Pre-Analysis Plan
Algorithmic Personalization and Digital Addiction: A Field Experiment on Douyin (TikTok)

Qianyun (Carole) Gao[a], Yu Gao[b], Juanjuan Meng[b], and Shizhe Yu[c]

[a]Department of Economics, University of Warwick

[b]Guanghua School of Management, Peking University

[c]Department of Economics, University of Wisconsin-Madison

1 Introduction

Recent developments in social media have influenced how people communicate and entertain. Social media has made it easier to connect with others and disseminate information. However, studies have shown that social media addiction can adversely affect individuals' well-being, including worsening mental health and increasing polarization (Lee, Choi, Kim, & Kim, 2014; Allcott, Braghieri, Eichmeyer, & Gentzkow, 2020; Hunt, Marx, Lipson, & Young, 2018).

Algorithm plays a central role in modern social media and entertainment, enhancing the attractiveness of digital platforms. However, this may lead to addiction, , making it difficult for some users to reduce their usage (McCluskey, 2022). Research has explored the self-control problem behind digital addiction (Allcott, Gentzkow & Song, 2022), and found that soft approaches can be effective (Hoong, 2021).

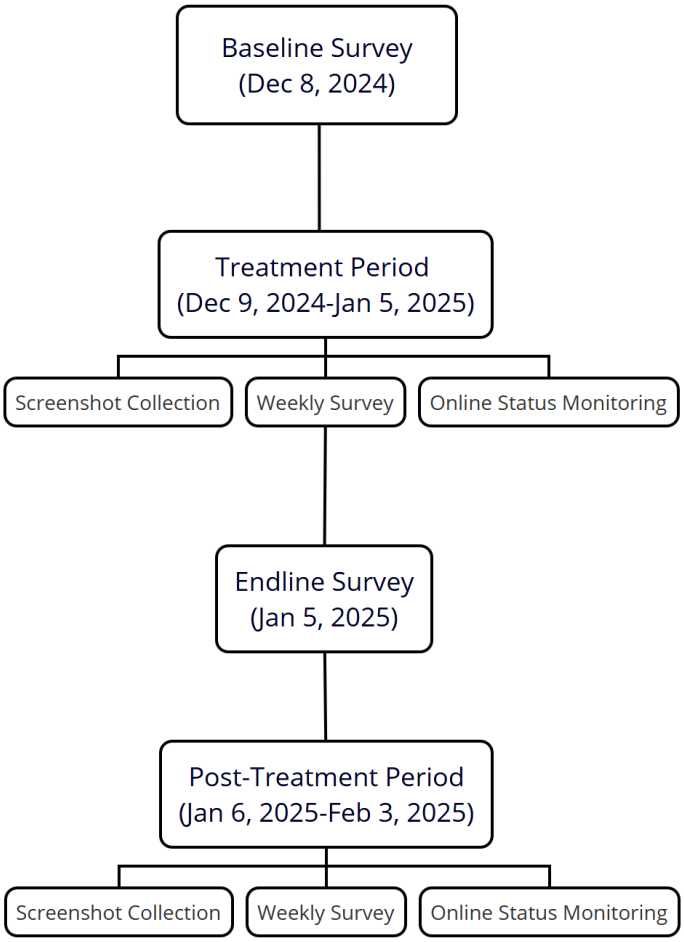
This study seeks to explore the effects of algorithms on participants' short-video consumption on Douyin, the Chinese counterpart of TikTok, and the platform's effect on subjective well-being, cognitive abilities, and opinion polarization. Recent research indicates that continuous exposure to short video streaming on platforms like TikTok can negatively influence users' attention spans and mental health (Lorenz-Spreen,

Mønsted, Hövel, & Lehmann, 2019). We will conduct a randomized experiment, and participants in three treatment groups will be instructed either to use Douyin as they like (Control), log out of their Douyin accounts and not receive personalized content (No Personalization), or to abstain from using Douyin entirely (No Usage). The outcomes will be evaluated through online surveys, direct monitoring of Douyin accounts, and analysis of participants' smartphone screen time records.

