

# Social Media Detox

W241 Experiments and Causality - Final Project

Stanley Ye, Shangyun Lv and Kuangwei Huang

School of Information, University of California, Berkeley  
{stanley.ye, shangyun.lv, kuangwei}@berkeley.edu

2019 April 21<sup>st</sup>

## 1. Abstract

This field experiment aims to evaluate whether message prompts promoting the benefits of limiting social media usage will affect subjects' daily social media mobile app usage and their emotional wellbeing, and whether such effects, if any, are sustained post-treatment. To this purpose, a two-week experiment was conducted in which participants were randomly assigned to either control or a one-week treatment period. Observations consisting of participants' daily app usage of ~~either~~ Facebook, Instagram or WeChat, and their emotional states were made pre-treatment, after the first week of treatment, and in the week after treatment. The treatment was administered via online surveys, and the specific app usage times were collected through screenshots from participants' screen time apps from their mobile phones' operating systems. The results of the study show that there are no statistically significant effect of message prompts being able to discourage usage times on Facebook, Instagram and WeChat, both during the one-week treatment period and in the post-treatment week. There is also no statistically significant treatment effect on participants' emotional states from the measurements made during the two weeks of the experiment.

## 2. Background

The boom of social media is an incredible phenomenon in Internet development. Within twenty years, we have lived through several generations of social media platforms, starting from Six Degrees (launched in 1997 and shuttered in 2000) to Friendster (launched in 2002 and closed in 2015), from Facebook (launched in 2004) to Instagram (launched in 2010) and Snapchat (launched in 2011). With its rapid growth and updates, social media has gone from relative obscurity in 2004 to playing an integral part of our daily life today. We use it to connect with family and friends, to share stories and experiences, to catch up on news and current events, and for entertainment. In 2018, about 68% of Americans and 2.6 billion people globally use social media<sup>1</sup>.

While the general public has embraced social media and it has become a fixture in many people's lives, recent research and studies are showing the negative effects of social media on society and individuals. Untrue information, cyberbullying<sup>2</sup>, depression and anxiety, unhealthy sleeping patterns, and general addiction<sup>3</sup> are a few among the widely accepted negative effects of social media. On the other hand, people have reported becoming happier and healthier after giving up social media<sup>4,5</sup>. Recently, researchers from the University of Pennsylvania conducted a study on 143 undergraduates, randomly selecting a subset of the students to limit their usage on Facebook and Instagram to less than ten minutes a day over a three-week period, and have

<sup>1</sup> Keith, *The History of Social Media: Social Networking Evolution!*, History Cooperative, 2019.

<sup>2</sup> Petersen, L. *The Negative Effect of Social Media on Society and Individuals*, Chron, 2019.

<sup>3</sup> Zhukova, A. *7 Negative Effects of Social Media on People and Users*, MakeUseOf, 2018.

found that the treatment group showed significant reductions in loneliness and depression compared to the control group<sup>4</sup>.

### 3. Research Question

Considering how many of these social media platforms are engineered to be addictive to increase engagement, perhaps acknowledging the negative effects and being mindful about using social media could alter behaviors and increase emotional wellbeing. Following this thought, our experiment seeks to answer the questions:

"Does being advised to reduce social media usage help reduce one's social media usage and improve his/her emotional status in the following days? And, are the effects sustained post-treatment?"

### 4. Experimental Design

#### 4.1. Recruitment

For both recruiting participants and conducting our experiment, we chose Qualtrics to enroll participants, collect information, administer treatment, and measure outcomes. The participants are recruited from the authors' personal and academic networks. The requirement for participating in the experiment was being able to successfully complete a pre-experiment survey described in the next section.

#### 4.2. Pre-Experiment Survey

A pre-experiment survey was conducted to screen potential subjects and ensure that they are able to meet the requirements for the experiment. The survey requires respondents to declare being over the age of 18 before allowing further continuation to collect contact and demographic information from the subject. The pre-experiment survey queries the respondents for self-reported usage times of five popular social media apps over the past seven days to allow us to determine a smaller subset of the most popular social media apps to target for the experiment.

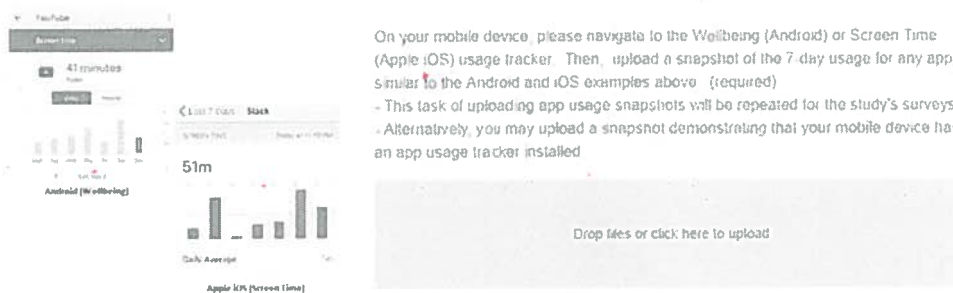


Figure 1: App Usage Data Collection

As shown in Figure 1, the pre-experiment survey requested subjects to upload mobile screen snapshots of their app usage trackers, i.e., Android Wellbeing and iOS Screen Time, to verify that they could provide social media app usage data during the experiment. We did not include the app usage data collected by the pre-experiment survey for any of our analyses, since we had not determined which social media apps to target.

<sup>4</sup> Hunt, et. al (2018). *No More FOMO: Limiting Social Media Decreases Loneliness and Depression*. Journal of Social and Clinical Psychology: Vol. 37, No. 10, pp. 751-768

"to measure this  
upgrades"

on sentence

Besides the treatment effect on participants' social media daily usage, another potential outcome we care about is whether the treatment changes the participants' emotional states, and what we need is a simple and quantifiable scale for emotion state that we could include as a potential outcome. However, the emotional state measurement is not yet a standardized practice. Some of the methods involve lengthy questionnaire such as Ryff's Psychological Well Being Scale<sup>5</sup> which potential increase difficulty for the participants to complete the survey; some other methods may be too specific to certain emotional states such as loneliness scale<sup>5</sup> or self-esteem scale<sup>5</sup>. After evaluating a number of well-known test methods, we decided to utilize the circumplex model for the emotional state measurement which map all emotional state or mood into two axis - "Activation" and "Pleasant" shown in Figure 2. In the model, every word of mood falls into one of the four quadrants of the system and has its corresponding weight on the axis. In the experiment, we carefully select 8 words from the model (two from each quadrant) and ask the participants to rate each of the 8 words in a scale from 0 to 10 based on their past 7-day feelings. The potential outcomes for each participant are the total activation score and pleasant score which are the weighted sum from the word rating. The benefit of using the circumplex model of measurement are multifold. First, we are able to avoid a lengthy questionnaire and measure participants emotional state with only one question in the survey; second, we only create two potential outcomes for the experiment -- "Activation" and "Pleasant", which simplify the following analysis; finally, using the weighted sum of the 8 selected words score avoids the domination of any specific emotional state, like loneliness or depression, and is able to capture more comprehensive emotional states.

