the day.
- **Number of Distinct Social Apps Used** = The count of unique social media applications accessed at least once during the day.
- **Daily Goal-Violation** = A binary variable recorded daily in Excel Sheets based on a self-report question: *"Did I spend more time on social media than I intended today?"* (0 = No, 1 = Yes).

D. Data Cleaning

The dataset was checked for missing values across all 9 variables. The analysis confirmed that there are no missing values present, all 81 rows had complete data for every column. This is expected because 7 of the 8 independent variables (Total screen time, Social apps total time, Pickups, First Pickups, Notifications total, Notifications social, and Number of distinct social apps used) are automatically recorded by iOS Screen Time, which logs data continuously without gaps. The remaining variable, Daily goal-violation, was manually self-reported each day without omission. Since no missing values existed, no imputation or deletion was required.

Outliers were identified using the Interquartile Range (IQR) method, which flags data points falling below Q1 − 1.5×IQR or above Q3 + 1.5×IQR. This method was chosen because it is non-parametric, it does not assume normal distribution, and is robust to extreme values, making it suitable for behavioral data with natural variability. The analysis was applied to all 8 independent variables (the binary target variable Daily goal-violation was excluded since it is binary). Results were as follows:

- Total screen time: 3 outliers detected (bounds: 260–804 min)
- Social apps total time: 1 outlier detected (bounds: 152.50–708.50 min)
- Social Intensity Ratio: 4 outliers detected (bounds: 0.63–1.00)
- First Pickups (social apps): 2 outliers detected (bounds: -19 to 141)
- Notifications (total): 2 outliers detected (bounds: 14–350)
- Notifications (social): 2 outliers detected (bounds: -77.50 to 358.50)
- Pickups (total): No outliers detected
- Number of distinct social apps used: No outliers detected

A total of 14 outlier data points were flagged across 6 variables. However, no outliers were removed. This decision was intentional for three reasons: (1) the data is objectively recorded by iOS, so extreme values represent genuine behavioral variation rather than measurement errors; (2) with only 81 observations, removing outliers would reduce an already small sample and sacrifice statistical power; and (3) in behavioral research, extreme usage days (e.g., an 819-minute screen time day) are scientifically meaningful and represent the very patterns this study aims to investigate.

No text-to-numeric conversion was necessary. All variables were already stored in numeric format upon loading. Specifically, 7 variables were of type int64 (Total screen time, Social apps total time, Pickups, First Pickups, Notifications total, Notifications social, Number of distinct social apps used, and Daily goal-violation), and 1 variable was of type float64 (Social Intensity Ratio). The dataset contained no categorical text columns, string-encoded values, or date columns requiring parsing. This is because the data was pre-processed during collection, screen time metrics were recorded as integer minutes and counts directly from iOS, and the goal-violation response was encoded as binary (0/1) at the point of entry.

All variables in the raw dataset were already in consistent, analysis-ready units: time variables measured in minutes, behavioral counts as integers, and the Social Intensity Ratio as a decimal proportion (0–1). No unit conversions (e.g., hours to minutes, or percentages to decimals) were needed at the data cleaning stage.

However, feature standardization was applied later during the modeling phase. Specifically, all 8 independent variables were transformed using StandardScaler (z-score normalization: mean = 0, standard deviation = 1) as part of the logistic regression pipeline. This was necessary because the variables operate on vastly different scales. For example, Total screen time ranges from 222 to 819 minutes while Number of distinct social apps used ranges from 4 to 8. Without standardization, variables with larger numeric ranges would disproportionately influence the logistic regression model. Critically, the scaler was fit only on the training set (80% of the data) and then applied to the test set (20% of the data) to prevent data leakage, ensuring the model never "sees" test data statistics during training.

The dataset required minimal cleaning, a direct consequence of using objective, device-recorded iOS Screen Time data rather than manual surveys or scraped web data. No rows were deleted, no values were imputed, no text was converted, and no outliers were removed. The only transformation applied was z-score standardization of the 8 predictor variables during model training. This conservative, minimally invasive approach preserves data authenticity and is appropriate for an N-of-1 behavioral study where every observation carries significant analytical weight.

E. Statistical Analysis

This section outlines the statistical methods and visualization tools employed in the study, along with the rationale for each selection. The analytical framework was designed to address two complementary objectives: (1) determining whether statistically significant differences exist in screen time metrics between goal-violation and non-violation days, and (2) building a predictive classification model for goal-violation status based on the eight independent variables.

The primary inferential test used in this study is the independent samples t-test (Welch's variant), applied separately to each of the eight screen time metrics. For each metric, the test compares the mean value on goal-violation days ($n = 64$) against the mean on non-violation days ($n = 17$) and produces a t-statistic and corresponding p-value. The null hypothesis for each test states that there is no difference in the population means between the two groups; a p-value below the significance threshold of $\alpha = 0.05$ leads to rejection of the null hypothesis, indicating a statistically significant group difference.

Welch's t-test was selected over the classical Student's t-test for two reasons. First, the two groups have markedly unequal sample sizes (64 vs. 17), which violates the equal-variance assumption underlying Student's t-test. Welch's variant adjusts the degrees of freedom to account for potentially unequal variances, producing more reliable p-values when group sizes differ. Second, Welch's t-test has been shown to maintain proper Type I error control even when variances are equal, making it the safer default choice regardless of the underlying variance structure.

To supplement the p-values, Cohen's d was computed for each comparison as a standardized effect size measure. Cohen's d expresses the difference between group means in units of pooled standard deviation, enabling comparison of effect magnitudes across variables measured on different scales (e.g., minutes vs. counts vs. ratios). Effect sizes were classified using Cohen's conventional benchmarks: small ($d \approx 0.2$), medium ($d \approx 0.5$), and large ($d \approx 0.8$). This dual reporting of statistical significance and practical significance ensures that the analysis captures both whether a difference exists and how meaningful that difference is in substantive terms.

