Does the Choice of Profile Picture Affect Online Conversation Perception?

Hejiang Wu, Yutao Luo, Yuhan Zu and Xingyu Yao

1. Introduction

Nowadays, online chat has become the most dominant way of communication. It has become increasingly important to understand how people perceive and interpret online conversations. Especially, users' profile photos are considered as a representative of their online persona and may significant affect how others perceive and interact with them in online environment.

From our literature review, few studies suggest that people form their first impressions of others based on individuals' profile photos, with certain types of pictures eliciting more positive or negative responses (Tong et al., 2008). Moreover, essentially, the text is less than 10% of what people think of you (Lee, 2023). However, relatively few studies have examined how personal profile photos affect people's perceptions of online conversations. Through the experiment, we aim to contribute to this gap and gain a deeper understanding on how profile pictures affect people's perceptions of online conversation.

2. Experiment design

2.1 Goal of experiment

We initiated this experiment to investigate how personal profile photo affect people's understanding and interpretation of online conversations. Our research assumption is that users who set their profile pictures with positive connotations, as compared to neutral ones, make people more likely to perceive his/her message as friendly and positive. To test our assumption, we draw following hypothesis.

Null Hypothesis: there is no difference between treatment and control group's interpretation on the online conversation by users with different profile pictures

Alternative hypothesis: there is a true difference between treatment and control group's interpretation on online conversation by users with different profile pictures

2.2 Survey design

We conducted a randomized experiment by using Qualtrics Survey (Appendix 1 – Preview of Survey) to test our hypothesis. Participants in our survey will be randomized into treatment and control groups, For participants in both groups, they will be randomly assigned to see the same ten short online conversations between two users. For treatment group, participants will see ten text-based conversations, each of which the user on the left-hand side of the screen has a profile picture with positive connotations (e.g., a picture

of a smiling Pikachu, a smiling dog, a smiling woman holding flowers, etc.). In contrast, participants in control group will see the exact same 10 conversations while each of which the user on the left-hand side of the screen has a default profile picture.

To reduce potential bias, we try to eliminate the impact of personal preference on the type of profile photos by using ten different profile photos for each conversation. Similarly, the content of each conversation in our experiment varies by different topic and different tone. The topics of the conversations range from couple breakups to daily conversations with friends, customer service complaints, and so on. By designing our surveys in this way, we try to randomize the inherent friendliness across different types of conversations. Moreover, we add some questions about personal information in our survey, such as age, gender, and ethnicity. These variables may also affect our outcome and could be used as a covariate in our further analysis.

2.3 Friendliness Ratings

Each respondent will rate the conversation friendliness after reading each conversation, and they are asked to answer that, "On a scale of 1 to 10, how friendly do you think the person on the left is during the conversation above?" where 1 is extremely unfriendly, 5 is neutral, and 10 is extremely friendly. It will count as our outcome variable that we will utilize in further analysis.

3. Package loading

When loading our dataset, we found that some participants didn't finish the survey. To clean the data, we use Excel to drop those rows. Then we use Excel to label our data into treatment and control groups. Further cleaning processes are shown in Appendix 2.

```
corrplot 0.92 loaded
```

```
You can cite this package as:
     Patil, I. (2021). Visualizations with statistical details: The 'ggstatsplot' approach.
     Journal of Open Source Software, 6(61), 3167, doi:10.21105/joss.03167
```

```
Attaching package: 'dplyr'
```

```
The following objects are masked from 'package:data.table':

    between, first, last
```

```
The following objects are masked from 'package:stats':

    filter, lag
```

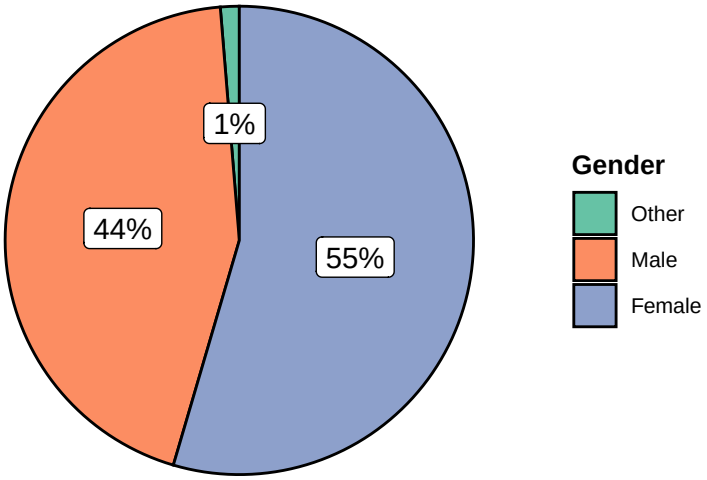
```
The following objects are masked from 'package:base':

    intersect, setdiff, setequal, union
```

4. Exploratory analysis

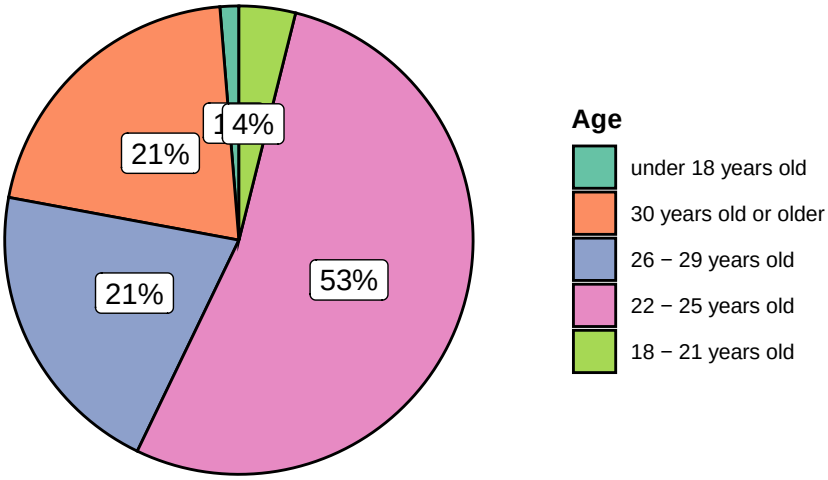
We firstly did an exploratory data analysis to get an overall understanding of our data. There are total 77 observations where 38 of them are in treatment group and 39 are in control group.