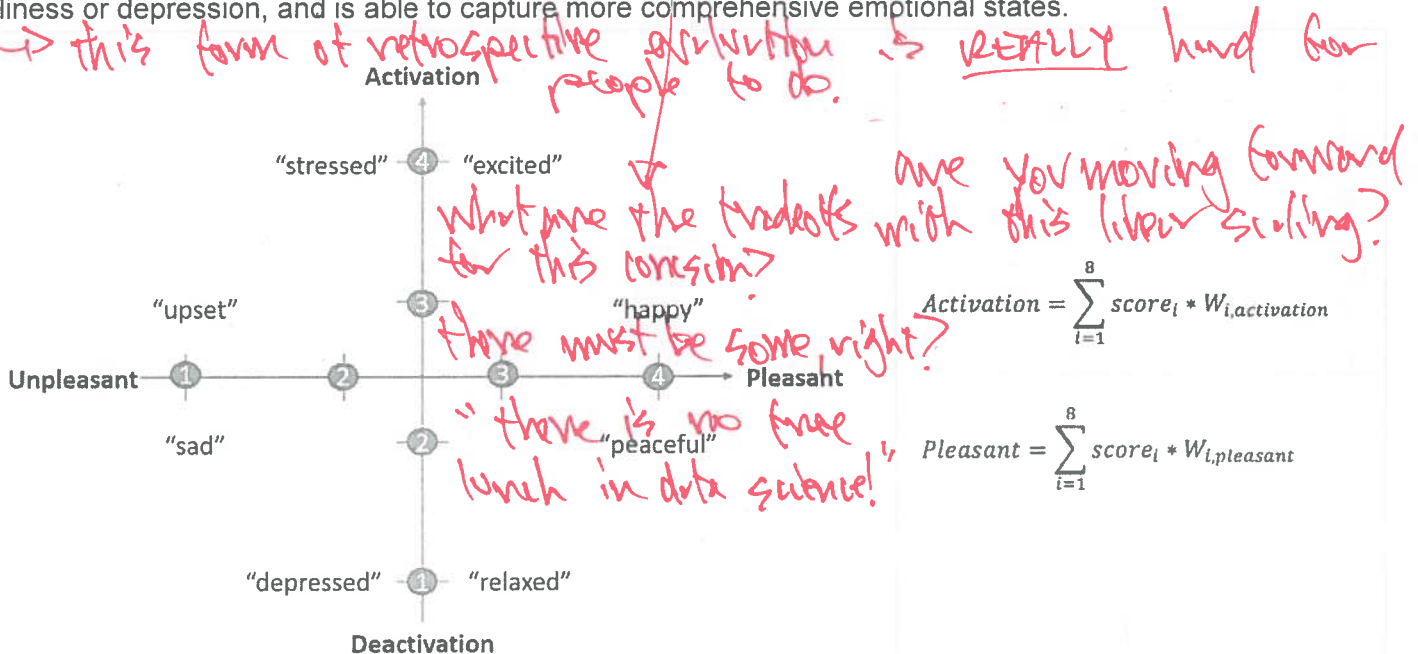


Figure 2: Circumplex mood model and "Activation" and "Pleasant" scoring function

The survey reached a total of 176 survey respondents, out of which, 47 respondents completed the full survey and included a valid screen capture of an app usage from their mobile phone. From the results of the pre-experiment survey, the three most popular social media apps are Facebook, Instagram, and WeChat.

<sup>5</sup> Ryff, C. D. (1989). *Beyond Ponce de Leon and life satisfaction: New directions in quest of successful ageing*. International Journal of Behavioral Development, 12, 35-55.



### 4.3. Pre-Test and Post-Test Design

Our experiment utilizes a pre-test and post-test measurement design, and the experiment laid out in ROXO grammar is shown in Figure 3. The experiment will be administered with a series of email links to different sets of survey questions, and acts both as treatment administration as well as the mechanism to collect outcome measurements. The experiment consists of a Day-0, Day-7 and Day-14 surveys sent to both treatment and control groups, which take a measurement on a subject's past 7-day social media app usage as well as the psychological assessment. Subjects assigned to treatment would receive an alternate version of the Day-0 survey with an added messaging prompt encouraging them to limit their time on social media. Subsequently, subjects in treatment would continue to receive a variety of such prompts from Day-2 through to Day-6. All treatment survey messages require subjects to acknowledge receiving the message at the end of each survey.

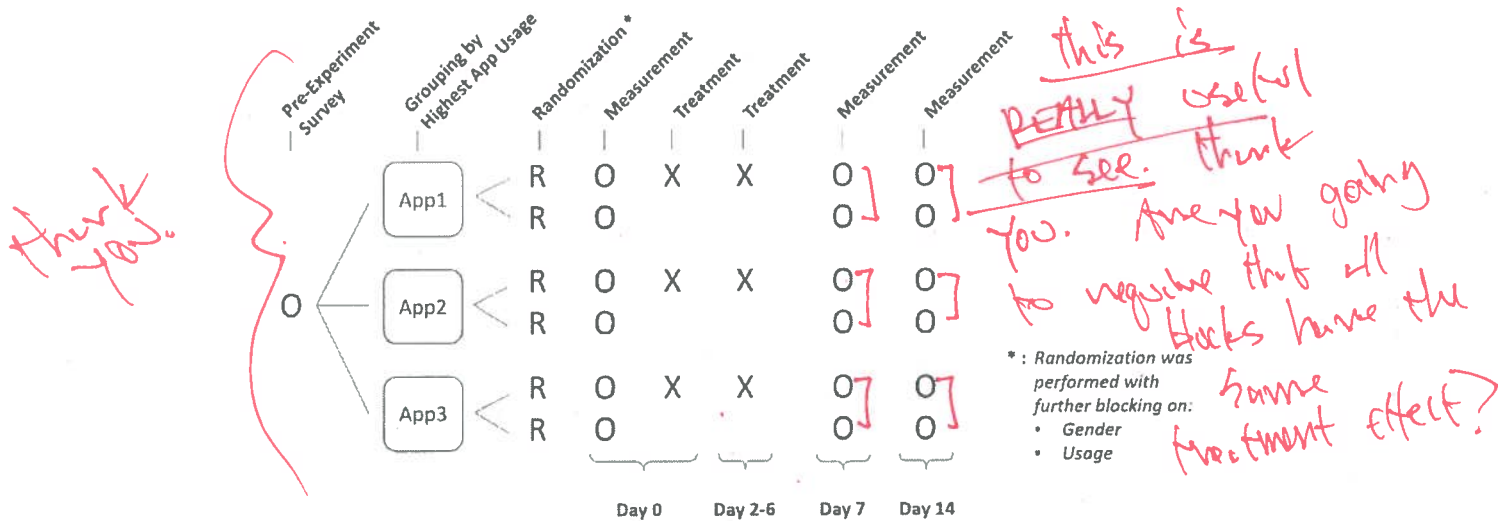


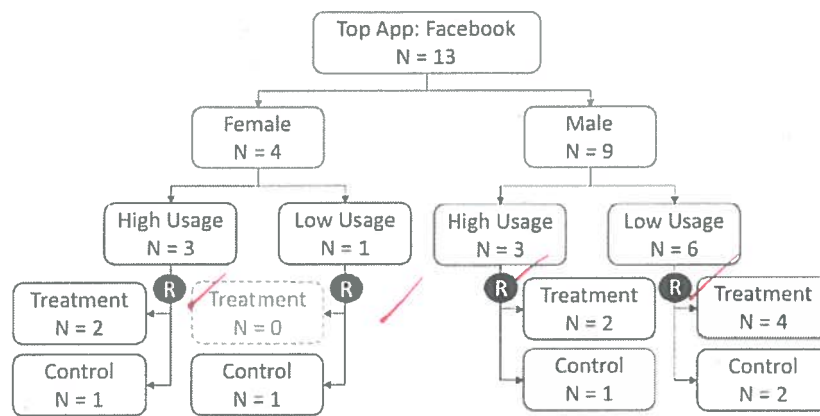
Figure 3: Experiment design using ROXO grammar.

We grouped our sample by their most highly used social media app out of the 3 most popular apps in our sample: Facebook, Instagram, and WeChat, so as to possibly observe larger treatment effects. This also allowed for the specific instructions for taking screen captures on the assigned app in every measurement survey, so that subjects grouped under Facebook, for example, would be specifically asked to submit in screen captures of their Facebook app usage, and mitigate the risk of submitting in a screen capture for another social media app.

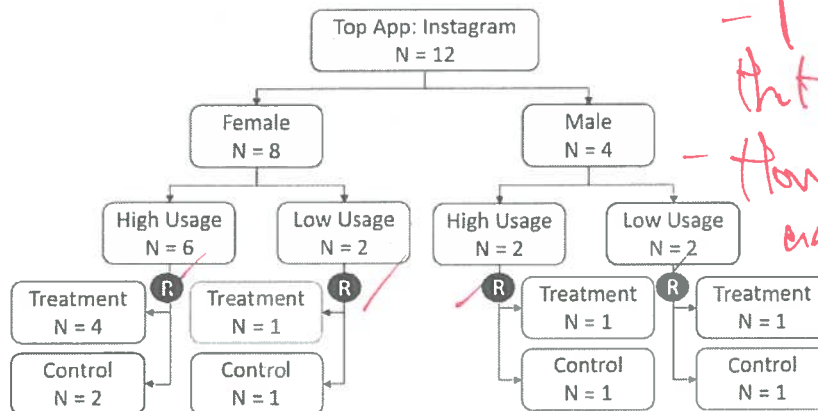
Handwritten note: randomization construction

### 4.4. Blocking and Randomization

Based on the demographic information collected in our pre-experiment survey, we chose to further block each app group by gender and whether the subject has a high usage of the social media app. We defined high usage of the social media app as more than two hours of usage per week. The subjects were then randomized using the R statistical software for treatment and control assignment with a probability of 60-70% of being assigned to treatment. The detailed scheme for randomization is provided in Figure 4.



- R** - Randomization
- 70% treatment for Facebook
  - 70% treatment for Instagram
  - 60% treatment for WeChat



- I also really like that you built these.

- How are you going to aggregate these treatment effects up to whole population Estimate.

