Lay Len Ching, Ethan Taylor, Brandon Zhou, Justin Poser
Professor Schwardmann
Behavioral Economics
04 December 2023

Paper Report on Digital Addiction

Section I: Research Design

Motivation

In recent years, the influence of smartphones, video games, and social media has raised concerns among researchers, who have drawn parallels to the addictive nature of well-known vices such as cigarettes, alcohol, and gambling. The paper "Digital Addiction," authored by Hunt Allcott, Matthew Gentzkow, and Lena Song, delves into an exploration of the addictive qualities of digital technology. Notably, the realm of digital self-control presents a unique landscape where market forces have given rise to commitment devices, including app blockers and screen time trackers. However, research proposes that there exists an unmet demand for such tools, suggesting a critical need for effective solutions to address the challenges posed by digital addiction.

Overview

The goal of this research was to provide empirical evidence on habit formation, temptation, and the perception of both in the context of digital technology. The study took place from March 22nd to July 26th, 2020, and involved an intake questionnaire and four subsequent surveys. Participant recruitment occurred between March 22nd and April 8th through Facebook and Instagram ads that displayed no information about the research question. 0.8% of users exposed to the ads clicked on them (26,101 out of 3,271,165 unique users), aligning with the average Facebook ad click-through rates. Of 18,589 eligible individuals, 8,514 consented and 5,320 successfully installed Phone Dashboard.[1] Participant ages ranged from 18-64 and they all had an Android device[2] as their primary phone,

Participants filled out four surveys about their current and expected use of six chosen technologies/services: Facebook, Instagram, Twitter, Snapchat, web browsers, and YouTube (FITSBY).

[1] Phone Dashboard is an app developed by Audacious Software specifically for this experiment. It records the app that is in the foreground of a smartphone every five seconds when the screen is on. Users can see their cumulative screen time by both day and week on the app.
[2] The reason the experiment only included Android users is because Phone Dashboard could not function properly on iOS.

The surveys were scheduled at three-week intervals on Sundays, from April 12 to June 14, starting with a baseline survey focusing on demographics. 4,038 participants completed Survey 1. Of these individuals, 1,912 were disqualified due to either reporting the use of another app to limit phone use or failing a data quality check. The remaining 2,162 participants were invited to survey 2, of whom 89% completed the study. All participants received $5 for baseline completion and $25 for subsequent surveys.

Participants were randomly assigned to the two treatment conditions, bonus and limit. Within these treatment conditions participants were further randomly assigned. In the bonus treatment, participants were assigned to the bonus, bonus control, or MPL group with probabilities of 25%, 75%, and 0.2%, respectively. In the limit treatment, participants were randomized into the limit or limit control groups with probabilities of 60% and 40%, respectively. The intersection of the bonus control groups and limit control groups is referred to as the control group. Randomization was balanced within eight strata defined by above-versus below-median baseline FITSBY use, restriction index, and addiction index. Treatment began after survey 2.

It is important to note here that the attrition rate for this study was exceptionally low. Several strategies contributed to this, including back-loading survey payments and explicit communication regarding the potential impact of dropping out on the integrity of the study. In addition, daily reminders were sent out for six days after the survey had been fielded, with additional payments offered after four days for completing all remaining surveys. Reminder emails were also sent to participants who had yet to respond to consecutive text messages.

Treatments

The bonus treatment aimed to explore projection bias and actual habit formation. Survey 2 clarified the bonus conditions, with participants informed about potential rewards based on reducing daily FITSBY screentime below a Bonus Benchmark [X][3] for 3 weeks. After the subsequent MPL, the bonus group was notified of their selection for the bonus in Period 3, while the bonus control group was informed otherwise. Analysis of the bonus group's responses in Periods 2, 3, 4, and 5 provide insights into projection bias, price response, and the magnitude of decay of habit.