Following are distributions of participants' age, gender and ethnicity. From our results, we can see that our data is quite imbalanced. For example, 78% of participants are Asians which may influence their understanding on English conversation.

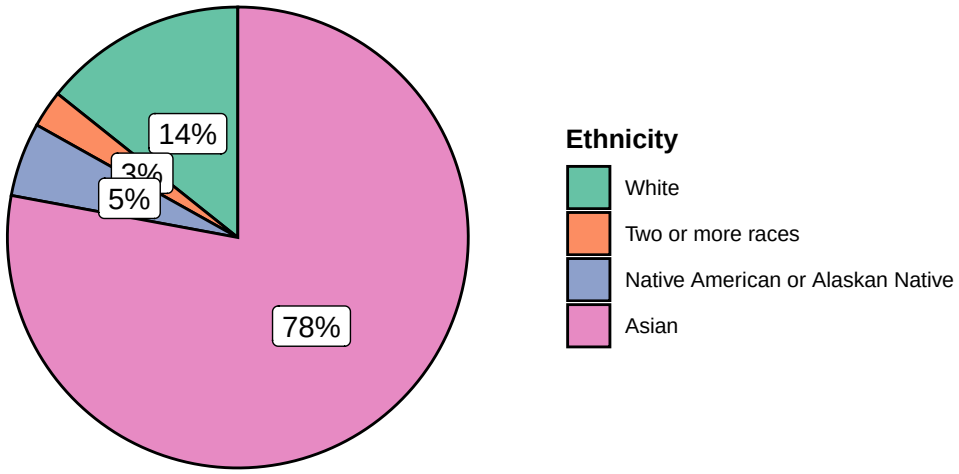
Gender Distribution



Age Distribution



Ethnicity Distribution



5. Experiment Analysis

5.1 Randomization Check

We randomized our experiment at a participant level. 50% of participants were randomized into the treatment while the remaining were randomized into control group. To check if our randomization is done properly, we first perform a proportion test to check if the proportion of units treated is what we expect. Also in step(b), we perform a balance check to see if treatment and control have similar pre-experiment characterisitics.

a). proportions test

From the prop.test result, we have a high p-value with 1 and our sample estimate 0.494 is within 95% confidence interval. It indicates that we fail to reject the null hypothesis at 5% level. Thus our randomization is done properly.

```
#prop_test
num <- nrow(df)
num_treat <- nrow(df[df$treatment==1])
prop.test(num_treat, num, p = .5)
```

```
    1-sample proportions test with continuity correction

data:  num_treat out of num, null probability 0.5
X-squared = 0, df = 1, p-value = 1
alternative hypothesis: true p is not equal to 0.5
95 percent confidence interval:
 0.3786715 0.6089936
sample estimates:
        p
0.4935065
```

b). check for differences between treatment and control

We perform a randomization check on gender and ethnicity. From the result, our co-variate do not exhibit any significant difference between treatment and control. So the randomization is done properly.

```
modelsummary(list("Asian" = reg_eth_check1,
                  "White" = reg_eth_check2,
                  "Two or more races" = reg_eth_check3,
                  "Native American or Alaskan Native" = reg_eth_check4),
             type = 'text', stars = T, output = 'markdown',
             coef_map = c("treatment"="Treatment"), gof_map = c("nobs", "r.squared"))
```

	Asian	White	Two or more races	Native American or Alaskan Native
Treatment	-0.084	0.082	0.001	0.001
	(0.095)	(0.081)	(0.037)	(0.051)
Num.Obs.	77	77	77	77
R2	0.010	0.014	0.000	0.000

Note: ^ + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

```
modelsummary(list("Female" = reg_gender1_check,
                  "Male" = reg_gender2_check,
                  "Non-binary" = reg_gender3_check),
             type = 'text', stars = T, output = 'markdown',
             coef_map = c("treatment"="Treatment"), gof_map = c("nobs", "r.squared"))
```

	Female	Male	Non-binary
Treatment	0.014	0.011	-0.026
	(0.115)	(0.115)	(0.026)
Num.Obs.	77	77	77
R2	0.000	0.000	0.013

Note: ^ + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

5.2 Estimated treatment effect

We perform our analysis using two different measurement levels. The first one is looking at average rating score of all 10 conversations for each participants. The second level is to measure the rating score for each conversation.

5.2.1 Measured by average rating scores

Since we have 10 conversations for each participants, we take the average of 10 rating scores as a measure of our outcome.

5.2.1a Simple regression

We begin with simple regression to estimate average treatment effect. Following are our regression results.

```
df1[,avg:= ((conversation1+conversation2+conversation3+
               conversation4+conversation5+conversation6+
               conversation7+conversation8+conversation9+
               conversation10))/10]

reg_avg <- feols(avg  ~ treatment,
                    data=df1, se='hetero')

modelsummary(list("Simple regression" = reg_avg),
             type = 'text', stars = T, output = 'markdown',
             coef_map = c("treatment"="Treatment"),
             gof_map = c("nobs", "r.squared"))
```

	Simple regression
Treatment	0.097
	(0.209)
Num.Obs.	77
R2	0.003

Note: ^ + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

Based on the results, we can see that the coefficient for treatment effect is 0.097 with a standard error of 0.209. For our SE, it may be considered as large since we have a small unit of observations of 77. It is not statistically significant and thus we fail to reject our hypothesis. In other words, there is no evidence to prove our hypothesis.