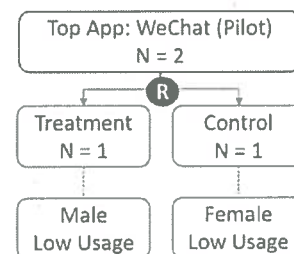
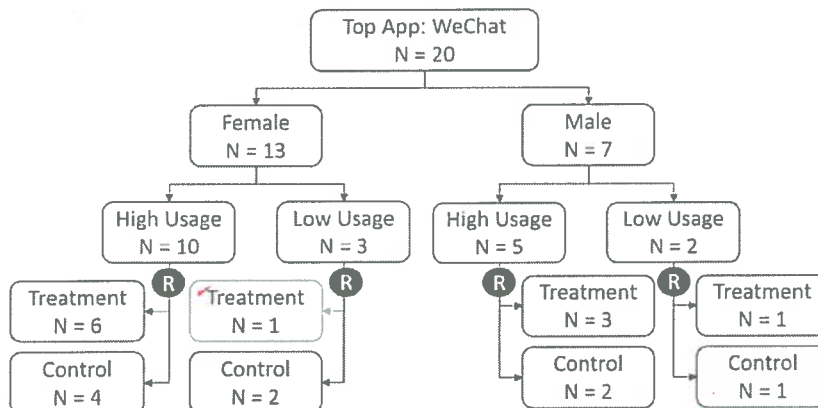


Figure 4: Blocking and randomization for all the 3 selected social media apps.

As we were only able to attain 47 subjects for the experiment, the randomization produced some blocks with only one subject, but we believed this to be the best way to achieve covariate balance in the randomization given the small sample size.

#### 4.5. Pilot Test

A couple of days before launching the experiment, we selected two subjects for a pilot test of the Day-0 survey. This was a final check to validate that actual Day-0 surveys could be completed and that screen captures could be successfully captured and uploaded to Qualtrics. We verified that we could extract app usage times from

the screen captures before releasing the Day-0 surveys to all subjects. No further changes were made to the survey after this short two-day pilot test.

#### 4.6. Pre-Treatment Compliance

Compliance in our experiment was determined by participants successfully completing the baseline measurement on Day-0 prior to treatment. Hence, the Day-0 survey was planned to be differentiated between control and treatment groups in that although both control and treatment groups had the same screen capture instructions and psychological assessment questions, the treatment group survey would have an additional message prompt at end of the survey to act as the first treatment message.

#### 4.7. Treatment and Control

The treatment consisted of daily email messages during the first week of the experiment. Figure 5a shows the first day's treatment, which cites the aforementioned University of Pennsylvania study findings, and subsequently requests that subjects consider curbing their social media app usage to under five minutes a day. Meanwhile, the first day's control message (Figure 5b) contains only a reminder that another app usage and emotional state survey would be sent in one week. Lastly, the end of every treatment and control message has an acknowledgement checkbox (Figure 6) to confirm receipt of the message.

Did you know a recent study in the University of Pennsylvania found that students who limited their social media usage to less than 10 minutes a day reported significant reductions in loneliness and depression over three weeks compared to a control group?

With the touted benefits of limiting social media use, please consider lowering your **Facebook** usage to less than 5 minutes a day. You could make use of the reminders in the iOS Screen Time or Android Wellbeing apps to help you achieve this.

We will be keeping in touch over the next few days and will check back with your usage times in 7 days!

Figure 5a: First Treatment Message

*Ok. This is a really polite manipulation. If I read this - and I am sorry to reason from a single observation (me) I would consider your manipulation... and then*

We will be keeping in touch over the next few days and will check back with your usage times in 7 days!

Figure 5b: First Control Message

Acknowledgement

☐ I have received this message!

Figure 6: Message Acknowledgement

*not change a damn thing about my life! You should be a little more bold in structuring the experience for your subjects.*

After the first day of the experiment, the control group would not receive any additional email messages until Day-7. Meanwhile, the treatment group would receive daily email messages, like the one in Figure 7, which provide an anecdote highlighting the positives of limiting social media usage, a multiple-choice question to get subjects to reflect on their own social media usage, and a gentle reminder to continue restricting their social media app usage for the first week of the experiment.

In 2018, CNBC health-tech reporter Christina Farr took a break from social media after realizing that she spent far more time on Instagram than she realized

In a few weeks, she started to enjoy life's more mundane moments and experience more gratitude for what she had

Following that, she had been off Facebook and Instagram for more than three months

Which of the following do you most agree with when a person stops using social media entirely for 1-2 weeks:

- ☐ Social media is now an integral part of life, people will eventually return to these platforms
- ☐ People would continue to stay off these platforms for the long term
- ☐ Undecided... and exercise my right to sit on the fence

I really really really like the "push-poll" nature of this question.

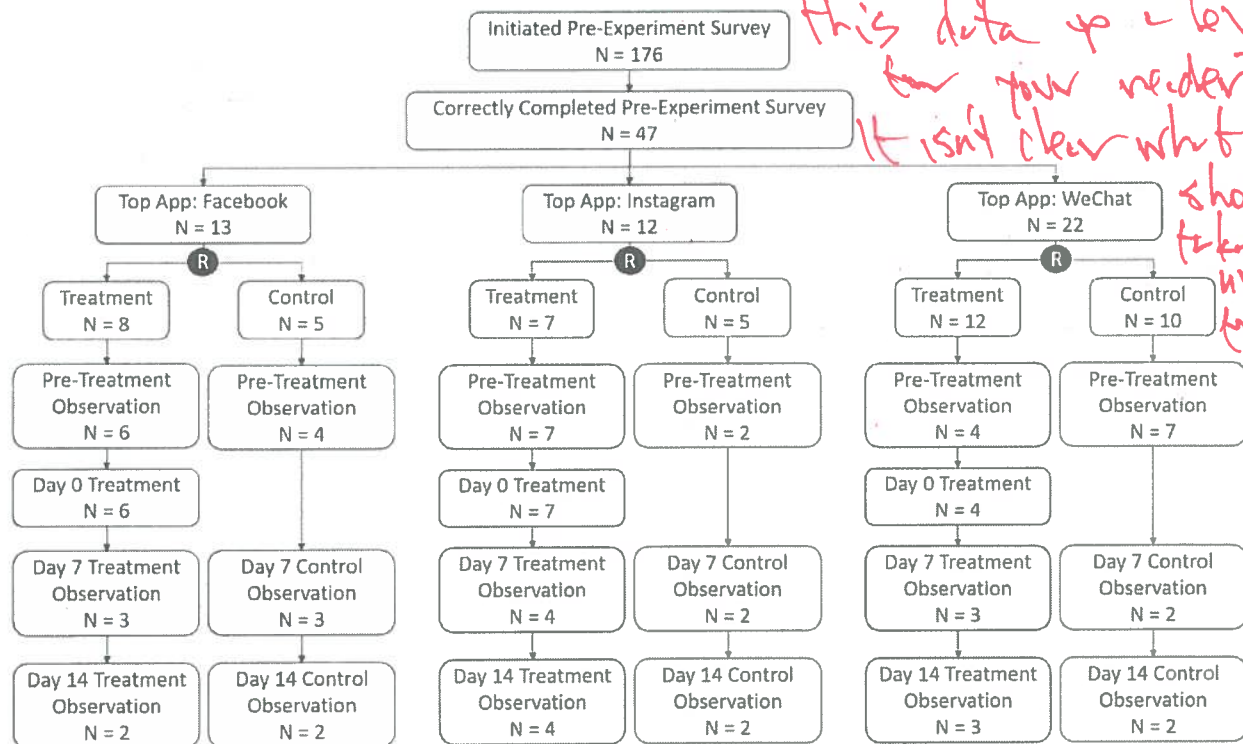
Continue to be mindful about limiting your social media usage this week, and experience the change!

Figure 7: Daily Treatment Messages

## 5. Observations

### 5.1. Attrition

We faced challenges on attrition for our experiment as we started out with a small sample size of 47 initially and subsequently observed subjects attrit out of the experiment either by not responding to the emails with the survey links, or by uploading incorrect screen captures. The flowchart for the attrition is shown in Figure 8, and we ended up only 17 subjects who completed all actions correctly at Day-7 and a further attrition down to 15 subjects who stayed through to complete Day-14.



R - Randomization with additional blocking on Gender and Usage

Figure 8: Flow diagram showing treatment assignment and attrition over the course of the experiment.



## 5.2. Screen Time Captures

word choice

We found that while most subjects are able to upload a correct screen time image for their social media app into Qualtrics, we did have cases whereby screen time was submitted for the usage in the current day instead of the past 7 days, and cases whereby screen time of all apps was submitted instead of the requested Facebook, Instagram or WeChat app. Examples of these cases are provided in Figure 9. We did benefit from having separate surveys for each app group with specific instructions for subjects to take screen captures of their particular app, and did not have any cases where subjects submitted a different social media app from their grouping.

Do people substitute one app for another?

Extracting the total and average social media app usage was performed manually, and getting the individual daily usage times required the pixel height of the bar charts to be measured and then interpolated against the total usage time indicated in the screen capture.