The limit treatment explored self-control through Phone Dashboard's time restrictions, which began after Survey 2. This feature allowed participants to set daily time limits for each mobile app. Changes to these limits took effect the following day. The app sent push notifications as the daily time

[3] Bonus Benchmark [X] was set as the participants' average FITSBY hours per day from Period 1, rounded up to the nearest integer.

limit approached, with the option for participants to snooze.[4] To reduce attrition and uninstallation, Phone Dashboard allowed participants to opt out of time restrictions.[5] The limit control group remained unaware of limits and continued using a version of Phone Dashboard without this functionality.

Bonus and Limit Valuations

Incentive-compatible multiple price list (MLP) mechanisms assessed values of the Screen Time Bonus and limit functionality. Participants made choices in an MPL survey to determine their perceived values, predicting screen time reductions and choosing between bonuses and payments. To ensure familiarity with the model, Survey 1 included a practice MPL that asked participants to decide between different survey completion payments at different times. In Survey 2, participants learned their past FITSBY screen time and predicted future usage after the introduction of the Screen Time Bonus. They then made hypothetical choices between the bonus and a payment matching their expected earnings (ranging from 0 to $150). To motivate honest responses, participants were informed of random selection for payment determination. A small percentage of participants[6] were placed in an MPL group where they received what they chose on a randomly selected row. In Survey 3, the limit group engaged in an MPL to assess the value of Phone Dashboard limit functions.

Surveys 2-4 gathered predictions of future FITSBY use, incentivizing accuracy with a "prediction reward," meaning that participants got an extra payout for accurately predicting their future FITSBY use. Participants were informed of their past FITSBY screen time before each prediction, bolstering their awareness of temptation.

High-Level Results

The analysis encompassed 1,933 participants, evenly distributed across bonus and control groups. Baseline data revealed an average daily screen time of 333 minutes, with 46% allocated to FITSBY apps. Qualitative insights from Figure 3 provide key insights into digital addiction from the baseline survey. Within the restriction index, 23% expressed interest in smartphone time limits, while 34% had no interest. Regarding ideal use change, 42% deemed their smartphone usage appropriate, with 0.5% indicating insufficient use. For those deeming usage excessive, the average ideal reduction was 34%.

---

[4] Snooze feature granted additional time after a specified delay. Participants were randomly assigned to different snooze delays (0, 2, 5, or 20 minutes) or a condition where snooze was disabled. Variations in snooze delay were not further explored in this paper.
[5] 4% of the participants opted out.
[6] 0.2% of participants were selected here.

