The study of how people of different ages react to different CAPTCHAs

Zehao Du
duze@oregonstate.edu
Oregon State University
Corvallis, Oregon

Jiawei Luan
luanj@oregonstate.edu
Oregon State University
Corvallis, Oregon

ACM Reference Format:
Zehao Du and Jiawei Luan. 2021. The study of how people of different ages react to different CAPTCHAs. In *Proceedings of* . ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 ABSTRACT

Due to the prevalence of web browsing, it is common for users of all age groups to encounter CAPTCHAs. We want to know how people of different ages react to different CAPTCHAs. It ease the user's experience and save more time when they are completing the CATPCHA.

In this study, we made a quiz which has four different types of CAPTCHAs and separated them into four sets according to their type. Then, we recruited two different ages groups to complete the quiz and one questionnaire. We used R studio to analyze and organize data to achieve our goal which to find a CAPTCHA format that is the most easily used or the most user-friendly for the age groups covered by this study. A relatively suitable CAPTCHA would entail using less response time and having a high accuracy rate. At the same time, participants have fewer questions when finishing the task.

After analyzing the data, we believe that the calculation question CAPTCHA is more suitable for the two age groups in the study.

2 INTRODUCTION

With the popularity of the Internet, more and more operations have been implemented on web pages. Browsing the web has gradually become something that people do in their daily lives. At the same time, due to the wide usage of web browsing, users of different age groups use CAPTCHA more frequently. This article will study how people of different ages react differently when using different CAPTCHAs. The goal is to find a CAPTCHA format that is relatively suitable for more age groups (the age groups covered by this study).

In order to better achieve the goal of this study, this study summarized the questions which makes this study have more specific data and information. Then we can analyze more directly to get the conclusion. To prevent confusion based on or caused by language and culture, Users will be given CAPTCHAs and questionnaire in their native language. The following are the research questions we relied on in our research.

RQ1: How does the same type of CAPTCHA perform for different ages of people?

Who has longer reaction time?

Who has lower grade?

RQ2: How do users of the same age group perform on the different types of CAPTCHAs?

Which type of CAPTCHA enables participants to gain high grade?

Which type of CAPTCHA enables participants to spend less time finishing CAPTCHAs?

RQ3: Which type of CAPTCHA causes the low mental load to participants?

After finishing the task, which type of CAPTCHA makes the participant mentally feel easier to recognize?

The first research question gets the responses of users from different age groups according to the same form of CAPTCHA. It analyzed the impact of CAPTCHA on users of different age groups and the factors that may lead to prolonged passage time.

The second research question was posed to single age groups reacting to different forms of CAPTCHAs. In order to find out what impact different CAPTCHA have on users of the same age and the factors that may lead to prolonged passage time.

The third research question analyzes which CAPTCHA will make users feel more comfortable. It is also an important factor to evaluate a good CAPTCHA.

Through the integration and analysis of the data obtained from the test of the above research questions, the most user-friendly CAPTCHA is determined from different angles. A good CAPTCHA means that participants use less response time to pass the CAPTCHA and the accuracy rate is high. At the same time, participants have fewer questions when finishing the task.

3 METHODOLOGY

To evaluate how the type of CAPTCHA affects different age groups, we conducted an experiment comparing the response time and grades via different ages of people using different CAPTCHAs. There are four types of CAPTCHAs: selecting the picture the title required, inputting the irregular words in the picture, answering the calculation question in the picture, and inputting the irregular words and numbers in the picture. From this we will answer the three research questions.

Participants were assigned to a group by age. Each participant finished four tasks.

, ,

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM…$15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

3.1 Participants and Treatments

The experiment is a two-by-four factorial design. Participants are between-subject factors with two levels:20<age<35 and 55<age<70, and CAPTCHA is the within-subject factor with four types: selecting picture the title required, inputting the irregular words and numbers in the picture, answering the question in the picture,input the irregular words in the picture. We recruited volunteers through email and social media applications. (Wechat/Facebook/Weibo) Finally, we recruited 8 participants and assigned them into two treatment groups according to their age.