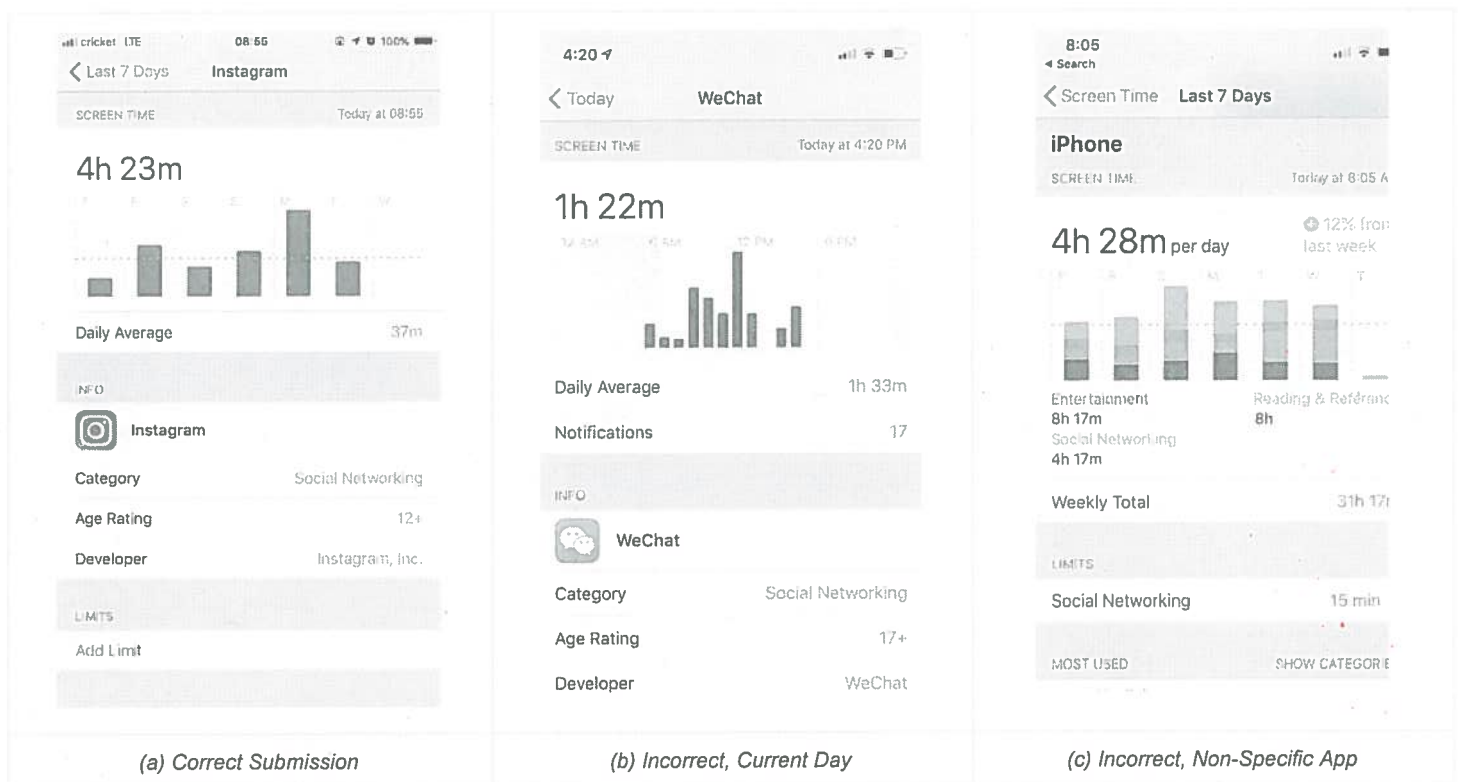


Figure 9: Showing the (a) correct screen time capture for the past 7-day usage of the social media app, (b) incorrect capture of the current day's app usage instead of past 7 days, and (c) incorrect capture without the specific app selected.

## 5.3. Usage Data Anomalies

Interestingly, many screen captures showed a zero or near-zero usage time for the final day of each week. This phenomenon can be attributed to screen time trackers resetting their daily counters at midnight. Hence, if a screen capture occurs early in the day before the social media app has been used, the app usage for the final day of that week would be misleadingly marked zero. Ideally, subjects should take screen captures at 11:59pm, right before the app usage counter resets; however, we decided not to overburden our subjects with this additional requirement. Due to subjects taking screenshots at various times of the day, rather than close

→ good!



to midnight, we chose to omit the usage data of the final day of the week-long screen captures for all three experiment surveys: Day-0, Day-7, and Day-14.

## 6. Results

Upon completion of the two week experiment period and manually transcribing all app usage data, we conducted statistical analysis to determine the treatment effects on our outcome variables of app usage time and emotional states. We initially blocked on gender and high initial app usage times, and assembled demographic data from the pre-experiment survey to run analyses with multiple covariates, such as gender, age, country, and mobile OS. However, we decided against including too many covariates in our analyses due to our small, final sample size.

### 6.1. Social Media App Usage

We plotted the mean daily social media app usage for the subjects over the 14 days of the experiment, and there was no discernible difference in overall usage trends between control and the treatment groups, as shown in Figure 10. Especially with the large standard deviation across subjects' usage times, the treatment administered from Day-0 to Day-6 does not seem to influence overall social media app usage.

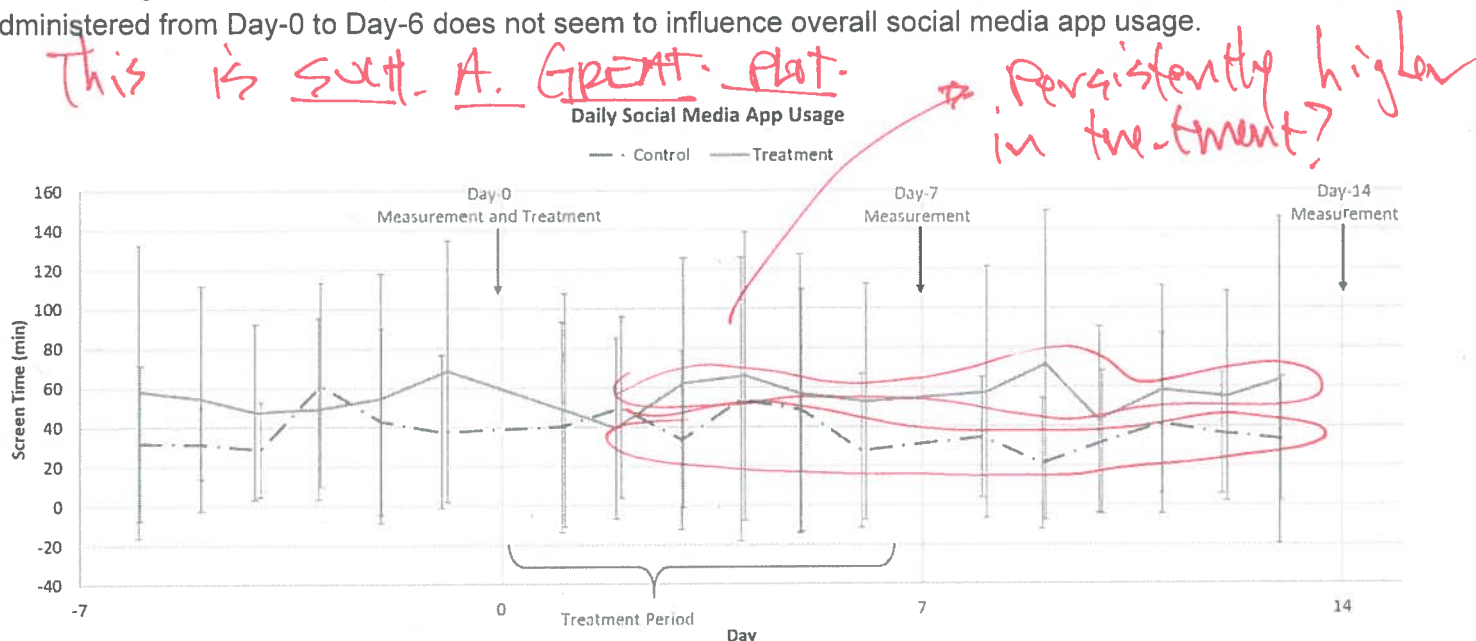


Figure 10: Daily social media usage for control and treatment groups over the course of the experiment.

We found that different social media apps have different mean usage times by subjects. Subjects assigned to report Facebook usage average 8.03 (12.87) minutes daily, whereas that for Instagram is higher at an average of 27.71 (30.83) minutes daily, and WeChat is the highest by far at an average of 106.88 (45.42) minutes daily. Referring to Figure 11 for Facebook usage, Figure 12 for Instagram usage, and Figure 13 for WeChat usage, when looking at the individual app subset data, there are no obvious changes in usage behavior for Facebook and Instagram, although WeChat subjects under treatment seem to have increased app usage times in the post-treatment period from Day-7 to Day-14.

*Why? That's a huge difference!*



Figure 11: Facebook usage for control and treatment over the course of the experiment.