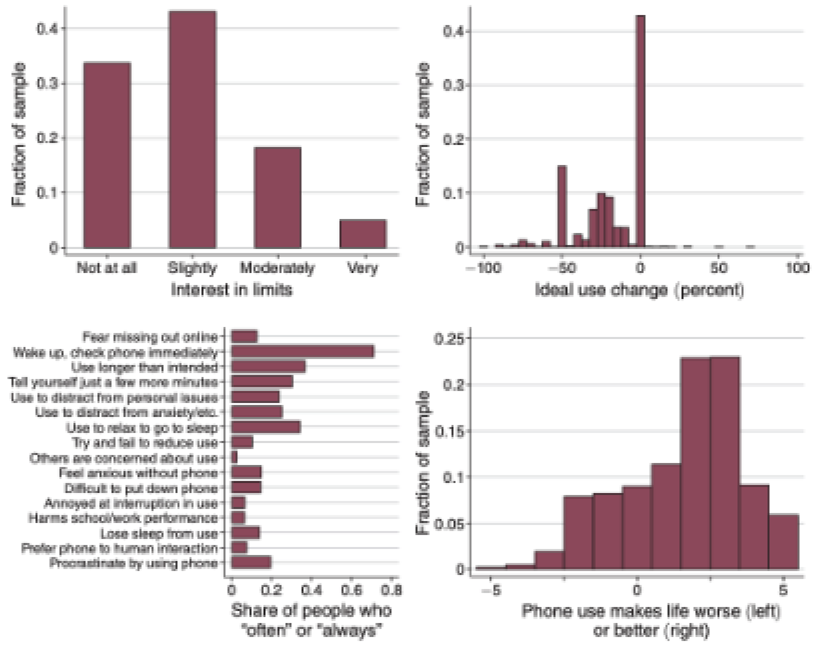

Figure 3. Baseline Qualitative Evidence of Self-Control Problems

In the addiction index, 33% frequently experienced moderate addiction elements, while 11% often or always experienced severe addictions. Notably, 19% perceived smartphone use as having detrimental effects despite a majority acknowledging smartphone usage as improving their lives. These results indicate a divergence in experiences related to digital addiction. Importantly, this study was conducted during the COVID-19 pandemic. Baseline survey questions gauged its potential impact. Findings revealed that 78%–a significant majority–reported having more free time due to the pandemic and 88% acknowledged increased phone usage.

Bonus Treatment and Habit Formation

Figure 4 interprets the impact of the bonus treatment on FITSBY use. A $50 incentive per average hour resulted in a substantial 39% reduction during the incentive period. Impressively, the bonus group sustained a 56-minute daily reduction post-incentive, suggesting the formation of enduring habits.

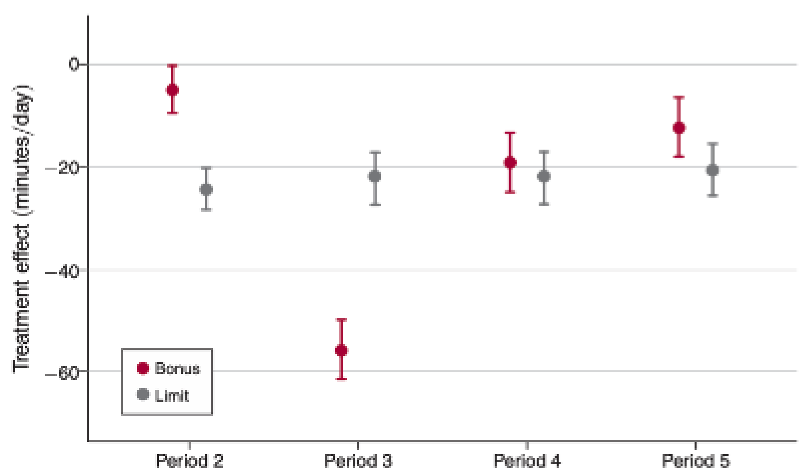

Figure 4. Treatment Effects on FITSBY Use

Limit Treatment and Temptation

Extensive engagement from the limit group underscored their commitment to restricted screen time. The "limit tightness" variable[7] indicated that 89% established positive limit tightness, averaging 53 minutes per day. Figure 4 demonstrates a 16% FITSBY use reduction in the limit group, with a sustained effect of 19 minutes per day in the final week of period 5,[8] suggesting enduring effects.

Substitution

Figure 5 explores the impact of bonus and limit the treatments on FITSBY apps. Notable reductions in FITSBY app usage are depicted, revealing a critical shift. Assessing effects on other apps, the rightmost coefficient explores participants' substitution of FITSBY time with alternative apps. During Period 3, the bonus showed no statistically significant impact on other apps, dismissing substantial substitution compared to FITSBY's observed 56-minute daily reduction. Conversely, the limit appeared to induce a 12-minute daily substitution, meaning that nearly half of FITSBY screen time shifted to apps where users were less likely to impose limits. An important caveat arises from the inability to directly monitor FITSBY use beyond smartphones. While Survey 4 sought insight into this substitution, it demonstrated a marginal increase of 4.2 minutes per day for the limit and an 8.1-minute reduction for the bonus on other devices. This implies that time on other devices acts as a mild complement in the case of the bonus.

Crucially, distinctions emerge in substitution effects between the bonus and limit. Despite expectations of similar proportions, a smaller share of the bonus effect of FITSBY use is substituted to other smartphone apps compared to the limit. Additionally, self-reported effects on FITSBY use on other devices exhibit opposite signs for the bonus versus the limit.