5.2.1b Treatment effect by controlling co-variate

Since we have a large standard error in previous regression, we want to increase precision by adding a co-variate. Following are our regression results.

```
modelsummary(list("Simple" = reg_avg,
                  "Simple + ethnitic" = reg_eth,
                  "Simple + gender + ethnicity" = reg_eth_gender),
             type = 'text', stars = T, output = 'markdown',
             coef_map = c("treatment"="Treatment"),
             gof_map = c("nobs", "r.squared"))
```

	Simple	Simple + ethnitic	Simple + gender + ethnicity
Treatment	0.097	0.159	0.180
	(0.209)	(0.207)	(0.211)
Num.Obs.	77	77	77
R2	0.003	0.114	0.139

Note: ^^ + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

By adding ethnicity as a covariate, we can see that our estimated treatment effect increased from 0.097 to 0.159 while SE decrease from 0.209 to 0.207, suggesting an increase in precision. However, when we try to add one more covariate - gender - the result does not meet our expectation as the SE increases. This may be due to a small number of observations compared to a large number of categorical values, which lead to less precise estimates.

5.2.2 Measured by each conversation

In the following measurement, we change our targeted variable to individual ratings of each conversations (10 regressions in total). In this way, we want to test if our treatment effect vary by each conversation.

Treatment effect plus covariate

In the following regressions, we can see that the coefficient of treatment effect vary by each conversation. Eight of them have a positive treatment effect, while two of them have a negative treatment effct. However, none of them are statistically significant. Therefore, we can conclude that there is no evidence to prove our assumption.

```
modelsummary(list("1" = reg1c,"2" = reg2c,"3" = reg3c,
                  "4" = reg4c,"5" = reg5c,"6" = reg6c,
                  "7" = reg7c,"8" = reg8c,"9" = reg9c,"10" = reg10c
                  ), type = 'text', stars = T, output = 'markdown',
             coef_map = c("treatment" = "ATE plus covariate"),
             gof_map = c("nobs", "r.squared"))
```

	1	2	3	4	5	6	7	8	9	10
ATE plus covariate	0.080	0.627	0.196	-0.235	0.216	0.236	0.647	-0.246	0.251	0.033
	(0.419)	(0.501)	(0.592)	(0.379)	(0.401)	(0.363)	(0.398)	(0.456)	(0.323)	(0.477)
Num.Obs.	77	77	77	77	77	77	77	77	77	77
R2	0.059	0.193	0.139	0.193	0.100	0.146	0.086	0.041	0.107	0.042

Note: ^^ + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

6. Limitation

After running the regression, we realized that there are several limitations in our experiment.

a). small sample size

By randomizing at a participant level, we only have 77 observations in our experiment. It leads to a small statistical power so that we do not have enough power to detect a significant effect. One potential improvement is to randomize at a question level so that we could have more sample size.

b). threats to external validity

One potential threat is that each participant may have a different understanding on conversation contents and profile photos, which could affect how they interpret the nature of the conversation and the emotional tone. For example, 78% of our participants are Asian whose first language are not English. While interpreting a English conversation, it will largely affect their perception of friendliness in the conversation.

Another threat is that participants may have different levels of attention on the profile photo stimulus. While asking feedback from survey participants, we found a few participants focus on profile photo while most participants indicates that they rate the conversation based on the context of story and form an initial impression based on the conversation content.

What's more, we may also need to carefully consider the generalizability of our findings to different populations or contexts. For example, our survey are mainly distributed among BU campus while most participants are students from MSBA program. Therefore, our participants may not be representative of overall population of interest. Additionally, it may introduce spillovers effect as students in MSBA program may interact with each other when they fill in the survey.

7. Conclusion

We conduct a casual experiment to test whether the choice of profile pictures affect how people interpret online conversations. We made a hypothesis that it is more likely for users who use profile pictures with positive connotations to be interpreted as more friendly and positive. From our regression analysis, we observed a positive average treatment effect, but it is not statistically significant. We also introduced 2 covariates - ethnicity and gender - to increase our precision. However the results shows that there is no significant evidence to reject our null hypothesis. Therefore our conclusion is that there is no significant difference between treatment and control group's interpretation on the online conversation by users with different profile pictures.

While observing the standard error, one potential reason for our insignificant results is that there is multicollinearity exists. It is also possible that our sample size is too small so that produce much uncertainty. Going forward, we can distribute our survey through more social media and collect more data to increase statistical power to detect differences between two groups. Another potential improvement is to randomized our experiment at question level and add some more variables to control for uncertainty such as participants' language background.