2 Experiment Design

Figure 1 presents the experimental design and timeline.

Figure 1: Experiment Overview



2.1 Recruitment and Baseline Survey

We recruit participants through one of the largest Chinese online survey platform.

The following content is presented at the start of our participant recruitment surveys (see Appendix A.1 for the original Chinese version):

Dear participant,

We are a research team from Guanghua School of Management, Peking University. We are conducting an academic study and would like to invite you to fill out a questionnaire.

There are five screening questions. If a participant is above 18, has installed Douyin and uses their smartphone as the main device to watch short videos, has used Douyin in the past 30 days, and has a smartphone from the following four brands: Huawei, vivo, OPPO, and Xiaomi[1], then they will be directed to a brief introduction of the main experiment and a consent form. After consenting, they will complete a series of questions regarding their demographic background and Douyin usage.

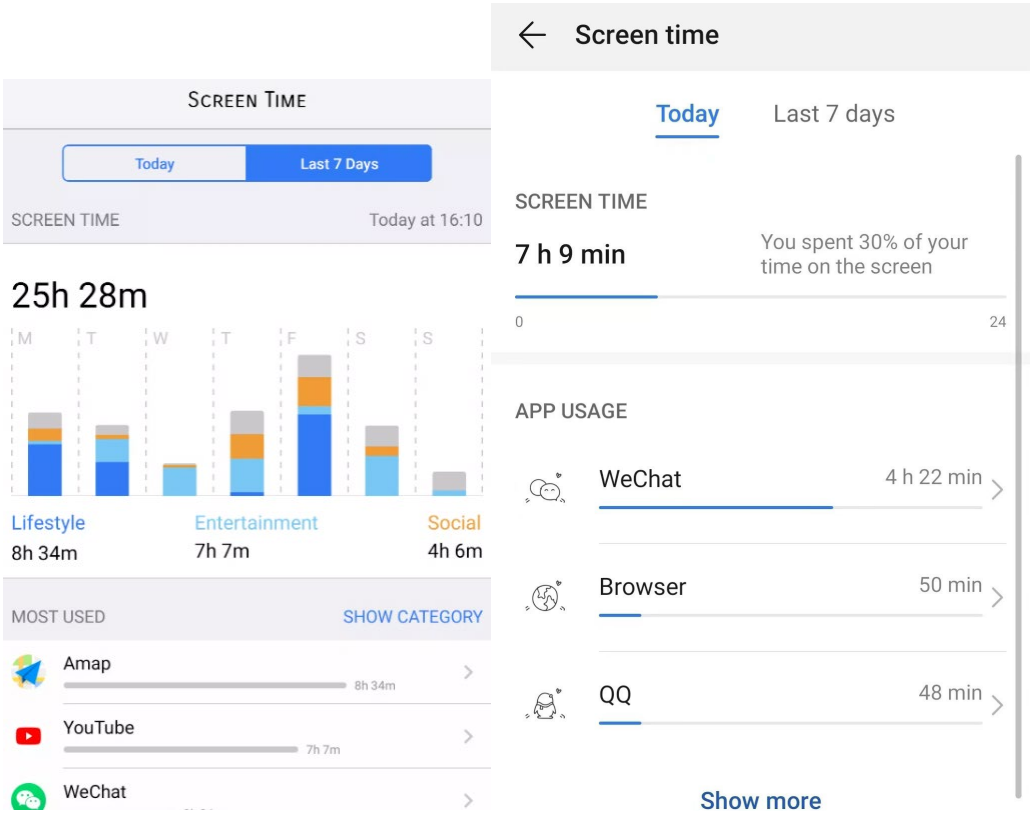
At the end of the recruitment survey, we give them instructions on 1) how to turn on the Screen Time function on the smartphone and take screenshots, and 2) how to turn on Douyin's "Display Status" function.