[7] Representing the hypothetical reduction in screen time if a user's limits were applied to their baseline
[8] This is noteworthy considering that the experimental procedures concluded weeks before period 5.

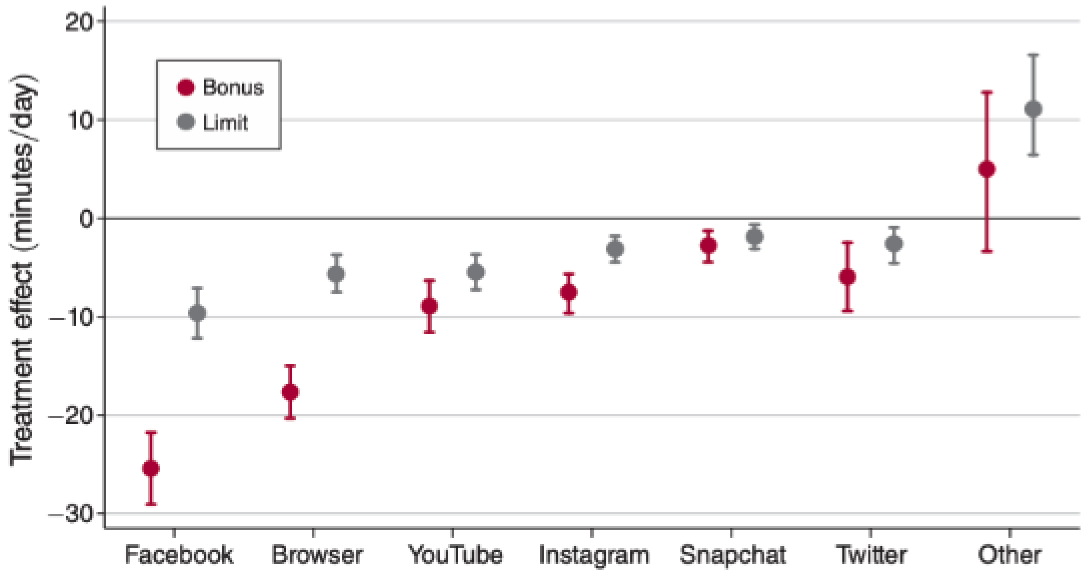

Figure 5. Effects on Smartphone Use by App

Predicted vs. Actual Use

Figure 6 illustrates the predicted and actual FITSBY use in the control condition, where neither the bonus nor the limit functionality was present. Predicted use is winsorized[9] within 60 minutes of actual use, revealing consistent underestimations across all future periods.

In Figure 7, treatment effects of the bonus on actual use are compared with participants' predictions. Before the MPL for the Screen Time Bonus, participants accurately estimated a 52-minute daily reduction. Subsequent predictions in Survey 3 align with the actual reduction in period 3, indicating an understanding that the bonus would persist beyond the incentive period. The time path of actual versus predicted effects suggests a potential overestimation of habit formation.

Bonus and Limit Valuations

In Survey 3 MPL, the average participant in the limit group expressed a willingness to forgo a $4.20 fixed payment for three weeks of access to the limit functionality. Notably, 58% of participants were open to sacrificing some money, with 20% willing to give up more than $10, signaling a demand for digital self-control tools and suggesting perceived self-control problems. Additionally, Survey 2 MPL

[9] Transformed by limiting extreme values in the statistical data to reduce the effect of possibly spurious outliers.

highlighted that individuals perceiving self-control issues favored the Screen Time Bonus over higher fixed payments to align future use with current preferences.

Effects on Survey Outcomes

Figure 8 shows the impact of bonus and limit treatments on survey outcomes. The interventions significantly reduced self-reported addiction measures, with the bonus exhibiting larger effects than the limit across most variables. The bonus decreased ideas use change by approximately 9 percentage points, while the limit decreased it by about 5 percentage points. In terms of Subjective Well-Being (SWB), the bonus and limit treatments increased SWB by 0.09 and 0.04 standard deviations respectively, with the effects appearing to be driven by improved concentration and reduced distraction.