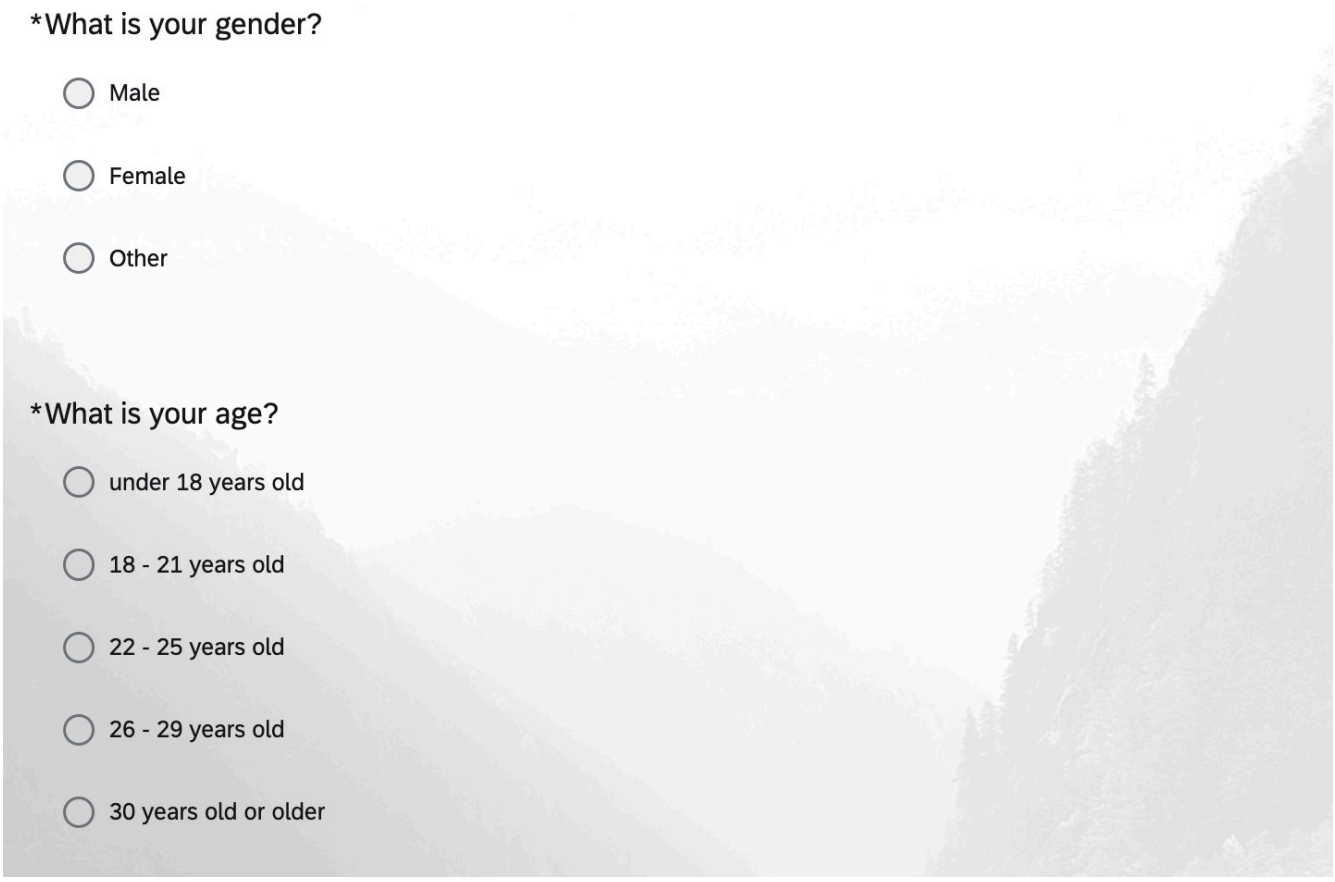
8. Reference

Tong,2008. ""Too Much of a Good Thing? The Relationship Between Number of Friends and Interpersonal Impressions on Facebook". Journal of Computer-Mediated Communication-Wiley online library. (n.d.). Retrieved March 16, 2023, from https://onlinelibrary.wiley.com/doi/full/10.1111/j.1083-6101.2008.00409.x

Lee, 2022 "What Research Says About the Best Profile Picture". Buffer.com. From(https://buffer.com/library/best-profile-picture-science-research-psychology/)

Appendix 1 – Preview of Survey

Demographic question for treatment and control group.