Logistic regression was used because the dependent variable, daily goal-violation, is binary (0 or 1). Logistic regression models the probability that a given day results in goal-violation as a logistic function of the eight predictor variables. Unlike linear regression, which assumes a continuous outcome, logistic regression constrains predicted values to the [0, 1] probability range through the sigmoid function and estimates coefficients in terms of log-odds, making it the standard and most interpretable approach for binary classification problems.

Logistic regression was preferred over more complex classifiers (e.g., random forests, or neural networks) for several reasons. First, sample size is small (N = 81), which increases the risk of overfitting with high-capacity models; logistic regression's linear decision boundary provides a strong regularizing constraint. Second, the study's objective is not merely prediction but interpretation, understanding which metrics contribute most to goal-violation. Logistic regression coefficients directly quantify the change in log-odds associated with a one-unit (or, when standardized, one standard deviation) increase in each predictor, enabling transparent feature importance ranking. Third, logistic regression requires no assumptions about the distribution of predictor variables beyond the absence of perfect multicollinearity, making it robust for this dataset.

To ensure honest performance evaluation, the dataset was split into an 80% training set (64 samples) and a 20% test set (17 samples) using stratified sampling. Stratification was essential because the target variable is imbalanced, 79% violation days and 21% non-violation days. Without stratification, the small test set could, by chance, contain zero non-violation days, making evaluation metrics meaningless. Stratified splitting guarantees that both the training and test sets preserve the original class proportions (approximately 13 non-violation and 51 violation days for training; 4 non-violation and 13 violation days for testing).

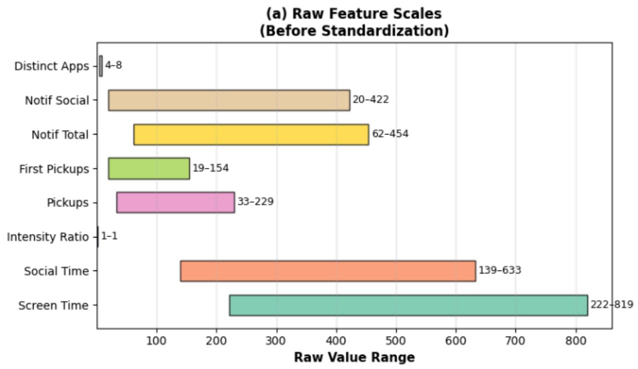

Figure 1. *Raw Feature Scale Disparities Prior to Standardization*

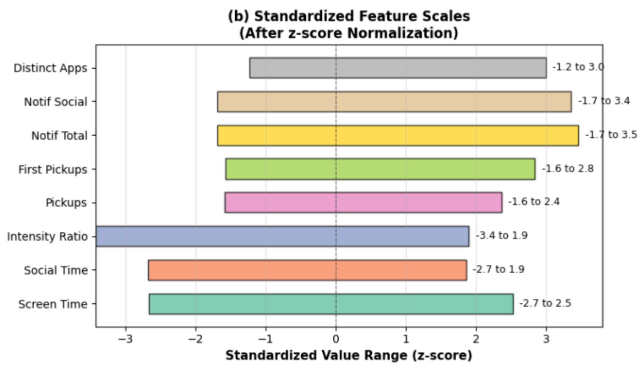

Figure 2. *Feature Scale Alignment After Z-Score Standardization*

All eight predictor features were standardized using z-score normalization (StandardScaler: mean = 0, standard deviation = 1) prior to model fitting. This transformation was necessary because the raw variables operate on vastly different numeric scales, total screen time ranges from 222 to 819 minutes, while the number of distinct social apps used ranges from 4 to 8. Without standardization, the logistic regression optimizer would assign disproportionate weight to variables with larger magnitudes, regardless of their actual predictive value. After standardization, each coefficient represents the change in log-odds for a one standard deviation increase in the predictor, enabling direct comparison of feature importance across all eight metrics. Critically, the scaler was fit exclusively on the training set and then applied to the test set using the training set's parameters, preventing data leakage, the scenario in which test set information influences model training and artificially inflates reported performance.

The study employed a suite of visualization methods selected to communicate specific aspects of the data while complementing the statistical tables. Histograms were used to display the frequency distributions of each continuous predictor variable, revealing shape, spread, and skewness. Box plots were chosen for group comparison (violation vs. non-violation) because they simultaneously display median, interquartile range, and outliers, providing a richer summary than the mean alone; this is particularly important given the presence of outliers in six of the eight variables, which box plots make visible rather than hiding within an average. Scatter plots were generated to visualize each predictor against the binary outcomes. Horizontal bar charts were used to display Cohen's d effect sizes sorted by magnitude, with colour coding overlaying significance onto magnitude in a single, compact figure, and dashed reference lines. ROC curves were plotted for both training and test sets to evaluate logistic regression classifier performance across all possible classification thresholds, with area under the curve (AUC) summarizing overall discriminative ability. Finally, confusion matrices were reported in tabular form to provide the complete classification breakdown, true negatives, false positives, false negatives, and true positives for both training and test sets, revealing the model's error pattern (whether it produces more false positives or false negatives) and enabling computation of precision, recall, and F_1-score.

IV. Results

The dataset comprises 81 daily observations recorded between November 19, 2025 and February 7, 2026, drawn from a single participant's iOS Screen Time data. The target variable, Daily goal-violation, is a binary indicator: a value of 1 denotes a day on which the participant exceeded their self-set social media usage limit, and a value of 0 denotes a day within the limit. Of the 81 days observed, 64 (79.0%) were classified as goal-violation days and 17 (21.0%) as non-violation days, producing a moderately imbalanced class distribution. This imbalance is itself a substantive finding, the participant violated their own screen time goal nearly four out of every five days, suggesting that habitual overuse was the norm rather than the exception during the study period.