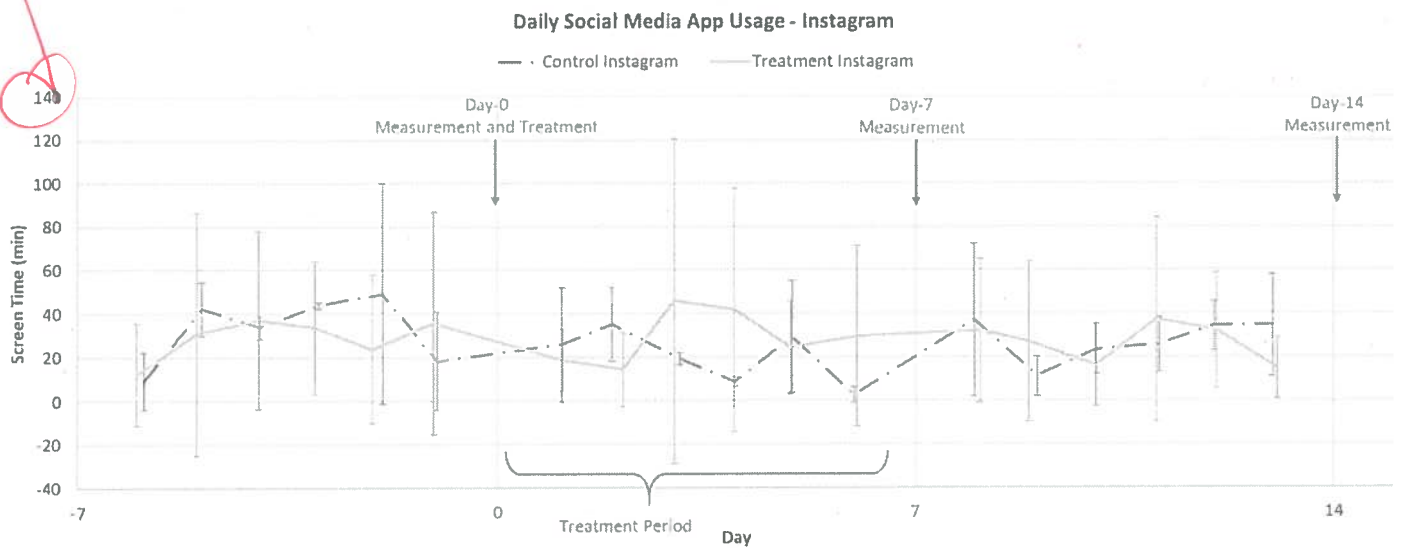


Figure 12: Instagram usage for control and treatment over the course of the experiment.

- Because these plots are all meant to be read together - make sure they all have the same y-axis.
- How many people are represented in each of these groups? FB, IG, WC?

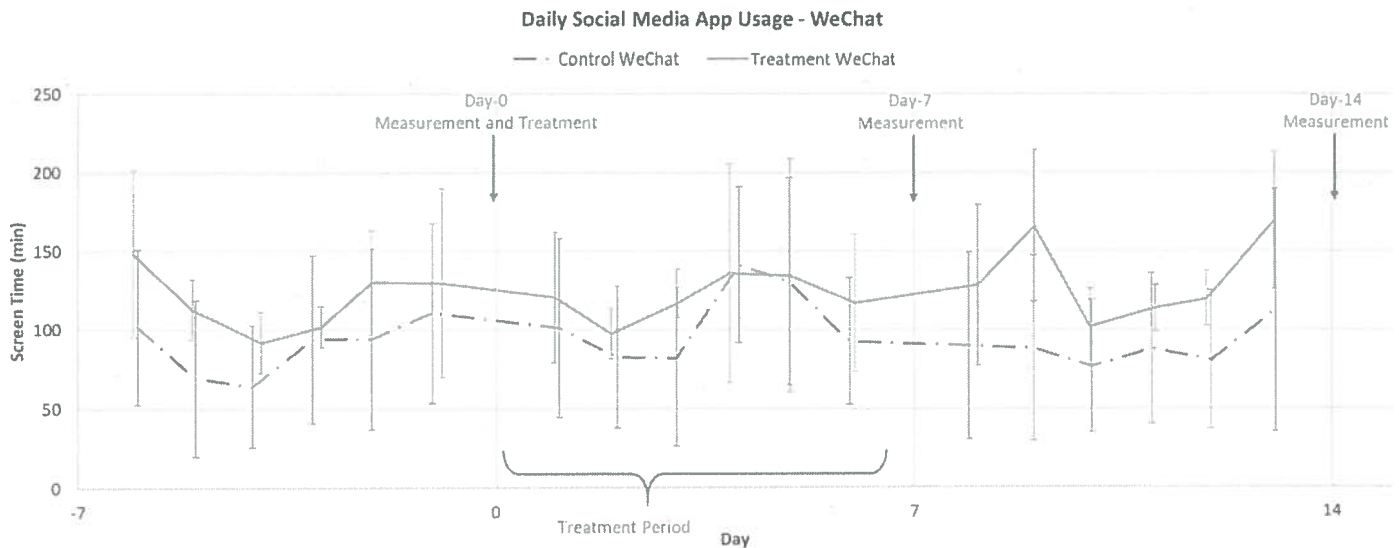


Figure 13: WeChat usage for control and treatment over the course of the experiment.

In order to test our hypothesis of message prompts being able to reduce social media app usage, we first performed a generic linear regression on the data with two outcome variables, the daily social media usage during Week-1 which is the treatment period between Day-1 and Day-6 inclusive and that during Week-2, which is the post-treatment period between Day-8 and Day-13 inclusive. We used clustered standard errors for the analysis, clustering on the subject identification number assigned. The results of the regression are shown in Table 1, and there is no statistically significant effect of the treatment on daily app usage whether it is during the treatment period or post-treatment period.

Table 1: Regression table of daily social media app usage outcomes over 2 weeks.

	Daily Social Media Usage (minutes)	
	during Week-1 (1)	during Week-2 (2)
Usage during Week-0 (minutes)	0.890*** (0.099)	0.793*** (0.102)
Treatment during Week-1	-2.625 (11.959)	11.901 (8.192)
Constant	7.476 (9.967)	2.137 (6.765)
Clustered SE on Subject	Yes	Yes
No. of Clusters	15	15
Observations	90	90
R <sup>2</sup>	0.625	0.599
Adjusted R <sup>2</sup>	0.616	0.590
Residual Std. Error (df = 87)	35.117	34.251
F Statistic (df = 2; 87)	72.445***	65.113***

Note:

\*p<0.1; \*\*p<0.05; \*\*\*p<0.01

REALLY  
GOOD

this regression  
is really nicely  
specified and  
intuitive.  
align on  
declined using  
align=T  
align=T



Next, we ran separate regressions on each of the three social media apps: Facebook, Instagram, and WeChat to see if there are app specific treatment effects. Referring the Table 2, we found no statistically significant treatment effect for Facebook and Instagram both in Week-1 and Week-2 app usage. WeChat was more interesting with no statistically significant treatment effect for Week-1 usage but a highly statistically significant effect on Week-2. The large coefficient of an increase in usage of 56.99 (9.55) minutes daily by the treatment group was also practically significant against the constant term of 52.37 (19.92) minutes.

We also obtained very low R-squared values in the separate regressions, and attribute this to the small number of subjects upon which we are clustering.

Table 2: Summary of separate regressions run on the 3 social media apps, and comparing the daily usage outcomes over 2 weeks

	Daily Social Media Usage (minutes)					
	Facebook (1)	during Week-1 Instagram (2)	WeChat (3)	Facebook (4)	during Week-2 Instagram (5)	WeChat (6)
Usage during Week-0 (minutes)	0.033 (0.095)	0.651*** (0.129)	0.419 (0.289)	-0.102 (0.072)	0.335** (0.146)	0.199 (0.230)
Treatment during Week-1	-0.609 (2.354)	11.515 (7.538)	0.950 (24.350)	4.305 (4.115)	0.239 (12.541)	56.991*** (9.550)
Constant	5.581*** (1.692)	-1.228 (6.068)	69.698* (36.454)	5.577*** (1.720)	16.716 (12.337)	52.370*** (19.918)
Clustered SE on Subject	Yes	Yes	Yes	Yes	Yes	Yes
No. of Clusters	3	6	5	4	6	5
Observations	24	36	30	24	36	30
R <sup>2</sup>	0.007	0.404	0.169	0.071	0.187	0.477
Adjusted R <sup>2</sup>	-0.088	0.368	0.108	-0.017	0.137	0.438
Residual Std. Error	8.034 (df = 21)	27.205 (df = 33)	42.255 (df = 27)	10.302 (df = 21)	23.600 (df = 33)	36.562 (df = 27)
F Statistic	0.073 (df = 2; 21)	11.206*** (df = 2; 33)	2.753* (df = 2; 27)	0.805 (df = 2; 21)	3.788** (df = 2; 33)	12.310*** (df = 2; 27)

Note: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01

→ Is this lack of finding because of being underpowered?