Section II: Strengths & Weaknesses of the Paper

Strength: Experimental Design

A strong aspect of this research paper is the cleanness of its treatment and control setup. Participants were randomly split into treatment and control groups, so the only differences between the treatment groups and the control groups were the treatment conditions.[10] This is thought of as a standard feature of experiments, but it can get a little muddled in some laboratory experiments. Here, whether or not the researchers' theories are correct, they can be fairly confident that any difference in FITSBY use between the groups is due to the treatments. This required a careful and detailed selection of language, specifications for Phone Dashboard, and survey design/scheduling from the researchers.

The researchers used this clean experimental design to convincingly obtain evidence for their theories, using a strong mathematical formulation of their experimental results. In Figure 5, the researchers highlight how the bonus treatment seemed to limit FITSBY use, as compared to the limit treatment and the baseline. They also compare the decrease in FITSBY use to the use of other applications, showing that the treatment was well-targeted and had the intended effect, particularly on the applications in question.

Also, although the researchers used surveys to ascertain some idea of the baseline use of FITSBY applications, they did not start the treatments until after Phone Dashboard had been installed for several weeks. This allowed them to get accurate data about how much participants were using FITSBY applications within the context of the experiment (i.e. after the participants knew more about what the experiment was about and would consist of). This is another example of removing potential confounding factors.

No experiment is perfect, but the researchers limited bias in many places and were highly diligent about ways to gather accurate information about their participants and results.

Weakness: Selection Bias

Within this paper, participants were pooled through advertisements within the Facebook platform. This included advertisements placed in the funnels of 2 primary social media platforms, Instagram and Facebook. Among these two platforms, over 18,000 individuals were shown the advertisement to participate in the study. Of those 18,000, there was a selected pool (those who completed the needed

[10] One could argue that there are complicating factors as constituted by the instructions that informed the bonus treatment group that they would receive bonuses and that the control group would not. However, this is a minor and essentially unavoidable aspect of the experiment, and only serves to emphasize the difficulty of a perfectly clean experiment.

surveys) of roughly 2000 American adults with Android smartphones. This design yields the risk of copious levels of selection bias. Primarily, those using Facebook and Instagram only account for a third of the platforms the authors of this study intended to implement. The authors chose to advertise on these platforms solely for ease of advertisement. This could cause a sort of selection bias, primarily in the form of a representation bias. This means that by only drawing their sample from users of 2 of the 6 social media platforms they aimed to study, the responses they get may not be accurately representative of the population they infer conclusions. Beyond this, they only advertised to American people with Android devices. This foreseeably caused a selection bias as there are differences within the characteristics and demographics of Android users and all other users. To keep things clear, we focus on the differences between Android and iOS users. In a study completed in 2014, researchers found that "Android users seem to be more secure and privacy aware, mostly because they notice Android permissions.'[11] Beyond this, the same study[12] found that women are more likely to have an iPhone, people who are more brand-aware are more likely to have an iPhone, and those with a technical background are more likely to have an Android phone. Further, a paper released on ExplodingTopics, a blog website focused on Google search data, by Josh Howarth discussed how iPhones and Androids differ in their market share statistics, with iPhone garnering roughly 58% of the US market share and Android having around 70% of the global market share.[13] Fundamentally, what these sources reveal is that the differences between iPhone and Android phone users, even at a demographic level, are fundamentally different. By only surveying those with Android phones they most probably experienced some level of selection bias due to their misrepresentation of the population they meant to study, further supported by the misalignment of characteristics above.