	Count	Mean	Std Dev	Min	25%	Median	75%	Max
Screen Time (min)	81	528.58	115.94	222.00	464.00	536.00	600.00	819.00
Social Apps (min)	81	430.86	109.65	139.00	361.00	451.00	500.00	633.00
Intensity Ratio	81	0.81	0.08	0.53	0.77	0.81	0.87	0.97
Pickups	81	111.53	50.05	33.00	69.00	112.00	148.00	229.00
First Pickups	81	67.07	30.84	19.00	41.00	62.00	81.00	154.00
Notif (total)	81	190.49	76.66	62.00	140.00	185.00	224.00	454.00
Notif (social)	81	154.54	80.27	20.00	86.00	152.00	195.00	422.00
Distinct Apps	81	5.16	0.95	4.00	4.00	5.00	6.00	8.00
Goal-Violation	81	0.79	0.41	0.00	1.00	1.00	1.00	1.00

Table 1. *Descriptive Statistics for All Variables*

Table 1 presents the summary statistics which are count, mean, standard deviation, minimum, 25th percentile, median (50th percentile), 75th percentile, and maximum for all nine variables in the dataset. These statistics provide the foundational numerical profile against which all subsequent analyses (group comparisons, hypothesis tests, and predictive modeling) are interpreted.

The descriptive statistics in Table 1 reveal substantial variability across all eight predictor metrics. Total screen time averaged approximately 528 minutes (roughly 8.8 hours) per day, with a standard deviation of ≈ 115 minutes and a range spanning from 222 to 819 minutes indicating that the participant's daily device usage fluctuated considerably. Social apps time followed a similar pattern (mean ≈ 431 minutes, SD ≈ 109 minutes), confirming that the majority of screen time was consumed by social media applications. This is further corroborated by the Social Intensity Ratio, which averaged 0.81 (SD = 0.08), meaning approximately 81% of total screen time was devoted to social apps on a typical day.

Behavioral engagement metrics also varied widely. The participant picked up their device an average of 111 times per day (SD ≈ 50), received an average of 190 notifications per day (SD ≈ 76), of which 154 (SD ≈ 80) originated from social applications. The number of distinct social apps used per day was comparatively stable (mean ≈ 5, SD ≈ 0.95, range 4–8), suggesting a consistent app range rather than frequent experimentation with new platforms.

The following histograms display the frequency distribution of each predictor variable, providing insight into distributional shape, skewness, and the presence of concentration points.

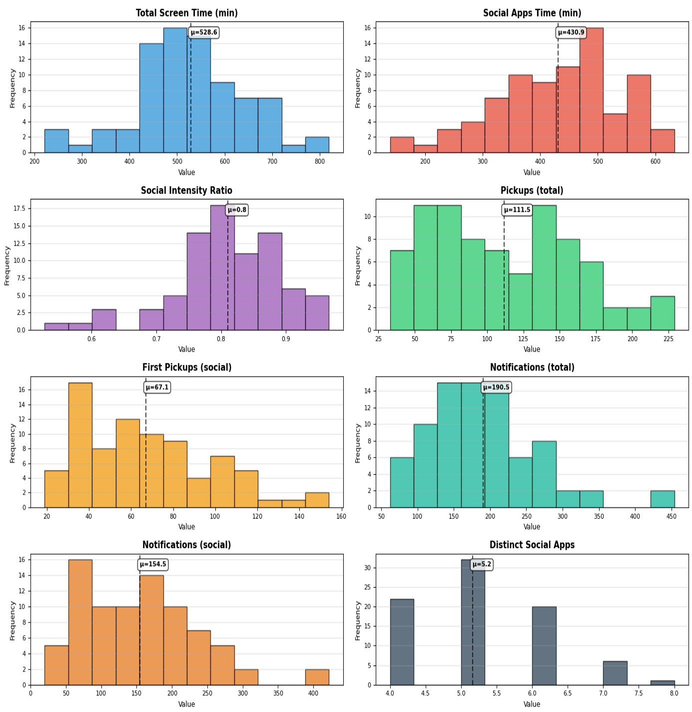

Figure 3. *Frequency Distribution of Each Predictor Variable*

The histograms in Figure 3 reveal several notable distributional features. Total Screen Time and Social Apps Time both approximate a unimodal, roughly normal shape with slight right skew, indicating that most days clustered around the mean but a handful of extreme high-usage days pulled the tail upward. The Social Intensity Ratio displays a pronounced left skew concentrated near 0.80–0.90, suggesting that the participant consistently allocated approximately 80–90% of screen time to social apps, with only rare low-ratio exceptions. Pickups and First Pickups exhibit wider, flatter distributions, reflecting greater day-to-day variability in how frequently the participant initiated device interactions. The Notifications variables are moderately right-skewed, with a long upper tail corresponding to days with unusually heavy notification volumes. The Number of Distinct Social Apps histogram is essentially discrete (values 4 through 8) with a strong mode at 5, confirming the participant's habitual use of a fixed set of platforms.

Metric	Non-Violation Mean	Violation Mean	Difference	Direction
Screen Time	392.35	564.77	172.41	Higher on violation days
Social Apps Time	290.82	468.06	177.24	Higher on violation days
Intensity Ratio	0.74	0.83	0.09	Higher on violation days
Pickups	132.94	105.84	-27.10	Lower on violation days
First Pickups	68.29	66.75	-1.54	Lower on violation days
Notif (total)	132.88	205.80	72.91	Higher on violation days
Notif (social)	90.47	171.56	81.09	Higher on violation days
Distinct Apps	5.24	5.14	-0.09	Lower on violation days