In attempting to rationalize a causal story behind this finding on WeChat, we hypothesize that during treatment, the message prompts to discourage social media usage had an actual cognitive effect on treatment subjects, positively influencing them to be desire to reduce app usage. However, the utility and social necessitated pressures of continuing the same levels of WeChat usage made the subjects under treatment unable to achieve any real reduction in usage times during the treatment week, even though they may be mentally trying to reduce usage. After treatment stops, the relieve of the cognitive stress of trying to limit WeChat usage may cause subjects to have a short period of increased usage on the platform. It would be interesting to see if these findings are repeatable, as well as to conduct further studies to see if this effect persists further on in time or does app usage regress back to the original levels prior to treatment.

## 6.2. Emotional States for "Activation" and "Pleasant"

The second outcome measure we are interested in is the emotional states of the subjects along two axes "Activation" and "Pleasant" across the 2 weeks of the experiment. In order to maximize the number of observations, we included all subjects who completed the emotional state surveys, regardless of whether they correctly submitted usage time screen captures.

We ran a regression across the data checking if there are treatment effects on the emotional states, and the results are presented in Table 3 and Table 4. We did not find any statistically significant treatment effects for

either emotional state, although the coefficients for treatment weakly suggest that treatment increases "Pleasant" emotional states. We would have expected "Activation" scores to increase if the message prompts are successful in influencing subjects to try to change their app usage behaviors. Since we have previously found that app usage was not affected by treatment, we did not expect any statistically significant effects on "Pleasant" scores.

Table 3: Regression table of emotional state "Activation" outcomes over 2 weeks.

	Emotional State "Activation" Score	
	on Week-1	on Week-2
	(1)	(2)
"Activation" Score Week-0	0.975*** (0.222)	0.629*** (0.224)
Treatment during Week-1	-2.284 (8.674)	-3.512 (8.740)
Constant	1.281 (19.921)	34.516* (20.074)
Observations	18	18
R <sup>2</sup>	0.635	0.437
Adjusted R <sup>2</sup>	0.587	0.362
Residual Std. Error (df = 15)	16.157	16.281
F Statistic (df = 2; 15)	13.061***	5.826**

Note:

\*p<0.1; \*\*p<0.05; \*\*\*p<0.01

Again, these regressions are informative lots of results.

Table 4: Regression table of emotional state "Pleasant" outcomes over 2 weeks.

	Emotional State "Pleasant" Score	
	on Week-1	on Week-2
	(1)	(2)
"Pleasant" Score Week-0	0.895*** (0.173)	0.733*** (0.194)
Treatment during Week-1	10.695 (9.635)	10.035 (10.773)
Constant	3.208 (18.990)	28.991 (21.234)
Observations	18	18
R <sup>2</sup>	0.651	0.497
Adjusted R <sup>2</sup>	0.605	0.430
Residual Std. Error (df = 15)	18.479	20.663
F Statistic (df = 2; 15)	14.007***	7.419***

Note:

\*p<0.1; \*\*p<0.05; \*\*\*p<0.01

We then ran separate regressions on data subsets of the 3 social media apps: Facebook, Instagram and WeChat. We did not find any statistically significant app specific treatment effects on both "Activation" and "Pleasant" scores as shown in Table 5 and Table 6. This result does not lend further support to the causal story rationalized during the analysis of the daily app usage on how WeChat app usage increased in the post-treatment week, however, we recognize the possibility of limitations on the chosen circumplex model to capture mood subtleties caused by the messaging prompts.

Table 5: Summary of separate regressions run on three social media apps, and comparing the emotional state "Activation" outcomes over two weeks

	Emotional State "Activation" Score					
	Facebook	on Week-1 Instagram	WeChat	Facebook	on Week-2 Instagram	WeChat
	(1)	(2)	(3)	(4)	(5)	(6)
"Activation" Score Week-0	0.812*** (0.236)	1.076*** (0.368)	1.235*** (0.305)	0.523 (0.417)	0.306 (0.260)	0.861** (0.171)
Treatment during Week-1	6.689 (11.949)	25.407 (18.568)	14.830 (15.995)	21.770 (21.595)	-4.669 (13.942)	-7.493 (8.228)
Constant	18.570 (21.951)	-31.233 (46.130)	-33.303 (32.438)	32.530 (37.164)	82.842 (32.289)	27.553 (18.491)
Observations	9	7	7	8	5	6
R <sup>2</sup>	0.665	0.699	0.805	0.323	0.741	0.908
Adjusted R <sup>2</sup>	0.553	0.548	0.707	0.053	0.481	0.847
Residual Std. Error	17.412 (df = 6)	16.290 (df = 4)	19.998 (df = 4)	29.457 (df = 5)	9.303 (df = 2)	9.261 (df = 3)
F Statistic	5.950** (df = 2; 6)	4.639* (df = 2; 4)	8.253** (df = 2; 4)	1.194 (df = 2; 5)	2.855 (df = 2; 2)	14.874** (df = 2; 3)

Note:

\*p<0.1; \*\*p<0.05; \*\*\*p<0.01

Table 6: Summary of separate regressions run on three social media apps, and comparing the emotional state "Pleasant" outcomes over two weeks

	Emotional State "Pleasant" Score					
	Facebook	on Week-1 Instagram	WeChat	Facebook	on Week-2 Instagram	WeChat
	(1)	(2)	(3)	(4)	(5)	(6)
"Pleasant" Score Week-0	0.812*** (0.236)	1.076*** (0.368)	1.235*** (0.305)	0.523 (0.417)	0.306* (0.171)	0.861*** (0.260)
Treatment during Week-1	6.689 (11.949)	25.407 (18.568)	14.830 (15.995)	21.770 (21.595)	-4.669 (8.228)	-7.493 (13.942)
Constant	18.570 (21.951)	-31.233 (46.130)	-33.303 (32.438)	32.530 (37.164)	82.842*** (18.491)	27.553 (32.289)
Observations	9	7	7	8	5	6
R <sup>2</sup>	0.665	0.699	0.805	0.323	0.741	0.908
Adjusted R <sup>2</sup>	0.553	0.548	0.707	0.053	0.481	0.847
Residual Std. Error	17.412 (df = 6)	16.290 (df = 4)	19.998 (df = 4)	29.457 (df = 5)	9.303 (df = 2)	9.261 (df = 3)
F Statistic	5.950** (df = 2; 6)	4.639* (df = 2; 4)	8.253** (df = 2; 4)	1.194 (df = 2; 5)	2.855 (df = 2; 2)	14.874** (df = 2; 3)

Note:

\*p<0.1; \*\*p<0.05; \*\*\*p<0.01

*This is so hard... people are really bad at follow through.*

## 7. Limitations and Future Enhancements

### 7.1. Limited Awareness of App Usage Trackers

Despite attracting nearly 200 people to our pre-experiment, enrollment survey, less than 50 subjects actually qualified for our experiment. One thing that hindered us was the relative newness of both the Android and iOS app usage trackers, released in March 2018 (Android Pie) and September 2018 (iOS 12). As a result, many



interested participants either did not have the app usage trackers installed or were unfamiliar with how to use them. We tried to compensate for this lack of awareness by providing links to both app usage trackers in the enrollment survey, as shown in Figure 14.

App Usage Tracker:

Android Wellbeing (Version: Pie or later): <https://wellbeing.google/>

Apple Screen Time (Version: iOS 12 or later): <https://support.apple.com/en-us/HT208982>

Figure 14: Webpage Links for App Usage Trackers

To counter this limitation, we could have recruited volunteers from a more tech-savvy demographic (e.g., college students), who would presumably be more knowledgeable and proficient with app usage trackers. Targeting such demographics could possibly have led to more qualified subjects.

## 7.2. Manual Data Collection

We considered having subjects fill in their daily app usage amounts but decided against this approach due to the high effort cost for participants. Instead, we asked subjects to upload to Qualtrics three separate screenshots (i.e., Day-0, Day-7, and Day-14) of their weekly app usage data. In hindsight, this approach has two major downsides. First, the manual process of converting dozens of screenshots into csv data points was very labor intensive and took us several hours to complete. Second, the iOS app usage tracker does not display a date; hence, we cannot determine if screenshots were taken on requested days (i.e., Day-0, Day-7, and Day-14) or on other, random days. For instance, Day-0 screenshots are intended to represent pre-treatment app usage; however, if Day-0 screenshots were actually taken on a later date (e.g., Day-3), then the app usage from the screenshot could potentially have been influenced by treatment on Day-0.

The ideal solution would be to replace manual data collection with an automated process that eliminates both human error and effort. For instance, given more time and resources, we would have considered writing an app to surreptitiously collect social media app usage data from subjects' mobile devices. Removing the requirement to take and upload screenshots would have lowered the bar of entry for people to participate in our experiment and potentially led to more subjects enrolling.