Weakness: Covid Implications (Temporal Effects)

Another shortcoming we noticed throughout evaluating this paper is that the study occurred during the beginning of the COVID-19 pandemic. A study by the National Library of Medicine revealed that during the COVID-19 pandemic, "people in rural and urban areas had a significant increase in their device use."[14] This variable may cause an interaction effect within what they are aiming to study. We feel as though, while impossible to predict at the time, it would have been more efficient to complete the study

---

[11] Reinfelder, Benenson, and Gassman. "Differences between Android and iPhone users in their security and privacy awareness."
[12] Reinfelder, Benenson, and Gassman. "Differences between Android and iPhone users in their security and privacy awareness."
[13] Howarth. "iPhone vs Android User Stats (2023 Data)."
[14] Jonathan, Seaton, Rush, Li, and Hasan. "Mobile Device Usage before and during the COVID-19 Pandemic among Rural and Urban Adults."

outside of a time in history when people were forced inside and, per the NLM, were on their phones at a statistically higher level than normal. While this is guaranteed to have skewed their results of screen time higher than they would typically, we hypothesize that it would also impact their desire to reduce such use as, quite literally, people often had nothing better to do.

Weakness: Definition Variance (social media addiction)

Finally, one of the largest shortcomings, in our opinion, of the paper being analyzed is its narrow-sighted definition of digital addiction. They define addiction as the combination of habit formation and self-control problems. They define habit formation as the increase of tomorrow's demand due to the consumption that happens today. This states if a person uses a habit-forming product on a particular day, then that person's demand for that product will be greater the next day (ceteris paribus). They define self-control problems as people consuming more or less today than they would have chosen for themselves in advance. If a product involves a self-control problem, a person using it at a particular time uses it more or less than they would have chosen to do so at a time before or after the particular moment (again, ceteris paribus). These definitions are constructed through the desire to have an operationalizable and analyzable data set. The problem with these definitions is that they are based solely on the consumption of some product (namely social media usage in this study) when that is not a holistic definition of addiction. Addiction is constructed by psychological, physiological, and environmental factors as well (as briefly touched upon in the COVID-19 segment). The researchers' model does not account for these factors, nor underlying factors that may affect the results of the study. In short, digital addiction is a complex and convoluted topic, and we believe their definition to be an oversimplification of the topic.

Section III: Types of FITSBY Use

Overview & Motivation

Missing from this research is any consideration of the type of actions the participants were taking on the FITSBY technologies. That is, researchers only measured the amount of time that participants were using FITSBY technologies and not what they were doing. A possible extension to this research would be to investigate the differences in digital addiction with some consideration about what participants are doing during their FITSBY use.

A critical difference and a good place to start would be investigating differences in digital addiction between users who primarily take more ‘substantive’ actions like messaging or posting, compared with more ‘passive’ users who primarily just watch or scroll through content (“liking” a post would probably fall more in the ‘passive’ category, as it is a very simple action that requires minimal thought). This would be important in terms of understanding which parts of social media contribute more or less to digital addiction. That knowledge could then be used to develop company-wide or governmental policies that regulate social media in a more fine-tuned way. This more detailed approach might be able to preserve more about what is good about social media while addressing its more pernicious aspects.

Technological Method of Implementation

To briefly address one simple (but possibly infeasible) way this further research might be achieved, the researchers’ creation of Phone Dashboard provides a foundational example. They were able to hire a development team to create software that fit their specifications and the ethical concerns of research; it is easy to imagine that the software could have been created differently to more closely track what participants were doing. The participants would be informed about exactly what was tracked and how the data would be aggregated to anonymize individual responses. These processes are even easier if the companies behind these technologies could be involved. They could directly add to their code to achieve exactly the necessary specifics of what the study participants were doing.[15] However, there are other ways that researchers could try to measure what mental state participants are in when using social media.

[15] There are obviously many difficulties and ethical concerns about these companies directly monitoring what particular users are doing. Their ability to change code is brought up mostly to highlight how detailed the information could be.

A More Traditional Method

One way that researchers could start to investigate people's use of digital technology correlated with their actions and mental state would be by examining night-time usage—specifically late-night usage of social media that occurs right before going to bed or during bedtime. Not all social media usage is the same, as each social media application has different and unique features, and the motivations and constraints for why an individual might use social media vary over a week, and even throughout the day.