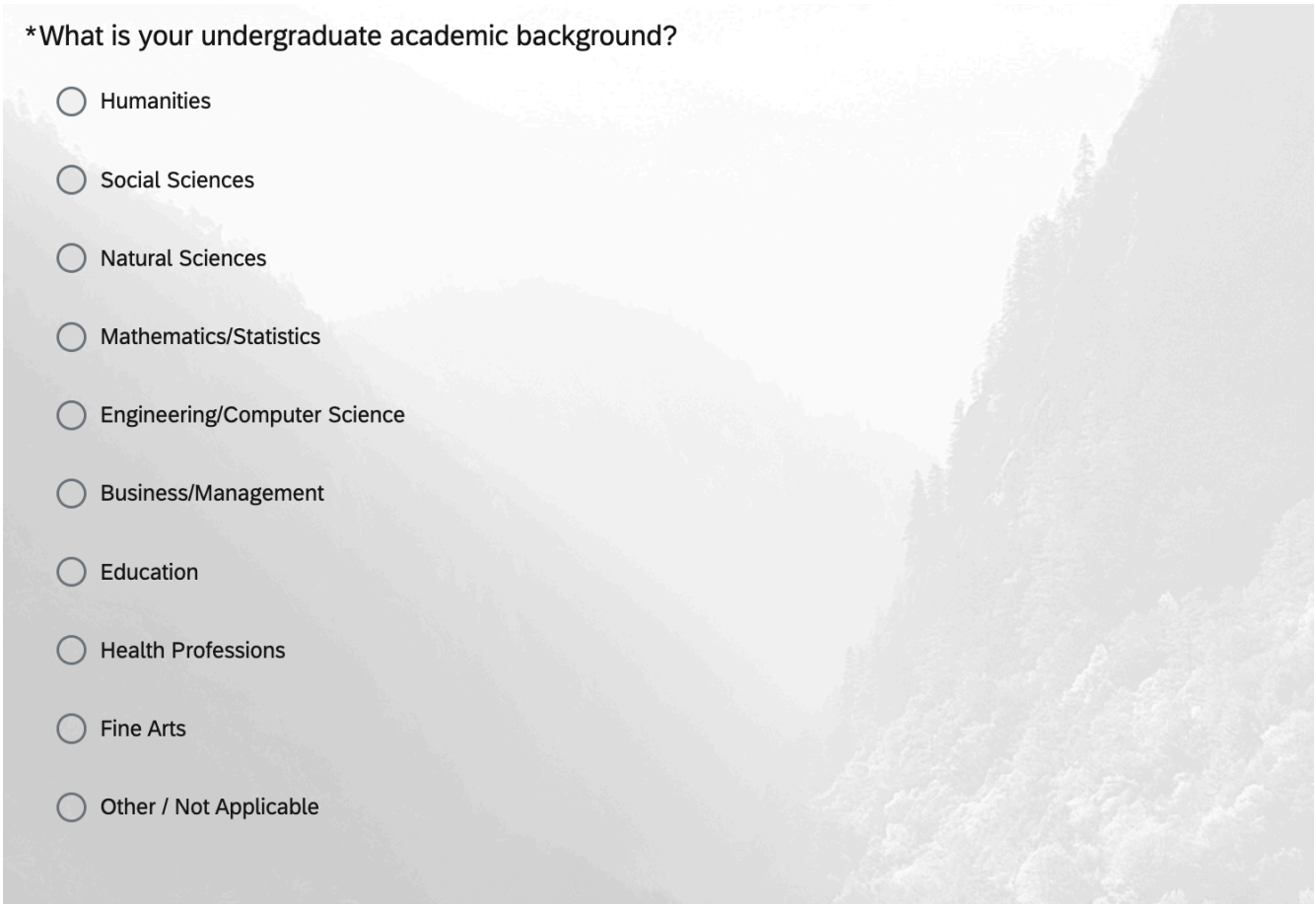


*Which ethnicity do you identify with?

- ◯ Hispanic or Latino
- ◯ White
- ◯ Black
- ◯ Asian
- ◯ Native American or Alaskan Native
- ◯ Native Hawaiian or other Pacific Islander
- ◯ Two or more races

*What is the highest level of degree you are pursuing or have already obtained?

- ◯ High school diploma or equivalent
- ◯ Associate degree
- ◯ Bachelor's degree
- ◯ Master's degree
- ◯ Doctorate degree
- ◯ Other



For treatment group, they will see:

From now on, you will face 10 different online conversations, each followed by the same question that you will need to answer.

Conversation #1

*On a scale of 1 to 10, how friendly do you think the person **on the left** is during the conversation above?
(1 = extremely unfriendly, 5 = neutral, 10 = extremely friendly)

extremely unfriendly neutral extremely friendly

Click to write Choice 1

1 10

Conversation #2

*On a scale of 1 to 10, how friendly do you think the person **on the left** is during the conversation above?
(1 = extremely unfriendly, 5 = neutral, 10 = extremely friendly)

extremely unfriendly neutral extremely friendly

Click to write Choice 1

1 10

Conversation #3

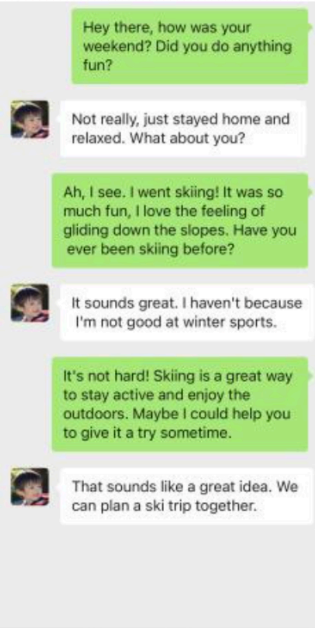
*On a scale of 1 to 10, how friendly do you think the person **on the left** is during the conversation above?
(1 = extremely unfriendly, 5 = neutral, 10 = extremely friendly)

extremely unfriendly neutral extremely friendly

Click to write Choice 1

1 10

Conversation #4



*On a scale of 1 to 10, how friendly do you think the person **on the left** is during the conversation above?
(1 = extremely unfriendly, 5 = neutral, 10 = extremely friendly)

extremely unfriendly neutral extremely friendly

Click to write Choice 1

1 10

Conversation #5

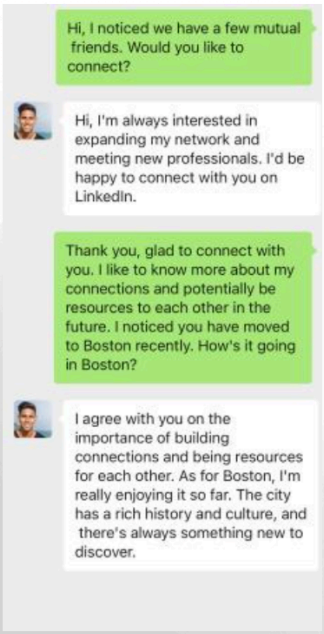
*On a scale of 1 to 10, how friendly do you think the person **on the left** is during the conversation above?
(1 = extremely unfriendly, 5 = neutral, 10 = extremely friendly)

extremely unfriendly neutral extremely friendly

Click to write Choice 1

1 10

Conversation #6



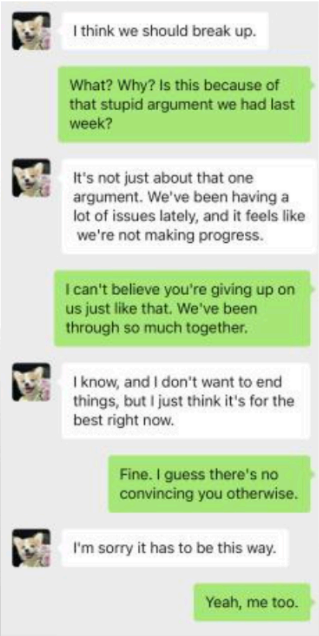
*On a scale of 1 to 10, how friendly do you think the person **on the left** is during the conversation above?
(1 = extremely unfriendly, 5 = neutral, 10 = extremely friendly)