## 7.3. Willingness to Detox

In hindsight, one useful pre-treatment data point would have been to ask subjects to rate, on a 5-point scale, their willingness to curb social media app usage. This willingness self-rating could have been a good covariate to explain the treatment effect for treated subjects. For instance, it would be logical to expect the treatment to have insignificant effects on subjects that were unwilling to reduce their social media app usage. On the opposite end of the spectrum, the treatment effect would be expectedly larger for subjects that are more open to curbing their social media app usage.

## 7.4. Email Distribution of Survey Links

We also believe that we could have reduced attrition by distributing the Qualtrics survey links via SMS messages instead of email, since people might not regularly check their inboxes. On one occasion, we even received a report of the survey email being labeled as spam by the recipient. An SMS message with a link to the Qualtrics survey would have been a viable alternative, had we collected mobile phone numbers during the initial pre-experiment survey.

## 8. Conclusions

From our experiment data and analyses, the treatment (i.e., one week of message prompts) did not have a statistically significant effect on either post-treatment social media app usage or emotional state. Due to our challenges in recruiting participants who were both cognizant of the relatively new app usage trackers and willing to submit screenshots of their social media app usage, we were only able to convert about 10% of experiment enrollees into qualified, compliant subjects over the two-week experiment period.

In addition, we neglected to survey how many subjects were willing to abstain from social media apps during the treatment week; hence, we suspect that relatively few, if any, subjects actually attempted to limit their social media app usage to the suggested limit of five minutes per day. This potential lack of cooperation in abstaining from social media would explain why a statistically significant treatment effect is not observed.

Despite the lack of an observable treatment effect, we feel satisfied with our learnings from this experiment. In addition, we are confident that our experiment design can be improved significantly based on our aforementioned observations and enhancement ideas. Lastly, we are optimistic that future research will provide actionable remedies to address social media addiction.

Sterling, Shangyun, Kuangwei -

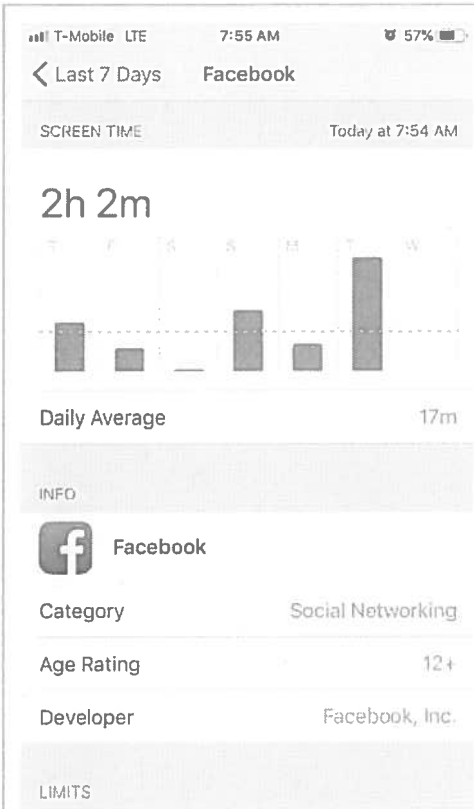
- This is a very well done experiment. From the perspective of the design, estimation, and reporting of the work this is sterling.
- Of course, you cannot help that people didn't sign up as much as you needed. This is a challenge that you all have acknowledged, and that I do not hold against you.
- Your documentation of design and randomization process is super solid. Thank you.
- As well, thanks for the analysis. It is hard to put too much faith in a regression w/ 6 individuals... but I'm sure you have thought this through.

## 9. References

- [1] Keith. The History of Social Media: Social Networking Evolution!. History Cooperative, 2019, <https://historycooperative.org/the-history-of-social-media/>
- [2] Petersen, L. The Negative Effect of Social Media on Society and Individuals. Chron, 2019, <https://smallbusiness.chron.com/negative-effect-social-media-society-individuals-27617.html>
- [3] Zhukova, A. 7 Negative Effects of Social Media on People and Users. MakeUseOf, 2018, <https://www.makeuseof.com/tag/negative-effects-social-media/>
- [4] Mulligan, O. Did Quitting Social Media Make Me Happier? Yes, Here's Why. IRIS LILLIAN, 2018, <https://irislillian.com/quitting-social-media-facebook-instagram/>
- [5] Yasin, K. I Quit Social Media for 65 Weeks. This Is What I Learned. HealthLine, 2018, <https://www.healthline.com/health/my-65-week-digital-detox#1>
- [6] Hunt, Marx, Lipson, and Young (2018). "No More FOMO: Limiting Social Media Decreases Loneliness and Depression.", Journal of Social and Clinical Psychology: Vol. 37, No. 10, pp. 751-768, <https://doi.org/10.1521/jscp.2018.37.10.751>
- [7] Ryff, C. D. (1989). Beyond Ponce de Leon and life satisfaction: New directions in quest of successful ageing. International Journal of Behavioral Development, 12, 35-55, <https://journals.sagepub.com/doi/10.1177/016502548901200102>



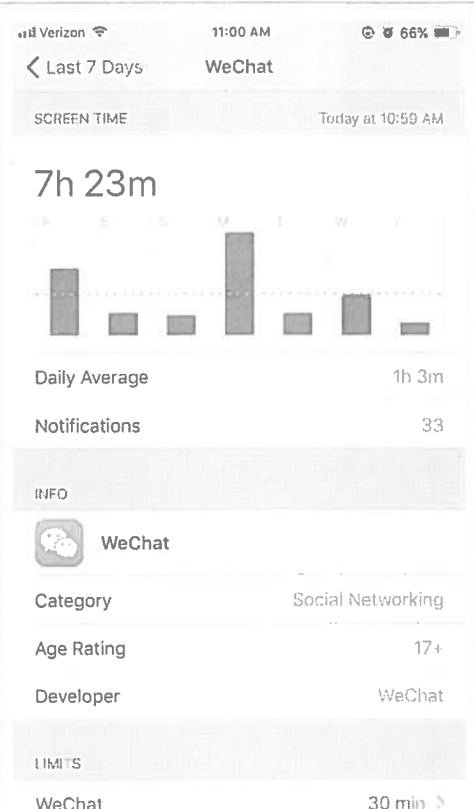
APPENDIX A: Sample Screen Time Snapshots