It is a within-subjects study and is held online via zoom. Firstly, we informed the purpose of the study participants and obtained their consent. Secondly, we handed out the tutorial and let them read and think aloud. When the group sessions with multiple participants, (according to the schedules of participants, it might have the situation that more than 1 participant have the same available time) If they have questions, they will send the text message, or we will bring them to a new session and unmute them and to answer their questions. Thirdly, we introduced each task and taught participants how to use Quizlet, and helped them set up the experiment environment. Then we did a demo task and started a 20-minute learning session. This research was conducted remotely through zoom, we recorded the audio and screen during the whole session.

Every participant did every treatment (CAPTCHAs). Each participant must complete both CAPTCHA tasks and a questionnaire. Each CAPTCHA task's order is random and there are different types of CAPTCHAs. Thus, it reduced any biases of anyone who might be good at one kind of CAPTCHA(counterbalancing).

Table 1 illustrates the demographic information of participants.

Age Group	Task1	Task2	Task3	Task4
The younger group (20<age<35) average age =23.4	CAPTCHA set1	CAPTCHA set2	CAPTCHA set3	CAPTCHA set4
The older group (55<age<70) average age =66.9	CAPTCHA set1	CAPTCHA set2	CAPTCHA set3	CAPTCHA set4

Table 1

3.2 Tasks

We created four sets for the four types of CAPTCHAs in Quizlet. Each participant took four tasks in random order. Each task included 10 CAPTCHAs. Every participant's task was the same. The CAPTCHAs were updated after participants input the answer. If they input the wrong answer, they will need to click the continue button and pass the CAPTCHA. We recorded the process of each task. For example, screen recording when they are doing the tasks. After completing the tasks, we sent a questionnaire to the participants, letting them answer some questions such as "Which CAPTCHA is more difficult or spend more time."

a.Each task is recorded

b. The participant must understand the submit.

c. The participants are instructed to maximize detail.

Captcha task group:

Task group: Each task includes 10 CAPTCHAs from the same type.

Task 1:

Type of CAPTCHAs: inputting the irregular words and numbers in the picture

Figure 1: type of CAPTCHA in task1

Task 2:

Type of CAPTCHAs: inputting just the irregular words

Figure 2: type of CAPTCHA in task2

Task 3:

Type of CAPTCHAs: selecting picture the title required

Figure 3: type of CAPTCHA in task3

Task 4:

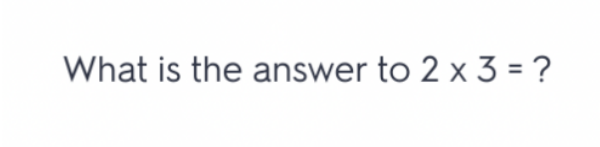

Figure 4: type of CAPTCHA in task4

Type of CAPTCHAs: answering the calculation problems in the picture

The order of 4 tasks and the order of 10 CAPTCHAs in each task was random. Those tasks include different types of CAPTCHAs. Then different groups of participants answered those tasks, we got the response time and compute the pass rate through video record.

Questionnaire: After participants finish the 4 tasks, they got a questionnaire that asked them some questions, for example, which kind of CAPTCHA is more difficult. Which kind of CAPTCHA was passed after one try?

Administer post session questionnaire.

Safeguard: In the lab, it is a remote lab. We contacted participants by Zoom. It allows participants to be in an environment they are familiar with, such as at home. We also gave an example question before the participants took the tasks. Then explained any questions they had about the study.

3.3 Data

Collected data:

a.response time of each CAPTCHA in each task.

b.The grade of each task.

c.The answers or scores to the survey.

We stored the data in the MySQL database/Excel.

How can the collected data address research questions?

The first research question was analyzed from the data of time and grade that from different ages people reacted to the same type of CAPTCHA. Different answers in surveys helped analyze the impact of CAPTCHAs on users of different age groups and the factors that may delay completion.

The second research question was analyzed the data related to completion times and pass rates from different types of CAPTCHAs from similar age groups. Different answers in surveys helped analyze the impact of CAPTCHA on users of the same age groups and the factors that may lead to delayed completion.

The third research question is analyzed by the questionnaire which has specific feedback from each participant.

Through the integration and analysis of the data obtained from the responses to the above research questions, a CAPTCHA with stable performance (less response time, higher correctness rate, and fewer questions when participants finish the task.) is obtained from different angles.