[1] Because each brand has a different Time Management interface, this aims to minimize training costs. According to IDC (Link: https://www.idc.com/getdoc.jsp?containerId=prAP52467524), in 2024 Q2, Huawei (including Honor), vivo, OPPO and Xiaomi take up 32.6%, 18.5%, 15.7% and 14.0% of the Chinese market respectively (80.8% in total).

2.2 Screen Time Function

The Screen Time function records the amount of time users spend on each application. By default, users can have a weekly report of their phone usage. It is available on both Android and HarmonyOS of Huawei. Figure 2 presents sample screenshots of the Android and HarmonyOS systems.

Figure 2: The Screen Time Function for Android (left), and Harmony OS (right)



2.3 Douyin's Display Status Function

The "Display Status" is a function of Douyin. If a user turns on this function, as a mutual follower, we can see their online status. When the user is online, the status is "online"; if not, the status is "online within the last xx minutes/hours". If a user has not logged in for more than 48 hours, the status is null.

During the experiment, a computer program automatically checks whether treatment group participants have logged in to Douyin by reading their online status eight times a day. We can also detect if a user disables the "Display Status" function.

2.4 Screenshots and Weekly Surveys

For four consecutive weeks (December 8, 2024-January 5, 2025), on each Sunday between 0:00 and 24:00, participants are required to complete a weekly survey, which include uploading screenshots of their mobile phone's screen time weekly summary and questions on subjective well-being, which we further explain in Section 3. Only screenshots that are not edited and include all apps used for more than 10 minutes are considered valid.

2.5 Treatment Assignment

One day before the treatment period, after completing the baseline survey, participants are informed of their specific treatment status. The texts are as follows (see Appendix A.2 for the original Chinese version):

Dear participant,

Thank you for your participation. Below are the requirements for the experiment. Please read carefully:

[only sent to the Control group]

Requirements:

During the experiment (Dec 9th 00:00 to Jan 5th 24:00):

You can use Douyin as you like, and please keep the "Display Status" function of Douyin active and keep following our account.

You will need to complete a weekly survey (which takes less than 5 minutes each week for 4 weeks) and upload screenshots of your screen time via the survey every Sunday between 0:00 and 24:00.

[only sent to the No Personalization group]

Requirements:

During the experiment (Dec 9th 00:00 to Jan 5th 24:00):

You should log out of your Douyin account, but you can use Douyin as a guest.

Please keep the "Display Status" function of Douyin active keep following our account.You will need to complete a weekly survey (which takes less than 5 minutes each week for 4 weeks) and upload screenshots of your app usage via the survey every Sunday between 0:00 and 24:00.

[only sent to the No Usage group]

Requirements:

During the experiment (Dec 9th 00:00 to Jan 5th 24:00):

You should not use Douyin. The research team will check your account's status.

Please keep the "Display Status" function of Douyin active and keep following our account.You will need to complete a weekly survey (which takes less than 5 minutes each week for 4 weeks) and upload screenshots of your app usage via the survey every Sunday between 0:00 and 24:00.

2.6 Endline Survey

At the end of the treatment period, we send participants the endline survey. The endline survey includes similar questions from the baseline survey, which we detail in Section 3.

2.7 Post-experiment Monitoring

After the treatment period, we will keep monitoring the online status of the Douyin accounts of participants for four weeks. Participants need to upload screenshots and complete surveys every Sunday during this period.

2.8 Payment

All the payments are sent to participants through the online platform.

All participants receive 10 yuan for completing the baseline survey, and 5 yuan for completing each weekly survey.

Participants who comply with their treatment requirements each week receive an

additional 10 yuan (not logging in to the Douyin account for the “No Personalization” group, or not using Douyin for the “No Usage” group).

Participants receive 10 yuan after completing the endline survey.