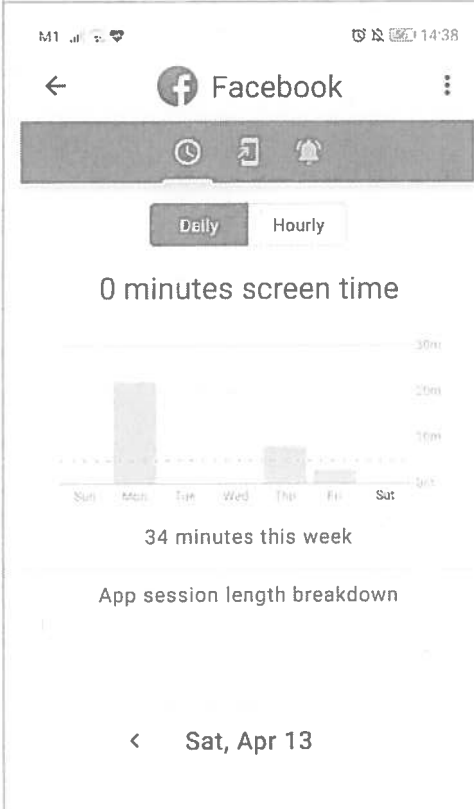
(a) Facebook, iOS



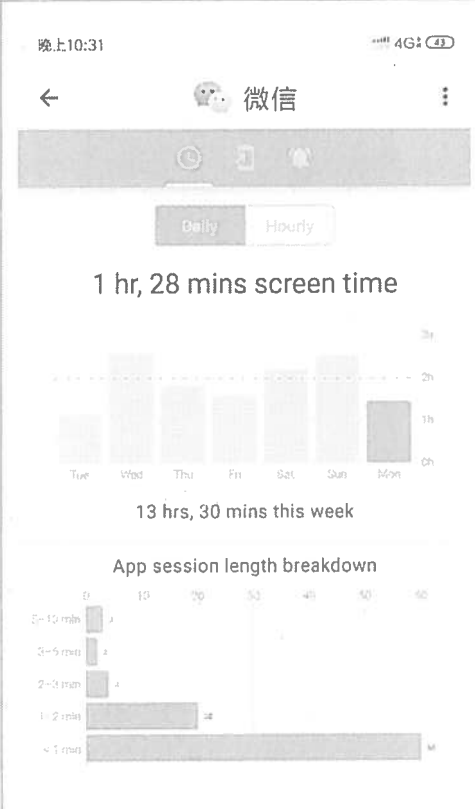
(b) Instagram, iOS



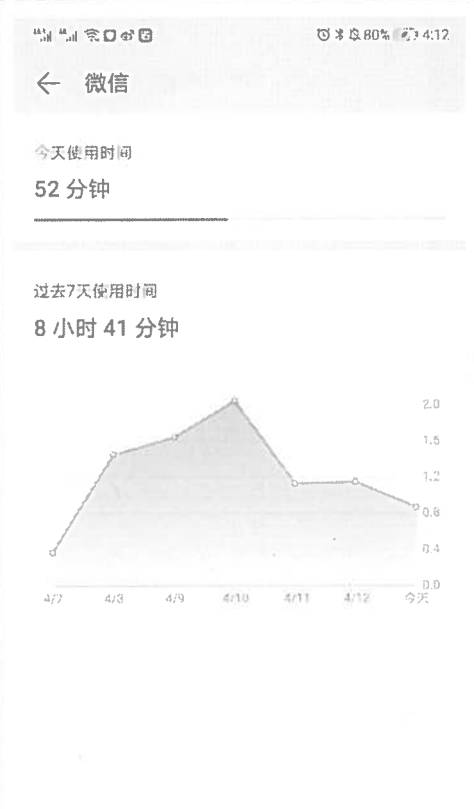
(c) WeChat, iOS



(a) Facebook app's built-in screen time



(b) WeChat, Android OS



(c) WeChat, Android OS

## APPENDIX B: Code

The `wrangling.R` and `aggregation.R` files contain R code for the transforming the Qualtrics response exported data into the data table for analysis. The `project\_rand\_ass.rmd` contains the code used for blocking and randomization, and the `project\_analysis\_0421.rmd` contains the regression and analysis code.

Main R and Rmd code files:

- data extraction and cleaning
  - wrangling.r
  - aggregation.r
- randomization
  - project\_rand\_ass.rmd
- final data for analysis
  - project\_analysis\_0421.rmd

The data folder contains various forms of the experiment including:

- treatment assignment:
  - data/block\_FB\_cntrl\_20190327.csv
  - data/block\_FB\_treat\_20190327.csv
  - data/block\_IG\_cntrl\_20190327.csv
  - data/block\_IG\_treat\_20190327.csv
  - data/block\_WC\_cntrl\_20190327.csv
  - data/block\_WC\_treat\_20190327.csv
- raw Qualtrics response exports
  - All the files in data/survey\_response/day0
  - All the files in data/survey\_response/day7
  - All the files in data/survey\_response/day14
  - data/Experiment Tracking - Main.csv
- transformed survey response data
  - data/survey\_response/response\_day0.csv
  - data/survey\_response/response\_day7.csv
  - data/survey\_response/response\_day14.csv
- daily usage charts in xlsx
  - data/Experiment Tracking - daily\_usage\_chart\_6\_day\_v2.xlsx

## APPENDIX C: Full Examples of Survey Questions

### Day-0: Control for Facebook

Note: Survey was labeled "Day 1" but this was corrected to be "Day-0" by convention in the report.

100% ^

BerkeleyHaas

Screen Time Study - Day 1


Thank you once again for participating in this 2-week study (with minimal participant effort) to monitor screen time for mobile apps!

< YouTube

Screen time

41 minutes  
Today

Daily Hourly



Mon Tue Wed Thu Fri Sat Sun

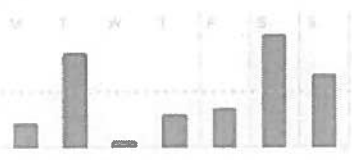
Sun, May 6

Android (Wellbeing)

< Last 7 Days Slack

SCREEN TIME Today 51:18:59 PM

51m



Mon Tue Wed Thu Fri Sat Sun

Daily Average 7m

Apple iOS (Screen Time)

On your mobile device, please navigate to the Wellbeing (Android) or Screen Time (Apple iOS) usage tracker. Then, upload a snapshot of the 7-day usage for **Facebook**, similar to the Android and iOS examples above.

Drop files or click here to upload



On a scale from 0-10 (with 10 being most applicable), how well do the following words describe your emotions today?

0      1      2      3      4      5      6      7      8      9      10

Excited

A horizontal scale bar with a slider at 0.

Happy

A horizontal scale bar with a slider at 0.

Peaceful

A horizontal scale bar with a slider at 0.

Relaxed

A horizontal scale bar with a slider at 0.

Upset

A horizontal scale bar with a slider at 0.

Stressed

A horizontal scale bar with a slider at 0.

Sad

A horizontal scale bar with a slider at 0.

Depressed

A horizontal scale bar with a slider at 0.

We will be keeping in touch over the next few days and will check back with your usage times in 7 days!

Acknowledgement

☐ I have received this message!

## Day-0: Treatment for Instagram

Note: Survey was labeled "Day 1" but this was corrected to be "Day-0" by convention in the report.

100% ^

BerkeleyHaas

Screen Time Study - Day 1

Thank you once again for participating in this 2-week study (with minimal participant effort) to monitor screen time for mobile apps!

< YouTube

Screen time

41 minutes

Today

Daily

Hourly

Mon

Tue

Wed

Thu

Fri

Sat

Sun

Sun, May 6

Android (Wellbeing)

< Last 7 Days Slack

SCREEN TIME

today at 11:59 PM

51m

Mon

Tue

Wed

Thu

Fri

Sat

Sun

Daily Average

7m

Apple iOS (Screen Time)

On your mobile device, please navigate to the Wellbeing (Android) or Screen Time (Apple iOS) usage tracker. Then, upload a snapshot of the 7-day usage for **Instagram**, similar to the Android and iOS examples above.

Drop files or click here to upload

22

On a scale from 0-10 (with 10 being most applicable), how well do the following words describe your emotions today?

0 1 2 3 4 5 6 7 8 9 10

Excited

A horizontal slider bar with a grey dot at the 0 position.

Happy

A horizontal slider bar with a grey dot at the 0 position.

Peaceful

A horizontal slider bar with a grey dot at the 0 position.

Relaxed

A horizontal slider bar with a grey dot at the 0 position.

Upset

A horizontal slider bar with a grey dot at the 0 position.

Stressed

A horizontal slider bar with a grey dot at the 0 position.

Sad

A horizontal slider bar with a grey dot at the 0 position.

Depressed

A horizontal slider bar with a grey dot at the 0 position.

Did you know a recent study in the University of Pennsylvania found that students who limited their social media usage to less than 10 minutes a day reported significant reductions in loneliness and depression over three weeks compared to a control group?

With the touted benefits of limiting social media use, please consider lowering your **Instagram** usage to less than 5 minutes a day. You could make use of the reminders in the iOS Screen Time or Andriod Wellbeing apps to help you achieve this.

We will be keeping in touch over the next few days and will check back with your usage times in 7 days!

Acknowledgement

☐ I have received this message!



## Screen Time Study - Day 2

Thank you once again for participating in this 2-week study (with minimal participant effort) to monitor screen time for mobile apps!

What would you do if you had an additional 30 mins a day?

- ☐ Read a book or listen to an audiobook
- ☐ Have coffee with a friend
- ☐ Start on that side project
- ☐ Meditate or exercise
- ☐ Other :

With time otherwise spent on social media, you could progress on one of these activities, so lower your social media usage today, and get inspired!

Acknowledgement

- ☐ I have received this message!



# BerkeleyHaas

## Screen Time Study - Day 3

Thank you once again for participating in this 2-week study (with minimal participant effort) to monitor screen time for mobile apps!

In 2018, CNBC health-tech reporter Christina Farr took a break from social media after realizing that she spent far more time on Instagram than she realized.

In a few weeks, she started to enjoy life's more mundane moments and experience more gratitude for what she had.

Following that, she had been off Facebook and Instagram for more than three months.

Which of the following do you most agree with when a person stops using social media entirely for 1-2 weeks:

- ☐ Social media is now an integral part of life, people will eventually return to these platforms.
- ☐ People would continue to stay off these platforms for the long term.
- ☐ Undecided... and exercise my right to sit on the fence.

Continue to be mindful about limiting your social media usage this week, and experience the change!

Reminder that you can help your efforts by setting screentime reminders!

### Acknowledgement

- ☐ I have received this message!



## Screen Time Study - Day 4

Thank you once again for participating in this 2-week study (with minimal participant effort) to monitor screen time for mobile apps!

When you see a post on social media, how often do you feel anxious about your life, and feel the need to plan for an important event or make a change?

- ☐ Not at all
- ☐ 1 out of 3 posts make me feel that way
- ☐ Half of all posts make me feel that way
- ☐ More than 2 out of 3 posts make me feel that way
- ☐ Almost all posts I see make me feel that way

People who limit social media usage go on to report feeling lighter and happier!

Continue to monitor and limit your social media usage this week to see what it does for you!

### Acknowledgement

- ☐ I have received this message!