extremely unfriendly neutral extremely friendly

Click to write Choice 1

1 10

Conversation #7



*On a scale of 1 to 10, how friendly do you think the person **on the left** is during the conversation above?
(1 = extremely unfriendly, 5 = neutral, 10 = extremely friendly)

extremely unfriendly neutral extremely friendly

Click to write Choice 1

1 10

Conversation #8

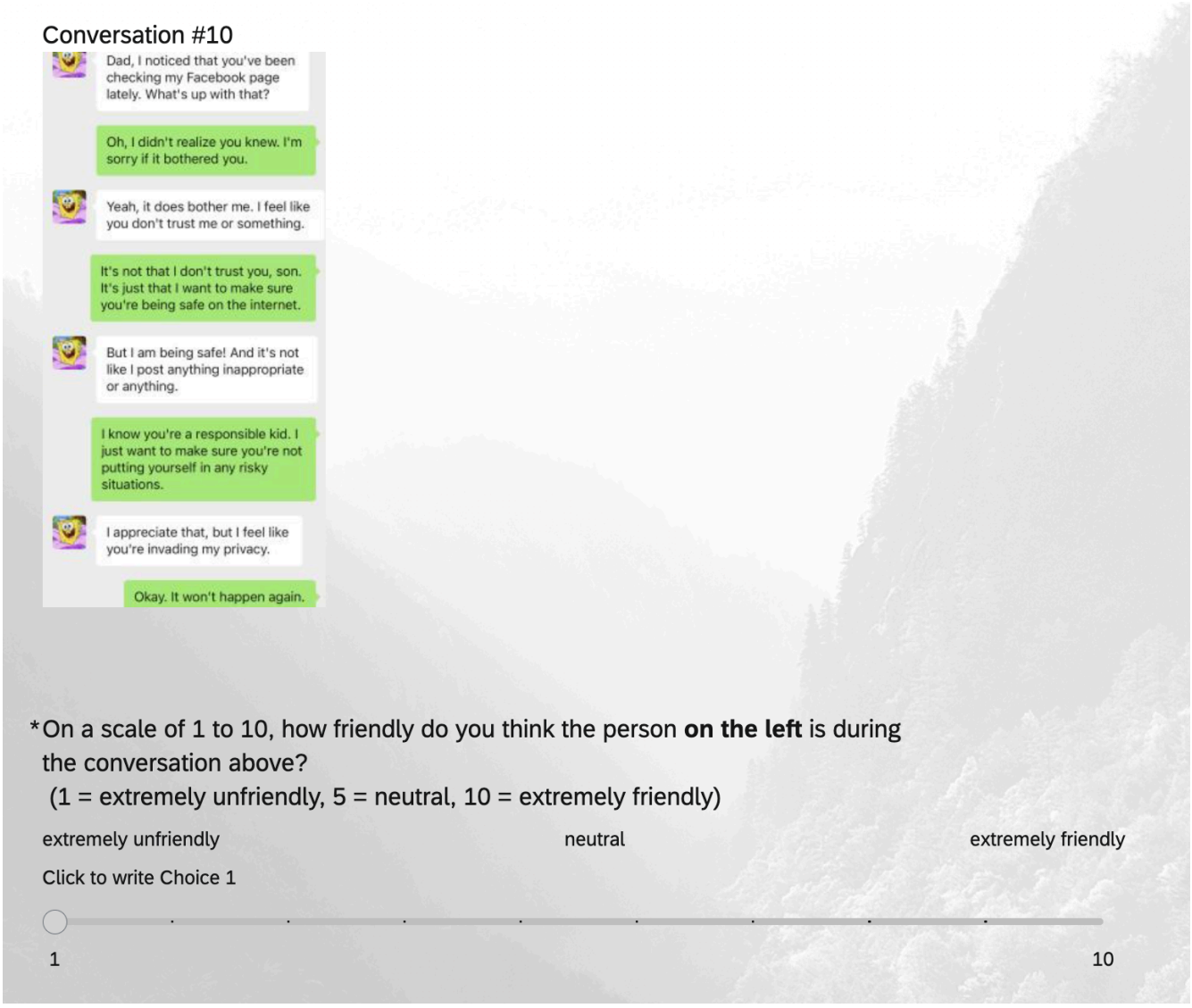
*On a scale of 1 to 10, how friendly do you think the person **on the left** is during the conversation above?
(1 = extremely unfriendly, 5 = neutral, 10 = extremely friendly)

extremely unfriendly neutral extremely friendly

Click to write Choice 1

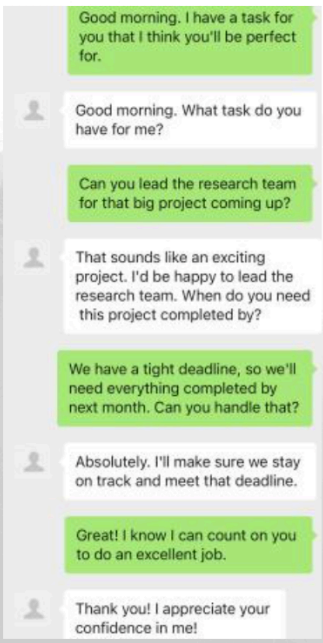
1 10

[](hahah%20/12.png



For control group, they will see:

Conversation #1



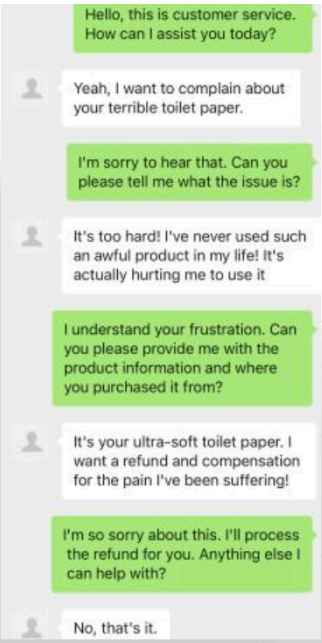
*On a scale of 1 to 10, how friendly do you think the person **on the left** is during the conversation above?
(1 = extremely unfriendly, 5 = neutral, 10 = extremely friendly)

extremely unfriendly neutral extremely friendly

Click to write Choice 1

1 10

Conversation #2



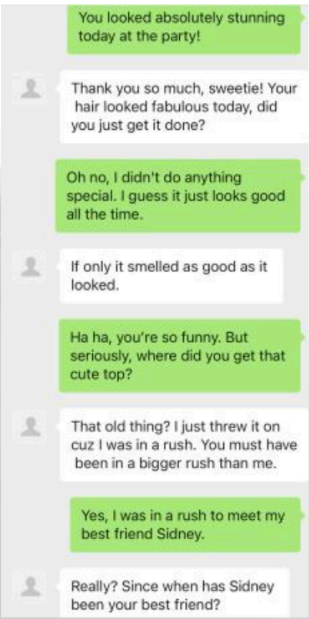
*On a scale of 1 to 10, how friendly do you think the person **on the left** is during the conversation above?
(1 = extremely unfriendly, 5 = neutral, 10 = extremely friendly)

extremely unfriendly neutral extremely friendly

Click to write Choice 1

1 10

Conversation #3



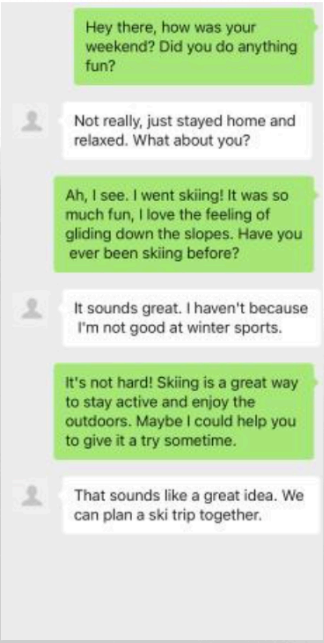
*On a scale of 1 to 10, how friendly do you think the person **on the left** is during the conversation above?
(1 = extremely unfriendly, 5 = neutral, 10 = extremely friendly)

extremely unfriendly neutral extremely friendly

Click to write Choice 1

1 10

Conversation #4



*On a scale of 1 to 10, how friendly do you think the person **on the left** is during the conversation above?
(1 = extremely unfriendly, 5 = neutral, 10 = extremely friendly)

extremely unfriendly neutral extremely friendly

Click to write Choice 1

1 10

Conversation #5

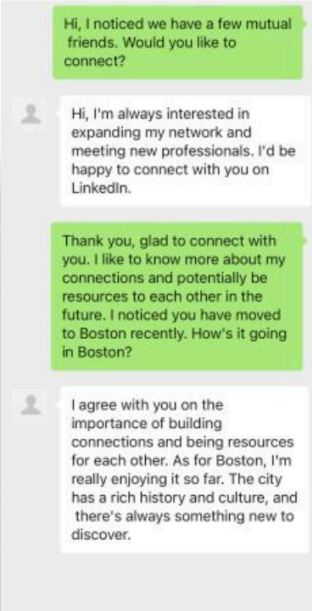
*On a scale of 1 to 10, how friendly do you think the person **on the left** is during the conversation above?
(1 = extremely unfriendly, 5 = neutral, 10 = extremely friendly)

extremely unfriendly neutral extremely friendly

Click to write Choice 1

1 10

Conversation #6



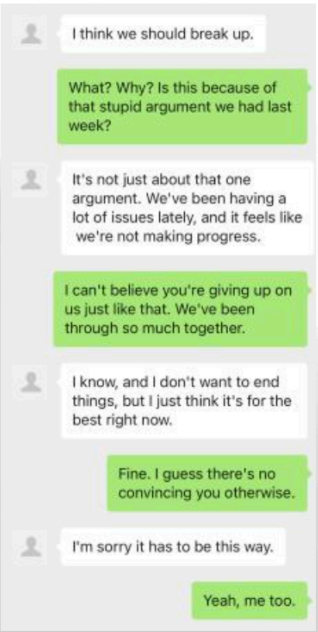
*On a scale of 1 to 10, how friendly do you think the person **on the left** is during the conversation above?
(1 = extremely unfriendly, 5 = neutral, 10 = extremely friendly)

extremely unfriendly neutral extremely friendly

Click to write Choice 1

1 10

Conversation #7



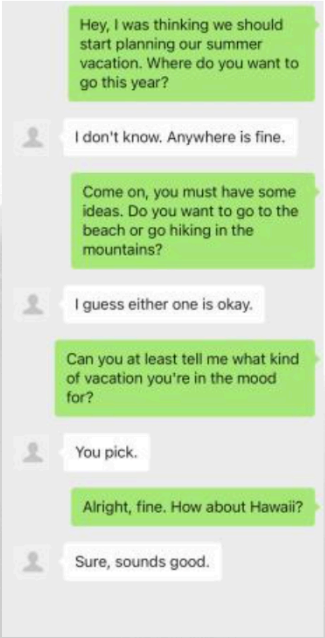
*On a scale of 1 to 10, how friendly do you think the person **on the left** is during the conversation above?
(1 = extremely unfriendly, 5 = neutral, 10 = extremely friendly)