Some usage of social media does not seem to fall under the definition of a digital addiction. Using a social media application to respond to some action of a close friend (e.g. answering a message that person sent) has a different tenor than scrolling through posts made by strangers. This more benign-seeming usage would contribute to total social media usage, but might not impact the findings related to the persistent effects of the study and the effects of setting limits. However, it is possible that there was a habit formation of limiting 'benign' usage and/or other (here, other implies digital addiction) social media usage, such as transferring social media usage to a different device, perhaps exaggerating the effect found within the study.[16] Night-time usage data would presumably contain much lower 'benign' usage, hence exploring more problematic and damaging forms of social media usage, meaning that outcomes might be more stark from this more focused approach.

This research strategy also has some benefits from an experimental design standpoint. Social media usage has a plethora of constraints that arise from various settings of usage over each day. The paper addressed this by analyzing usage over three weeks, but to understand the underlying mechanisms of habit formation and self-control problems, information with a consistent setting is important. Late-night social media usage has a largely constant setting: the participant's bedroom. A constant setting is linked to habit formation.[17] Also, self-control is not the only factor of night-time social media usage, given that biological factors[18] have been shown to impact late-night procrastination; for this and other reasons, studying late-night social media usage would provide more specific insights into the mechanisms and reality of digital addiction.

[16] It would also be possible to do the opposite: find a time of day when people are more likely to be taking 'benign' actions.
[17] Ouellette and Wood. "Habit and intention in everyday life: The multiple processes by which past behavior predicts future behavior."
[18] Kühnel, Syrek, and Dreher. "Why Don't You Go to Bed on Time? A Daily Diary Study on the Relationships between Chronotype, Self-Control Resources and the Phenomenon of Bedtime Procrastination."

Analyzing night-time usage could be achieved by replicating the study and by collecting data on the part of the day that social media usage occurs, having participants label their night-time social media usage in the third week before being randomly assigned treatments and then analyzing the change in night-time usage over the following weeks. Importantly, this could prompt the participants to think that night-time usage should be reduced in the study. It is also worth noting again that the paper's study was over the pandemic, so it is plausible the overall results would be different for this proposed study.

Works Cited

1. Hunt Allcott, Matthew Gentzkow, and Lena Song. “Digital Addiction.” American Economic Review 2022. Accessed at www.aeaweb.org/articles?id=10.1257/aer.20210867
2. Lena Reinfelder, Zinaida Benenson, and Freya Gassman. “Differences between Android and iPhone users in their security and privacy awareness.” Springer, 2014. Accessed at www.researchgate.net/publication/281031486_Differences_between_Android_and_iPhone_Users_in_Their_Security_and_Privacy_Awareness
3. Josh Howarth. “iPhone vs Android User Stats (2023 Data).” Exploding Topics, 2023. Accessed at explodingtopics.com/blog/iphone-android-users
4. Livia Jonnatan, Cherisse L. Seaton, Kathy L. Rush, Eric P. H. Li, and Khalad Hasan. “Mobile Device Usage before and during the COVID-19 Pandemic among Rural and Urban Adults.” Creative Commons Attribution, 2022. Accessed at www.ncbi.nlm.nih.gov/pmc/articles/PMC9315523/#:~:text=Our%20findings%20indicated%20that%20people,health%20compared%20to%20rural%20participants
5. Judith A.Ouellette and Wendy Wood. “Habit and intention in everyday life: The multiple processes by which past behavior predicts future behavior.” Psychological Bulletin, 1998. Accessed at psycnet.apa.org/record/1998-04232-003
6. Jana Kühnel, Christine J. Syrek, and Anne Dreher. “Why Don’t You Go to Bed on Time? A Daily Diary Study on the Relationships between Chronotype, Self-Control Resources and the Phenomenon of Bedtime Procrastination.” Frontiers, 2018. www.frontiersin.org/articles/10.3389/fpsyg.2018.00077/full