3 Survey Outcome Variables

We consider the following main outcome variables: time allocation, subjective wellbeing, cognitive skills, and issue polarization, as presented in Table 1. The original Chinese version is presented in Appendix B.

3.1 Use of Time

3.1.1 Douyin

we monitored participants' Douyin use through two methods: by directly checking participants' Douyin online status, and by asking them to provide weekly screenshots of their Screen Time reports.

3.1.2 Other App Activities

To assess participants' time allocation to Douyin and other apps, we collected screenshots of their weekly Screen Time reports through weekly surveys. We categorize the apps into eight categories, as presented in Table 1 Column (3).

3.1.3 Offline Activities

In the past month, compared to the situation before this experiment, did you spend your time differently?

[1(Significantly less)-5(Significantly more)]

- *Watching TV or movies*
- *Reading books newspapers or magazines*
- *Sports and exercises*
- *Other Hobbies (such as cooking or crafts)*
- *Spending time with family*
- *Spending time with friends*

Table 1: Outcome Variables

(1)	(2)	(3)	(5)
Time Allocation	Douyin	Time spent on Douyin App	Screenshots from Weekly Surveys
		Douyin Login Frequency	Daily Direct Monitoring of "Online Status"
	Other App Activities	Other Short-video Apps; Long-video Apps; Entertainment; Social Media and News; Life Service; Study and Efficiency; Others	Screenshots from Weekly Surveys
	Offline Activities	TV or Movies; Reading; Sport Exercise; Other Hobbies; Family; Friends	Endline Survey
Subjective Well-being		Loneliness; Happiness; Stress; Anxiety; Concentration; Sleep Quality	Baseline and Weekly Surveys
Cognitive Competence	Subjective Productivity	Self-perceived Work and Study Efficiency	Baseline and Endline Survey
	Objective Cognitive Skills	Raven's Matrices; Stroop Test	Baseline and Endline Survey
Issue Polarization		Hard work vs. Luck; Economy vs. Environment; Government Regulation of Firm; Military vs. Diplomacy	Endline Survey

3.2 Subjective Well-being

Below are six statements that you may agree or disagree with. Indicate your opinion on each item based on your feeling this month.

[1 (not at all) - 10 (exactly)]

- *I felt lonely.*
- *I felt happy.*
- *I felt stressed.*
- *I felt anxious.*
- *I could not focus on what I am doing.*
- *I slept well.*

3.3 Cognitive Skills

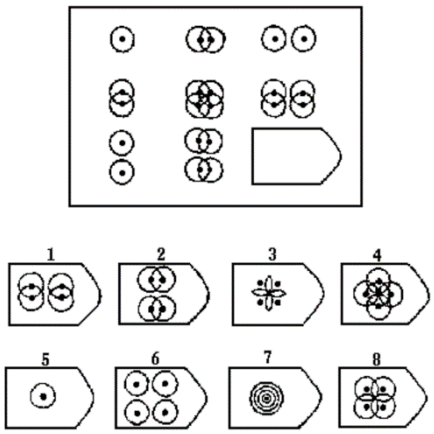
3.3.1 Self-perceived Work and Study Efficiency

Think of your worst level of efficiency ever and your best possible efficiency and rate how efficient you felt you were this month compared to your worst ever (1) and best possible (10).

For example: You feel that you are so efficient this month that it is close to being your best possible performance. In this case, your answer would fall somewhere between 8 and 9.

3.3.2 Raven's Matrices

Figure 2: Example of Raven's Matrices



Each question below has a certain theme, but part of the main pattern in each large image is missing. Please select the most suitable small image from the options provided under each question to complete the pattern logically and coherently.

3.3.3 Stroop Test

Figure 3: Example of Stroop Test

Red

You will need to complete a simple task.

You will see a series of words displayed in different colors. Your task is to quickly and accurately identify the color in which each word is displayed, rather than the meaning of the word itself.