extremely unfriendly neutral extremely friendly

Click to write Choice 1

1 10

Conversation #8



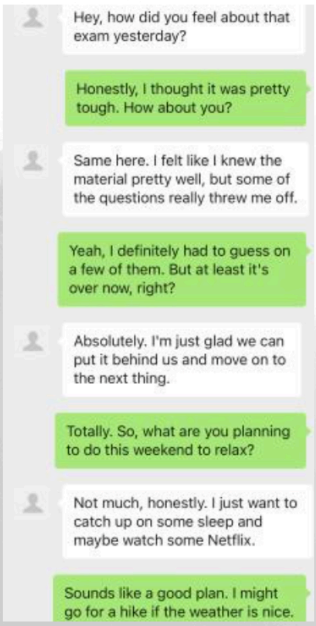
*On a scale of 1 to 10, how friendly do you think the person **on the left** is during the conversation above?
(1 = extremely unfriendly, 5 = neutral, 10 = extremely friendly)

extremely unfriendly | neutral | extremely friendly

Click to write Choice 1

1 10

Conversation #9



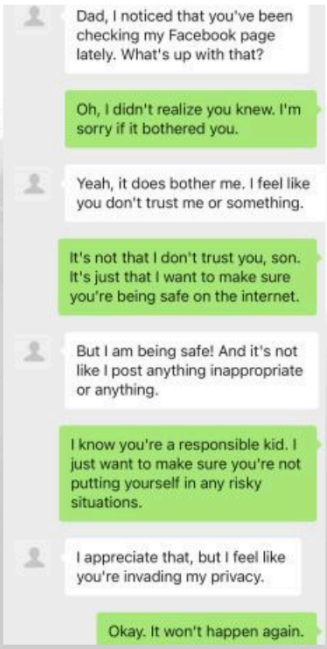
*On a scale of 1 to 10, how friendly do you think the person **on the left** is during the conversation above?
(1 = extremely unfriendly, 5 = neutral, 10 = extremely friendly)

extremely unfriendly neutral extremely friendly

Click to write Choice 1

1 10

Conversation #10



*On a scale of 1 to 10, how friendly do you think the person **on the left** is during the conversation above?
(1 = extremely unfriendly, 5 = neutral, 10 = extremely friendly)

extremely unfriendly neutral extremely friendly

Click to write Choice 1

1 10

Appendix 2

Below are our data processing code.

```
#loading and data cleaning
survey <- fread('830Online Conversation.csv') #choice value
#merge values
survey$conversation1 <- ifelse(survey$treatment == "1",
survey$Conversation_1, survey$Conversation_1_1)
survey$conversation2 <- ifelse(survey$treatment == "1",
survey$Conversation_2, survey$Conversation_2_1)
survey$conversation3 <- ifelse(survey$treatment == "1",
survey$Conversation_3, survey$Conversation_3_1)
survey$conversation4 <- ifelse(survey$treatment == "1",
survey$Conversation_4, survey$Conversation_4_1)
survey$conversation5 <- ifelse(survey$treatment == "1",
survey$Conversation_5, survey$Conversation_5_1)
survey$conversation6 <- ifelse(survey$treatment == "1",
survey$Conversation_6, survey$Conversation_6_1)
survey$conversation7 <- ifelse(survey$treatment == "1",
survey$Conversation_7, survey$Conversation_7_1)
survey$conversation8 <- ifelse(survey$treatment == "1",
survey$Conversation_8, survey$Conversation_8_1)
survey$conversation9 <- ifelse(survey$treatment == "1",
survey$Conversation_9, survey$Conversation_9_1)
survey$conversation10 <- ifelse(survey$treatment == "1",
survey$Conversation_10, survey$Conversation_10_1)
survey$Q6 <- ifelse(survey$treatment == "1",
survey$Q6, survey$Q6_1)
#rename columns
setnames(survey, "Q1", "gender")
setnames(survey, "Q2", "age")
setnames(survey, "Q3", "ethnicity")
setnames(survey, "Q4", "education_level")
setnames(survey, "Q5", "undergrad_major")
```

```
setnames(survey, "Q6", "survey_purpose_guess")
#make a copy on cleaned dataset
library(dplyr)
survey %>%
select(treatment,
conversation1, conversation2,conversation3,
conversation4, conversation5, conversation6,
conversation7,conversation8, conversation9,
conversation10,gender,age,ethnicity,
education_level,undergrad_major,survey_purpose_guess) ->  df
#convert data type
df$treatment <- as.numeric((df$treatment))
df$conversation1 <- as.numeric((df$conversation1))
df$conversation2 <- as.numeric((df$conversation2))
df$conversation3 <- as.numeric((df$conversation3))
df$conversation4 <- as.numeric((df$conversation4))
df$conversation5 <- as.numeric((df$conversation5))
df$conversation6 <- as.numeric((df$conversation6))
df$conversation7 <- as.numeric((df$conversation7))
df$conversation8 <- as.numeric((df$conversation8))
df$conversation9 <- as.numeric((df$conversation9))
df$conversation10 <- as.numeric((df$conversation10))
#encoding categorical variable in randomization check and merge table
encoded_eth <- model.matrix(~ ethnicity + 0, data = df)
encoded_gender <- model.matrix(~ gender + 0, data = df)
encoded_age <- model.matrix(~ age + 0, data = df)
encoded_edu <- model.matrix(~ education_level + 0, data = df)
encoded_major <- model.matrix(~ undergrad_major + 0, data = df)
df1 <- cbind(df, encoded_eth, encoded_gender, encoded_edu, encoded_major)
setnames(df1, "ethnicityNative American or Alaskan Native", "ethnicityNative_American_or_Alaskan_
setnames(df1, "ethnicityTwo or more races","ethnicityTwo_or_more_races")
```