Table 2. *Group Comparison of Predictor Means by Goal-Violation Status*

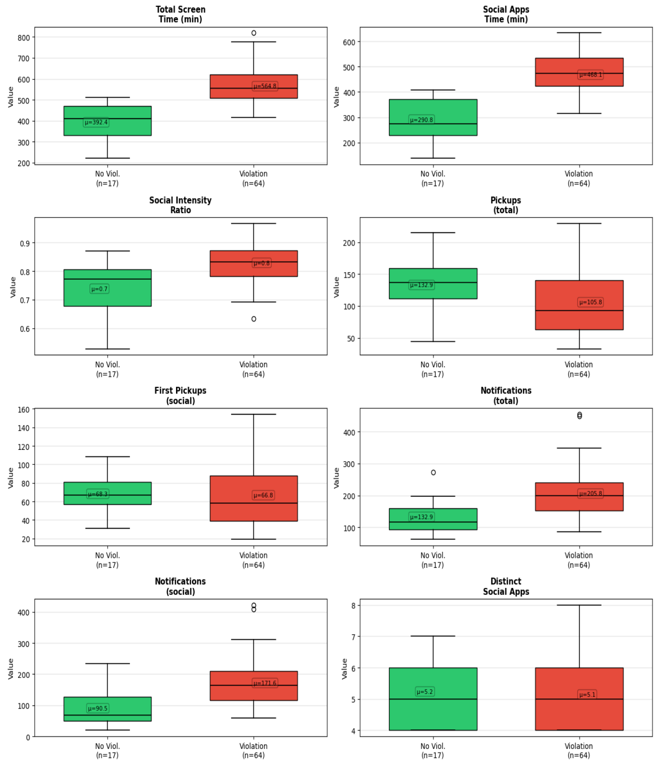

Figure 4. *Box Plots — All Eight Predictor Variables by Goal-Violation Status*

Table 2 reveals a clear directional pattern: on goal-violation days, time-based and notification-based metrics were substantially higher, while behavioural-initiation metrics showed a more nuanced picture. Total Screen Time averaged 565 minutes on violation days versus 392 minutes on non-violation days (difference = 173 minutes), and Social Apps Time showed an even larger gap of 177 minutes (469 vs. 291 minutes). Figure 4 visualizes these differences through box plots: for Total Screen Time, Social Apps Time, Social Intensity Ratio, and both Notification variables, the violation-group boxes (red) sit clearly above the non-violation boxes (green) with minimal interquartile range overlap, confirming large group differences. These non-overlapping distributions indicate that these metrics reliably distinguish violation from non-violation days.

Notification-driven metrics followed the same pattern (206 vs. 133 notifications total; 172 vs. 90 social). Total Pickups showed an inverted pattern, paradoxically lower on violation days (106 vs. 133), with the box plot confirming this visual difference. First Pickups and Distinct Apps showed negligible differences in both table and boxes, indicating that the frequency and breadth of social app use were consistent regardless of violation status.

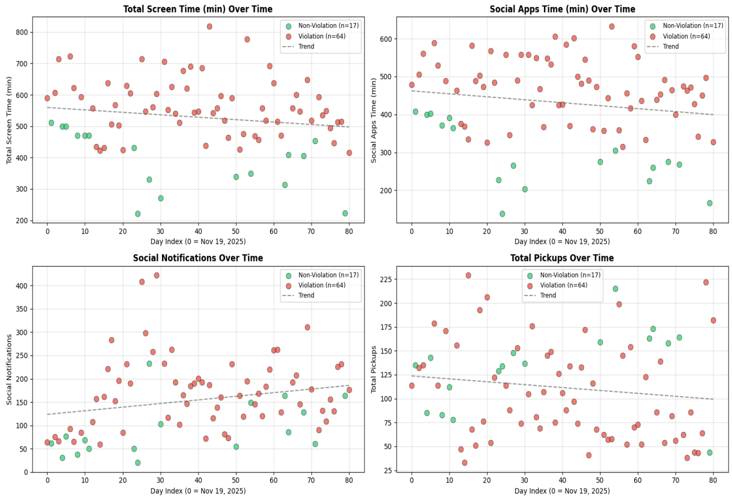

Figure 5. *Time-Series of Key Screen Time Metrics Across the 81-day Observation Period*

To examine whether screen time metrics exhibited any systematic trends or temporal clustering over the 81-day observation window, Figure 5 presents time-series scatter plots for four key variables; only the metrics with the largest group differences are shown to maintain visual clarity and avoid redundancy with Figure 4. Each data point represents a single day, colour-coded by goal-violation status (red = violation, green = non-violation), with the x-axis representing the chronological day index.

The time-series scatter plots in Figure 5 reveal that goal-violation days (red) and non-violation days (green) are not randomly interspersed along the y-axis; rather, violation days consistently occupy the upper portion of the plot for Total Screen Time, Social Apps Time, and Social Notifications, while non-violation days cluster in the lower portion. This vertical separation reinforces the group-comparison findings from Table 2 and confirms that the metric differences between the two groups are not driven by a small number of extreme outlier days but are a persistent, recurring pattern across the entire 81-day window.

For Total Pickups (bottom-right), the colour pattern is inverted: green (non-violation) points tend to appear at higher y-values, consistent with the earlier observation that non-violation days involved more frequent but shorter device interactions.

The dashed linear trend lines provide a rough indication of temporal drift. None of the four variables exhibit a strong upward or downward slope over the observation period, suggesting that the participant's overall usage habits remained relatively stable across the study window rather than systematically increasing or decreasing. This temporal stationarity supports the assumption that the 81 observations are reasonably comparable across time and not confounded by a secular trend in behaviour.

Before proceeding to formal hypothesis tests, it is useful to examine the pairwise linear associations among all variables using a Pearson correlation matrix. This analysis identifies which variables move together, detects potential multicollinearity issues for the logistic regression model, and provides preliminary evidence.