3.4 Issue Polarization

Please share your opinions on the following questions. On the scale, 1 represents complete agreement with the statement on the left, while 10 represents complete agreement with the statement on the right. Please indicate your opinion on the scale.

[1 (statement on the left) - 10 (statement on the right)]

• Government regulation of business usually does more harm than good vs. Government often does a better job than people give it credit for

• The best way to ensure peace is through military strength vs. Good diplomacy is the best way to ensure peace

• In the long run, hard work usually brings a better life vs. Hard work doesn't generally bring success—it's more a matter of luck and connections

• Protecting the environment should be given priority, even if it causes slower economic growth and some loss of jobs vs. Economic growth and creating jobs should be the top priority, even if the environment suffers to some extent.

We also consider the following secondary outcomes for exploratory purposes.

3.5 Algorithm

3.5.1 Attitude towards Algorithm

Douyin provides a feature to "turn off personalized recommendation algorithms," which means you will no longer receive content recommendations based on your personal preferences and usage habits. What is your choice regarding this feature?

• I am willing to turn off personalized recommendation algorithms.

• I am not willing to turn off personalized recommendation algorithms.

• Uncertain.

3.5.2 Change in Algorithm Quality

[1(strongly disagree)- 10(strongly agree)]

Based on your Douyin usage during the research period, please rate your level of agreement with the following statements. Use the scale where 1 represents "completely disagree" and 10 represents "completely agree."

- *During the research period, my dependence on Douyin changed.*
- *After the first month, I clearly changed my Douyin usage habits.*

• *After the first month, the quality of Douyin's recommendation algorithms declined, and the content was not as engaging.*

3.6 Guess of Experiment Aim

In your opinion, what is our experiment aim?

3.7 Time Preference

• *Assume you receive an income of 10,000 yuan today. You can choose to take a portion of the money one year later, and this portion will earn a 10% bonus. How much money will you choose to take one year later?*

• *Assume you will receive an income of 10,000 yuan one year later. You can choose to take a portion of the money two years later, and this portion will earn a 10% bonus. How much money will you choose to take two years later?*

• *Imagine you have the following lottery opportunity: Participating in the lottery gives you a 50% chance of losing 10 yuan and a 50% chance of winning 11 yuan. Would you like to participate?*

4 Analysis

4.1 Outcome Variable Construction

For survey outcomes, we reverse code negative statements such as loneliness, so that higher scores indicate more positive treatment effects. For the construction of three composite indexes - the Offline Activity Index, Cognitive Competence Index, and Issue Polarization Index -- we utilized Anderson (2008)'s Inverse Covariance Weighting method. For each index, this method uses the inverse of the covariance matrix of the contributing variables (as outlined in Table 1 Column 3) to assign weights.

The Subjective Well-being Index is an exception, and it is not constructed using the Inverse Covariance Weighting method. Instead, we create this index by summing

up the relevant outcome variables in this category. This approach is based on the literature on the construction of the Subjective Well-being Index, which posits that each variable contributes equally and independently to the overall measure, therefore negating the need for weighting (Huppert et al., 2009). This method is also consistent with approaches used in previous research, such as that conducted by Allcott et al. (2020).

Finally, all outcomes are standardized relative to the Control group (except for cellphone app use time variables and Douyin login frequency, for interpretation purpose), ensuring that the Control group has a standard deviation of 1.

4.2 Empirical Strategy

To estimate the treatment effects, we employ the following regression equation:

$$Y_{it} = \alpha + \gamma Y_{i0} + \tau_1 T_1 + \tau_2 T_2 + \delta X_i + \varepsilon_{it}, \quad (1)$$

where Y_{it} represents the outcome variable for individual *i* at period t, while Y_{i0} is its baseline value if the outcome is also measured before the treatments, otherwise Y_{i0} is not included. Specifically, we have weekly measurements for variables such as Douyin time, usage time for other apps, and subjective well-being, which include pre-treatment baseline values. The cognitive competence variable also has a baseline value. For Douyin login frequency, Y_{it} is measured daily throughout the pre-treatment, treatment and post-treatment periods and is presented as the percentage of checks that individual *i* is observed online. Measurements for offline activities and issue polarization are only taken during the endline survey.