3.4 Recruitment strategy

There were two ways participants were recruited.

1. Email

We recruited participants by sending emails from public mailing list. This email included the study's goal and what participants would do to help us complete this study. After scheduling participants, we invited them to join the task when they were available according to the information they had provided to us.

2. Invite participants from social media applications. (Wechat/Facebooks/Weibo) 3. All participants are selected randomly.

For the study, we recruited people of different ages. Since this study does not have special requirements for participants. We sent the invitation email to a mailing list or publish a questionnaire in Wechat moment (which helps invite people). During the conversation, we introduced what the study would be about (the study goal and the tasks), answering the questions they didn't understand. Then we scheduled a specific time to complete the tasks.

3.5 Treatment groups

There were two groups of participants, the younger group consisted of people aged between 20 and 35 years old, and the older group were people aged from 55 years old to 70 years old. We focused on the comparison between young adults and the middle-aged people. We believed with the age span, the contrast to response time would be more obvious.

3.6 Data analysis techniques

After the participants completed all four tasks and the survey, we organized the raw data from the recorded screen video. These data included the total response time and grade of each participant doing each task, then we drew charts that indicated the relationship between these two parameters. As for the questionnaire, the link is in Appendix.

Firstly, we sorted those data by using an Excel spreadsheet to classify by different ages and different types of CAPTCHAs. Secondly, we used Anderson-Darling to determine whether the distribution of the mean grade and the mean time was normal. Then we compared the mean response time of four tasks between our treatment groups and within each treatment group in order to answer research questions 1 and 2.

As we drew the charts of the relationship between ages and CAPTCHA types to give a guide to the designing of CAPTCHA and answer the research question 3.

During the whole process, we used charts to highlight the problems when participants finished different tasks to analyze our research questions.

3.7 Study procedures

1. Recruiting the participants by emailing or posting on social applications.
2. Let participants complete four tasks which have four different types of CAPTCHAs and one questionnaire. There would be a tutorial before participants took the tasks and questionnaire. The specific tutorial is in the Appendix.
3. Organizing and analyzing the raw data from participants.
4. Discussing the results and generating the conclusion.

4 ANALYSIS

In this study, we have two factors: CAPTCHA sets and age groups. This is a multi-factor experiment and the data was conformed of the

assumptions of ANOVA, we applied two-way ANOVA to determine whether there is a significant difference in mean response time and grades of different age groups finishing the different types of CAPTCHA sets.

According to Fig.5 and Fig.6 we can see both of the two factors have the significant influence on the time and grade.

Then, we used post-hoc ANOVA to see whether each factor has the significant influence on the results.

The means and grades of different age groups are shown in Figure 7 and Figure 8. In Figure 7, Grade is defined as how many right answers they made during each task. The green columns represent the mean grade (the number of correct responses) of the older age group in each task. And the purple columns represent the younger group. In Figure 8, Time means the response time to complete each task and was collected in seconds. The green columns represent the mean response time of the older group in each task while the purple columns are represented the younger groups. From the pictures, we can see the set 4 of tasks (calculation question) has less time spent and higher grades than other sets of tasks.

As it is shown in Figure 6, the p-value of CAPTCHA sets and age groups are less than 0.05. Thus, it indicates that the different ages and different sets show a significant effect on time. The p-value of set and group is less than 0.05, which means there is no interaction between these two factors. Then, in order to figure out which groups differ from each other, We used Tukey HSD ("honestly significant difference) to see the influence of each factor on the result. The post hoc ANOVA test result indicates that the difference between set 3 and set 1, set 4 and set 1, set 3 and set 2, set 4 and set 2 are statistically different, with a p-value less than 0.05. There is also a significant difference between the younger group and the older group. As is shown in Figure 6, the p-value of the factor set is 8.59e-13, and the p-value of the factor group is 0.00484, which is less than 0.05. We can determine that mean grades vary between CAPTCHA sets and age groups. The Tukey HSD indicates that since the p-value of set3 and set1, set3 and set2, set4 and set3 are less than 0.05, the grade of CAPTCHA set3 differs from the other sets. And the mean grade of the younger group is different from the older group because the Tukey HSD results determine that the p-value is less than 0.05.