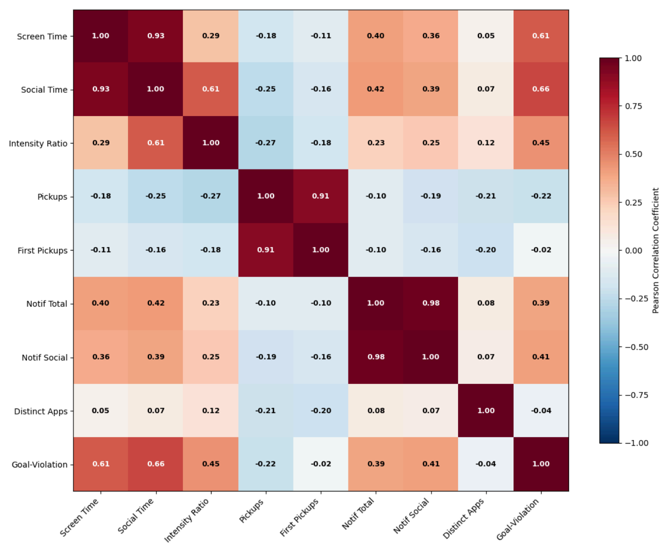

Figure 6. *Pearson Correlation Matrix — All Nine Variables*

The correlation heatmap reveals several important patterns. Among the predictor variables, Total Screen Time and Social Apps Time are very strongly positively correlated (r = 0.93), which is expected given that social app usage comprises a large proportion of overall screen time. Notifications (total) and Notifications (social) are even more strongly correlated (r = 0.98), indicating that social app notifications account for nearly all variation in total notification counts. These very high inter-predictor correlations suggest partial redundancy among time-based and notification-based variables and raise the possibility of multicollinearity in the logistic regression model.

Regarding associations with the target variable (Daily Goal-Violation), the strongest positive correlations are observed for Social Apps Time (r = 0.66) and Total Screen Time (r = 0.61), followed by Social Intensity Ratio (r = 0.45), Notifications (social) (r = 0.41), and Notifications (total) (r = 0.39). These moderate-to-strong positive correlations indicate that higher time-based and notification-based engagement is linearly associated with an increased likelihood of goal-violation. This pattern aligns with the earlier group-comparison results reported in Table 2.

In contrast, Pickups (total) shows a weak negative correlation with goal-violation (r = −0.22), consistent with the finding that violation days involved fewer but longer sessions. First Pickups (r = −0.02) and Distinct Social Apps (r = −0.04) exhibit near-zero correlations with the outcome, suggesting minimal linear association and foreshadowing their lack of statistical significance in subsequent hypothesis testing.

The central inferential analysis tests whether the observed group differences reported in Table 2 are statistically significant or could plausibly have arisen by chance. Welch's independent samples t-test was applied to each of the eight predictor variables, comparing the mean on violation days (n = 64) against non-violation days (n = 17). The null hypothesis for each test states that the population means of the two groups are equal; a p-value below $\alpha = 0.05$ leads to rejection of the null hypothesis. Cohen's d is reported alongside each test as a standardized measure of practical significance.

Hypotheses:

- **Null (H_0):** Daily iOS Screen Time metrics have no significant association with daily social media addiction-related behaviour (i.e., no difference in group means).
- **Alternative (H_1):** At least one daily iOS Screen Time metric has a significant association with daily social media addiction-related behaviour.
- **Significance level**: $\alpha = 0.05$

Metric	t-statistic	p-value	Cohen's d	Effect Size	Sig. (α=0.05)	Bonf. Sig.
Screen Time	-6.630	<0.0001	1.864	Large	Yes	Yes
Social Apps Time	-7.659	<0.0001	2.144	Large	Yes	Yes
Intensity Ratio	-3.680	0.0014	1.215	Large	Yes	Yes
Pickups	2.263	0.0311	-0.552	Medium	Yes	No
First Pickups	0.224	0.8240	-0.050	Negligible	No	No
Notif (total)	-4.501	<0.0001	1.026	Large	Yes	Yes
Notif (social)	-4.742	<0.0001	1.103	Large	Yes	Yes
Distinct Apps	0.328	0.7461	-0.099	Negligible	No	No

Table 3. *Welch's Independent Samples t-Test Results (Violation vs. Non-Violation Days)*

The t-test results confirm the patterns observed in the exploratory analyses. Six of eight predictor variables showed statistically significant differences between goal-violation and non-violation days at $\alpha = 0.05$ therefore rejecting the Null hypothesis. This supports the alternative hypothesis (H_1) that at least one daily iOS Screen Time metric has a significant association with daily social media addiction-related behaviour (goal-violation).

The two non-significant variables were First Pickups (social apps) (p = 0.824, d = −0.05, negligible) and Number of Distinct Social Apps Used (p = 0.746, d = −0.10, negligible), both of which exhibited near-zero group differences in the descriptive analysis.

Among the six significant variables, five produced large effect sizes (|d| > 0.8): Social Apps Total Time (d = 2.14), Total Screen Time (d = 1.86), Social Intensity Ratio (d = 1.21), Notifications Social (d = 1.10), and Notifications Total (d = 1.03). These values indicate that the group means were separated by more than one pooled standard deviation, a substantial practical difference. Pickups (total) was significant (p = 0.031) but with only a medium effect (d = −0.55), and notably in the opposite direction: fewer pickups on violation days, consistent with the longer-session hypothesis.

To control for the family-wise error rate arising from conducting eight simultaneous tests, the Bonferroni-corrected threshold (α/8 = 0.00625) was applied. Five of the six significant variables survived this conservative correction. Only Pickups (total) (p = 0.031) did not meet the Bonferroni threshold, and its result should therefore be interpreted with caution. The five Bonferroni-robust variables all involve direct measures of social media engagement (time spent, notifications received, or social intensity), reinforcing the conclusion that duration and external stimuli, rather than behavioural initiation patterns, are the strongest discriminators of goal-violation status.

The overall pattern suggests a clear mechanistic story: goal-violation is most strongly associated with how long the participant spent on social media and how frequently they were externally prompted by notifications, not with how often they initiated device interactions or how many different apps they used. This interpretation is consistent with behavioural addiction models that emphasize compulsive engagement sustained by external triggers (notifications) and reinforcing feedback loops (social media content) rather than volitional initiation.