Define T_1 and T_2 as indicators for the No Personalization treatment and No Usage treatment. X_i is a vector of control variables measured in the baseline survey. We use robust standard errors in all regressions. For dependent variables measured multiple times, we use panel regressions with time fixed effect and cluster robust standard errors at the time level to accommodate potential patterns possibly influenced by external events, trends or time-varying algorithms.

References

Allcott, Hunt, Matthew Gentzkow, and Lena Song. 2022. "Digital Addiction." *American Economic Review*, 112 (7): 2424–63.

Allcott, H., Braghieri, L., Eichmeyer, S., & Gentzkow, M. (2020). The welfare effects of social media. *American Economic Review*, *110*(3), 629–76.

Anderson, M. L. (2008). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the abecedarian, perry preschool, and early training projects. *Journal of the American Statistical Association*, *103*(484), 1481-1495.

Dutcher, G., & Rodet, C. S. (2018). Which two heads are better than one? uncovering the positive effects of diversity in creative teams. *Journal of Economics & Management Strategy*.

Hoong, R. (2021). Self-control and smartphone use: An experimental study of soft commitment devices. *European Economic Review*, 140, 103924. doi:10.1016/j.euroecorev.2021.103924.

Hunt, M. G., Marx, R., Lipson, C., & Young, J. (2018). No more fomo: Limiting social media decreases loneliness and depression. *Journal of Social and Clinical Psychology*, *37*(10), 751–768.

Huppert, F. A., Marks, N., Clark, A., Siegrist, J., Stutzer, A., Vittersø, J., & Wahrendorf, M. (2009). Measuring Well-being Across Europe: Description of the ESS Well-being Module and Preliminary Findings. *Social Indicators Research*, *91*(3), 301–315. https://doi.org/10.1007/s11205-008-9346-0

Lee, J. K., Choi, J., Kim, C., & Kim, Y. (2014). Social media, network heterogeneity, and opinion polarization. *Journal of communication*, *64*(4), 702–722.

Lorenz-Spreen, P., Mønsted, B. M., Hövel, P., & Lehmann, S. (2019). Accelerating dynamics of collective attention. *Nature communications*, *10*(1), 1–9.

McCluskey, M. (2022, January 4). How addictive social media algorithms could finally face a reckoning in 2022. https://www.editorandpublisher.com/stories/how-addictivesocial- media-algorithms-could-finally-face-a-reckoning-in-2022,214063

Appendix-Original Chinese Version

A.1 Questionnaire for Recruitment

尊敬的被访者您好，我们是北京大学光华管理学院的研究团队，请您参与一项问卷调查。

A.2 Treatment Assignment

尊敬的受访者您好。

感谢您参与我们有关抖音使用习惯的研究。以下是下一阶段的实验要求，请您仔细阅读。

[Only sent to the Control group]

【本阶段要求】

本阶段期间（12月9日0:00-1月5日24:00）：

您可以自由使用抖音。

继续保留项目小组抖音好友，并持续开启“好友活跃状态”。

每周日0:00-24:00上传截图并回答简单问题（共四周，完成问卷所需时间在5分钟以内）。

【本阶段报酬】

上传截图并回答问题，每次可得5元，四周共20元。

[Only sent to the No Personalization group]

【本阶段要求】

本阶段期间（12月9日0:00-1月5日24:00）：

停止登录抖音账号，但可以在不登录的情况下使用抖音。

继续保留项目小组抖音好友，并持续开启“好友活跃状态”。

在周日0:00-24:00上传截图并回答简单问题(完成问卷所需时间在5分钟以内)。

【本阶段报酬】上传截图并回答问题，每次可得5元，四周共20元。

不登录抖音账号，每周可得10元，四周共40元。

[Only sent to the No Usage group]