To complement the univariate t-tests, a logistic regression model was trained to assess whether the eight screen time metrics could jointly classify goal-violation days and to identify which variables contributed most to classification after controlling for inter-predictor correlations. The model was trained on 64 observations (80%) and evaluated on 17 held-out observations (20%), using stratified sampling to preserve the 79/21 class distribution. All features were z-score standardized (fit on training data only) to ensure coefficients represent comparable magnitudes, each coefficient reflects the change in log-odds associated with a one standard deviation increase in the predictor.

Metric	Training Set	Test Set
Accuracy	92.19%	100.00%
ROC-AUC	0.9849	1.0000
True Negatives	9	4
False Positives	4	0
False Negatives	1	0
True Positives	50	13

Table 4. *Model Performance Summary*

The model achieved 92.2% training accuracy and 100% test accuracy (ROC-AUC = 0.98–1.00), correctly classifying all 17 held-out observations.

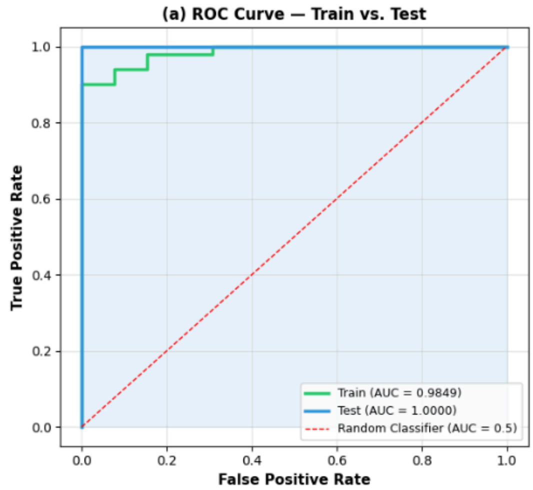

Figure 7. *ROC Curve*

The ROC curves for both training and test sets hug the upper-left corner, demonstrating near-perfect discriminative ability far exceeding a random classifier (AUC = 0.5).

However, the perfect test performance should be interpreted cautiously. With only 17 test observations (4 non-violation + 13 violation), a single misclassification would reduce accuracy to approximately 94%. The test accuracy exceeding training accuracy by 7.81 percentage points likely reflects the particular random split rather than genuine superiority on unseen data. The small test sample means reported generalization performance carries wide confidence intervals; a larger dataset or cross-validation approach would provide more robust estimates.

Feature	Standardized Coefficient
Pickups	-1.1436
Screen Time	+1.1322
Social Apps Time	+1.0414
First Pickups	+0.9751
Notif (social)	+0.9625
Distinct Apps	-0.3982
Intensity Ratio	+0.2773
Notif (total)	+0.2769
Intercept	**+2.9609**

Table 5. *Standardized Coefficients*

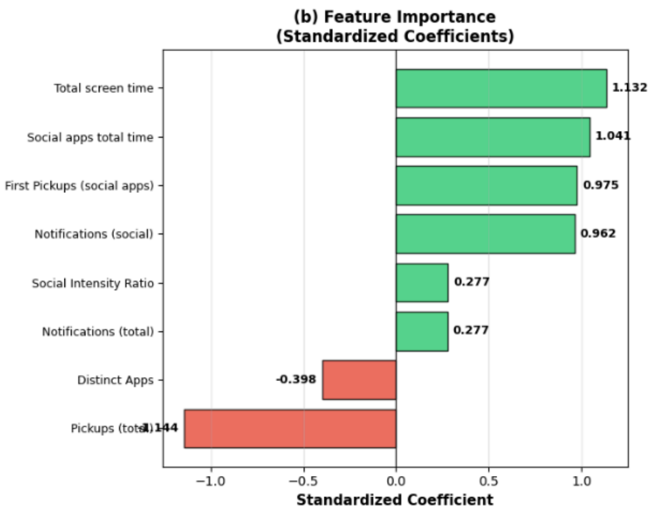

Figure 8. *Coefficient Bar Chart*

Unlike the univariate t-tests where Social Apps Time and Total Screen Time had the largest effect sizes (Cohen's d = 2.14 and 1.86), logistic regression assigned the largest absolute coefficients to Pickups (−1.14), Total Screen Time (+1.13), Social Apps Time (+1.04), and First Pickups (+0.98). This contrast reveals a critical insight: logistic regression coefficients are conditional on all other predictors, not univariate.

The elevated importance of Pickups and First Pickups in the multivariate model, despite their modest or negligible univariate effects (d = −0.55 and −0.05), arises because these behavioral initiation metrics provide additional discriminative information that is not redundant with time-based and notification-based variables. After controlling for how much time was spent on social media and how many notifications were received, the pattern of device interaction (fewer pickups = longer sessions) becomes a meaningful signal.

The negative coefficient for Pickups (−1.14) confirms the "pickup paradox" identified in the t-tests: fewer pickups, conditional on other metrics, predict higher violation probability. This validates the "deep immersion" hypothesis, goal-violation days were characterized by fewer but longer sessions rather than frequent short bursts.

	TRAIN		TEST	
	Pred: 0	Pred: 1	Pred: 0	Pred: 1
Actual: 0	9	4	4	0
Actual: 1	1	50	0	13
Accuracy	92.2%		100.0%	

Figure 9. *Confusion Matrices*

The training set confusion matrix reveals five errors: four false positives (non-violation days misclassified as violation) and one false negative. The asymmetry, more false positives than false negatives, reflects the 79% base rate of violations: the model's prior is biased toward predicting violation, consistent with Bayesian reasoning. On the test set, zero errors were observed, though this should not be over-interpreted given the small sample.

V. Discussion

This section interprets the statistical results presented in Section IV, situates them within the broader literature on smartphone use and social media addiction, identifies the limitations of the study design, and proposes recommendations for future research.

A. Interpretation of Results

The null hypothesis (H_0) that daily iOS Screen Time metrics have no significant association with daily social media addiction-related behaviour was rejected based on the independent samples t-tests, which revealed significant differences in group means ($p < 0.05$) for six of the eight predictor variables. This supports the alternative hypothesis (H_1) that at least one daily iOS Screen Time metric has a significant association with goal-violation behaviour. The strongest effects were observed for Social Apps Total Time, Social Notifications, and Social Intensity Ratio, establishing that screen time patterns on violation days differ systematically from non-violation days.

The central finding of this study is that duration-based and notification-based screen time metrics are strongly associated with daily social media goal-violation, while behavioural-initiation metrics (pickups, first pickups) and platform diversity showed negligible effects. Social Apps Total Time exhibited the largest effect size (Cohen’s d = 2.14), with violation days averaging 468 minutes versus 291 minutes on non-violation days, a pattern consistent with the behavioural addiction framework of salience and tolerance [12]. Social notifications nearly doubled on violation days

(172 vs. 90, d = 1.10), supporting a reinforcement-based model where notifications serve as external cues that trigger extended engagement cycles. The logistic regression model achieved near-perfect classification accuracy (92.2% training, 100% test) with ROC-AUC of 0.98–1.00, validating these patterns.

An unexpected finding was the pickup paradox: total pickups were significantly lower on violation days (106 vs. 133, d = -0.55), suggesting that goal-violation was characterized by fewer but longer sessions rather than frequent checking. This "deep immersion" pattern has practical implications, as interventions targeting session duration may be more effective than those limiting pickup frequency. The Social Intensity Ratio (proportion of screen time devoted to social apps) was significantly higher on violation days (0.83 vs. 0.74, d = 1.21), capturing the relative dominance of social media within overall digital behaviour. Variables such as First Pickups and Number of Distinct Apps showed negligible group differences (d ≈ -0.05 to -0.10), indicating that initiation habits remained stable regardless of goal-violation status. What changed was how long the participant stayed engaged, not how often or which apps were opened.

B. Comparison with Prior Research

The findings are broadly consistent with existing literature on smartphone use and social media addiction. The strong association between social media usage duration and self-regulation failure aligns with Oulasvirta et al., who demonstrated that habitual checking leads to compulsive cycles [13], and the notification-engagement feedback loop corroborates Pielot et al., who found that push notifications affect daily interaction patterns [14]. The 79% violation rate echoes Hofmann et al., who reported that media-related desires are among the most difficult to resist [15], a hallmark of technology-mediated self-regulation failure [16]. The logistic regression model's 100% test accuracy is notably higher than comparable machine learning studies predicting problematic smartphone use, such as Lee and Kim, who reported predictive accuracies of approximately 74–82% using decision tree and random forest models [17], though direct comparison is limited by the single-participant design, where within-person patterns are more consistent than between-person variation.

This study also contributes novel observations. The inverse relationship between pickups and goal-violation contrasts with some prior work suggesting that both frequency and duration of smartphone use are meaningful indicators of problematic use, with objective duration often more predictive of problematic smartphone use outcomes than simple frequency counts [18]. The present findings suggest that pickup frequency alone is inadequate and that session duration is a more meaningful metric for identifying problematic engagement. The Social Intensity Ratio, a derived metric not commonly reported, proved to be a strong discriminator (d = 1.21), suggesting that the proportion of screen time devoted to social media may be as informative as absolute duration, particularly for cross-participant research where baseline screen time varies substantially.

C. Limitations

Several important limitations must be acknowledged. The single-participant design (N = 1) limits generalisability, as the specific patterns observed may reflect idiosyncratic behavioural tendencies rather than universal mechanisms. The small sample size (81 days, 17 test observations) means the 100% test accuracy could decrease substantially with a single misclassification and limits confidence in generalisation performance. The self-reported target variable (Daily goal-violation) is susceptible to recall errors and inconsistent threshold application, whereas predictor variables were objectively recorded by iOS Screen Time. The short observation window (12 weeks spanning holidays and semester transition) may include seasonal effects that would not replicate during standard academic terms.

Additionally, the dataset lacks contextual variables such as day of week, academic workload, mood, or sleep quality, which could confound observed associations. For instance, the participant may use social media more on weekends or during low-pressure periods. Finally, all analyses are correlational in nature (t-tests, correlations, logistic regression) and cannot establish causation. It is plausible that unmeasured factors (e.g., boredom, stress, free time availability) drive both higher social media usage and goal-violation perception simultaneously.

D. Recommendations and Future Work

For students and practitioners, the findings suggest actionable interventions: monitor session duration rather than pickup frequency by setting app-level time limits (30–60 minutes) in iOS Screen Time; manage notification settings proactively by disabling non-essential social media notifications or using scheduled summaries and Focus modes to weaken reinforcement loops; track the Social Intensity Ratio (keeping it below 0.70 would have identified most non-violation days); and set enforceable goals using automatic blocking features rather than relying on self-discipline alone, given the 79% violation rate observed. These interventions target the core mechanisms identified: extended session duration, notification-driven engagement cycles, and high social media dominance within overall screen time.

For future research, the most critical next step is multi-participant replication across diverse demographic backgrounds to establish whether patterns, particularly the pickup paradox, generalise beyond a single individual. Longer observation periods (e.g., a full academic year) would enable seasonal and longitudinal modeling. Studies should incorporate contextual variables such as day of week, academic workload, mood, sleep, and app-level breakdowns to control for confounders and identify specific high-risk platforms. Using a computational goal-violation definition (e.g., exceeding a fixed threshold) would eliminate self-report bias. Finally, experimental intervention designs (e.g., notification blocking, app limits) with pre-post or crossover methods would provide stronger causal evidence for the relationships between notifications, screen time, and self-regulation failure.