【本阶段要求】

本阶段期间（12月9日0:00-1月5日24:00）：

停止使用抖音。

继续保留项目小组抖音好友，并持续开启“好友活跃状态”。我们会检查您的使用情况。

在周日0:00-24:00上传截图并回答简单问题（完成问卷所需时间在5分钟以内）。

【本阶段报酬】

上传截图并回答问题，每次可得5元，四周共20元。

不使用抖音，每周可得10元，四周共40元。

B Survey Questions Appendix – Original Chinese Version

B.1 Offline Activities

在过去的一个月，与之前的大多数情况相比，您在工作之外的时间投入了更多，更少，还是相同的时间在以下活动？

[1(更少)-5(更多)]

- 看电视或电影
- 阅读纸质书籍、报刊或杂志
- 体育锻炼
- 烹饪、手工等兴趣爱好
- 陪伴家人
- 与朋友聚会

B.2 Subjective Well-being

请回想您过去一周的生活，对以下描述与您状态的相符程度做出评价。

1(完全不符合)-10(完全符合)

- 我感到孤独
- 我感到快乐
- 我感到压力
- 我感到焦虑
- 我无法专注于正在做的事
- 我的睡眠质量很好

B.3 Cognitive Skills

B.3.1 Self-perceived Work and Study Efficiency

请将您这个月对自己平均工作学习效率的评估与您过去的最高水平和最低水平相比较。

例：如果您在本月感到自己的效率非常高，以至于接近您的最高水平，则您的答案应该位于8-9之间。

[1(最低水平)-10(最高水平)]

B.4 Issue Polarization

请您告诉我对下列问题的看法，量表中1表示完全同意左侧的看法，10表示完全同意右侧的看法，请在量表上标出您的看法。

- 努力工作通常能带来更好的生活-努力并不一定能成功，运气和关系更重要

[1(左边的看法)-10(右边的看法)]

- 经济增长有时候会伴随着对环境的破坏，但同时环境保护有时候会以放慢经济增长速度甚至增加失业为代价。我们应该：

[1(环境保护优先)-10(经济增长和就业优先)]

- 政府对企业的监管通常弊大于利-政府对企业的监管通常利大于弊

[1(左边的看法)-10(右边的看法)]

- 军事力量是确保和平的最佳方式-外交手段是确保和平的最佳方式

[1(左边的看法)-10(右边的看法)]

B.5 Algorithm

B.5.1 Attitude towards Algorithm

抖音平台提供了一个“关闭个性化推荐算法”的功能，意味着您将不会收到根据您的个人喜好和使用习惯推荐的内容。对于这个功能，您的选择是？

- 我愿意关闭个性化推荐算法
- 我不愿意关闭个性化推荐算法
- 不确定

B.5.2 Change in Algorithm Quality

请根据您研究期间的抖音使用情况，对您对下列说法的认同程度作出评价。

[1(完全不同意)-10(完全同意)]

- 在研究期间，我对抖音的依赖程度发生了变化。
- 在第一个月结束后，我明显改变了抖音的使用习惯。
- 在第一个月后，我的抖音的推荐算法质量有所下降，内容不那么吸引人了。

B.6 Guess of Experiment Aim

您认为我们在研究什么问题？

B.7 Time Preference

- 假设您今天有一笔10000元的收入。您可以选择一部分钱在一年后拿到，且这部分钱会有10%的红利。请问您会选择多少钱在一年后拿到？
- 假设您一年后会有一笔10000元的收入。您可以选择一部分钱在两年后拿到，且这部分钱会有10%的红利。请问您会选择多少钱在两年后拿到？
- 假如您有以下的抽奖机会：

参加抽奖后，您有50%的可能会损失10元，同时也有50%的机会赢得11元。
请问您是否愿意参加？