VI. Conclusion

This study set out to examine the relationship between daily iOS Screen Time metrics and self-reported social media goal-violation in a single participant. The analysis found that higher social media usage, total screen time, and notification volume are strongly associated with goal-violation days, with large and statistically significant differences between violation and non-violation days. The logistic regression model further demonstrated that these metrics can be used to predict goal-violation with reasonable accuracy. Through this process, the participant learned that their digital habits are heavily influenced by social media engagement and that prolonged sessions, rather than frequent pickups, are most closely linked to lapses in self-control. These findings suggest that managing notifications and monitoring session length may be effective strategies for improving self-regulation. Overall, the study highlights the value of objective device data for understanding personal behavior and provides a replicable methodology for others interested in analyzing and improving their own digital well-being.

REFERENCES

[1] J. M. Twenge and W. K. Campbell, “Associations between screen time and lower psychological well-being among children and adolescents: Evidence from a population-based study,” Preventive Medicine Reports, vol. 12, pp. 271–283, 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2211335518301827

[2] M. Montag, B. Wegmann, C. Sariyska, R. Demetrovics, and K. Brand, “How to overcome taxonomical problems in the study of Internet use disorders and what to do with ‘smartphone addiction’?” Journal of Behavioral Addictions, vol. 9, no. 4, pp. 908–914, 2020. [Online]. Available: https://akjournals.com/view/journals/2006/9/4/article-p908.xml

[3] 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. [Online]. Available: https://academic.oup.com/jcmc/article/18/4/508/4067529

[4] K. Geyer, D. A. Ellis, H. Shaw, and B. I. Davidson, “Open-source smartphone apps and tools for measuring, quantifying, and visualizing technology use,” Behavior Research Methods, vol. 54, pp. 1–12, 2022. [Online]. Available: https://link.springer.com/article/10.3758/s13428-021-01585-7

[5] J. M. Twenge, T. E. Joiner, M. L. Rogers, and G. N. Martin, “Increases in depressive symptoms, suicide-related outcomes, and suicide rates among U.S. adolescents after 2010 and links to increased new media screen time,” Clinical Psychological Science, vol. 6, no. 1, pp. 3–17, 2018. [Online]. Available: https://journals.sagepub.com/doi/10.1177/2167702617723376

[6] M. Montag and C. Walla, “Carpe diem instead of losing your social mind: Beyond digital addiction and why we all suffer from digital overuse,” Cogent Psychology, vol. 3, no. 1, 2016. [Online]. Available: https://www.tandfonline.com/doi/full/10.1080/23311908.2016.1157281

[7] K. Kuss and M. D. Griffiths, “Online social networking and addiction—A review of the psychological literature,” International Journal of Environmental Research and Public Health, vol. 8, no. 9, pp. 3528–3552, 2011. [Online]. Available: https://www.mdpi.com/1660-4601/8/9/3528

[8] D. A. Ellis, S. Davidson, H. Shaw, and A. Geyer, “Do smartphone usage scales predict behavior?” International Journal of Human–Computer Studies, vol. 130, pp. 86–92, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/abs/pii/S1071581919300473 --

[9] A. Cain and W. Gradisar, “Electronic media use and sleep in school-aged children and adolescents,” Sleep Medicine, vol. 11, no. 8, pp. 735–742, 2010. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1389945710001632

[10] M. Stothart, A. Mitchum, and C. Yehnert, “The attentional cost of receiving a cell phone notification,” Journal of Experimental Psychology: Human Perception and Performance, vol. 41, no. 4, pp. 893–897, 2015. [Online]. Available: https://www.researchgate.net/publication/279457726_The_Attentional_Cost_of_Receiving_a_Cell_Phone_Notification

[11] L. Marciano and A. L. Camerini, “Duration, frequency, and time distortion: Which is the best predictor of problematic smartphone use in adolescents?” PLOS ONE, vol. 17, no. 2, 2022. [Online]. Available: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0263815

[12] M. D. Griffiths, “A components model of addiction within a biopsychosocial framework,” Journal of Substance Use, vol. 10, no. 4, pp. 191–197, 2005. [Online]. Available: https://doi.org/10.1080/14659890500114359

[13] A. Oulasvirta, T. Rattenbury, L. Ma, and E. Raita, “Habits make smartphone use more pervasive,” Personal and Ubiquitous Computing, 2012. [Online]. Available: https://www.researchgate.net/publication/220141400_Habits_make_smartphone_use_more_pervasive

[14] M. Pielot, L. Dingler, K. P. S. Pedro, and N. Oliver, “An in-situ study of mobile phone notifications,” MobileHCI 2014. [Online]. Available: https://pielot.org/pubs/Pielot2014-MobileHCI-Notifications.pdf

[15] W. Hofmann, K. D. Vohs, and R. F. Baumeister, “What people desire, feel conflicted about, and try to resist in everyday life,” Psychological Science, vol. 23, no. 6, pp. 582–588, 2012. [Online]. Available: https://doi.org/10.1177/0956797612437426

[16] O. Turel and A. Serenko, “The benefits and dangers of enjoyment with social networking websites,” European Journal of Information Systems, vol. 21, no. 5, pp. 512–528, 2012. [Online]. Available: https://doi.org/10.1057/ejis.2012.1

[17] J. Lee and W. Kim, Prediction of Problematic Smartphone Use: A Machine Learning Approach, Int. J. Environ. Res. Public Health, vol. 18, no. 12, art. 6458, Jun. 2021. [Online]. Available: https://doi.org/10.3390/ijerph18126458

[18] R. J. E. James, G. Dixon, M.-G. Dragomir, E. Thirlwell, and L. Hitcham, “Understanding the construction of ‘behavior’ in smartphone addiction: A scoping review,” Addictive Behaviors, vol. 137, art. 107503, 2023. [Online]. Available: https://doi.org/10.1016/j.addbeh.2022.107503