Comparing the mean response time of the four tasks between our treatment groups is determining whether any statistically significant differences will indicate whether age is a factor in the efficiency of those CAPTCHAs. By comparing the average grades of each task, we can determine whether age is a factor that affects the grades of CAPTCHA verification. Thus the research question1 could be answered.

Comparing the mean response time of each task within each treatment group can determine whether there are any statistical differences between each task that indicates the CAPTCHA type is a factor in the efficiency of those CAPTCHAs, and which type of CAPTCHA enables participants to gain high correctness. This answers the research question2.

As we drew the charts of the relationship between ages and CAPTCHA type, we had a further discussion in Discussion section combined with the survey to indicate which type of CAPTCHA causes the low mental load to participants and to compare the usability of each type of CAPTCHA, which may give a guide to the designing of CAPTCHA and answer the research question3.

```
             Df  Sum Sq Mean Sq F value Pr(>F)
set           3 1081700  360567   14380 <2e-16 ***
group         1  455833  455833   18179 <2e-16 ***
set:group     3  166061   55354    2208 <2e-16 ***
Residuals   192    4814      25
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = time ~ set + group, data = alltime)

$set
                 diff        lwr        upr    p adj
set2-set1    9.564204   -5.77749   24.90590 0.372308
set3-set1  137.957622  122.61593  153.29932 0.000000
set4-set1  -64.928113  -80.26981  -49.58642 0.000000
set3-set2  128.393419  113.05172  143.73511 0.000000
set4-set2  -74.492317  -89.83401  -59.15062 0.000000
set4-set3 -202.885735 -218.22743 -187.54404 0.000000

$group
                  diff       lwr       upr p adj
younger-older -95.48125 -103.7376 -87.22488     0
```

Figure 5: time

```
            Df Sum Sq Mean Sq F value   Pr(>F)
set          3  163.3   54.42  25.802 4.55e-14 ***
group        1   19.2   19.22   9.113  0.00288 **
set:group    3   56.4   18.79   8.910 1.48e-05 ***
Residuals  192  405.0    2.11
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = accurate ~ set + group, data = allaccuracy)

$set
           diff        lwr        upr     p adj
set2-set1 -0.60 -1.3971584  0.1971584 0.2106599
set3-set1 -2.40 -3.1971584 -1.6028416 0.0000000
set4-set1 -0.56 -1.3571584  0.2371584 0.2668617
set3-set2 -1.80 -2.5971584 -1.0028416 0.0000001
set4-set2  0.04 -0.7571584  0.8371584 0.9992124
set4-set3  1.84  1.0428416  2.6371584 0.0000001

$group
              diff       lwr      upr     p adj
younger-older 0.62 0.1909972 1.049003 0.0048382
```

Figure 6: grade

5 RESULTS

According to the data we analyzed, there are results we summarized for the research questions.

RQ1: How does the same type of CAPTCHA perform for different ages of people?

From set 1 to set 4, the older people generally needed more time to answer tasks than the younger group.In set 3 (selecting picture the title required), the older people have a significantly lower correct rate and cost more time than younger people

Mean of grade (set1):

older group is 9.25, younger group is 9.5.

Mean of time (set1):

older group is 143s, younger group is 73s.

Mean of grade (set2):

older group is 8.75, younger group is 9.5.

Mean of time (set2):

older group is 338.2s, younger group is 155.2s.

Mean of grade (set3):

older group is 5, younger group is 7.75.

Mean of time (set3):

older group is 338.2s, younger group is 155.2s.

Mean of grade (set4):

older group is 9.75, younger group is 9.75.

Mean of time (set4):

older group is 55s, younger group is 30s.

Then, for the RQ2: How do users of the same age group perform on the different types of CAPTCHAs?

For both groups(older and younger), the set 4 (calculation problems) has the highest correctness and cost the least amount of time.

Set 3 (selecting picture the title required) cost the most time to finish.

At last, we answered the RQ3: Which type of CAPTCHA causes the low mental load to participants?

According to the questionnaire, 88.9% of people chose the lowest level for set 4 (calculation problems) when testing the comfort level of different tasks.

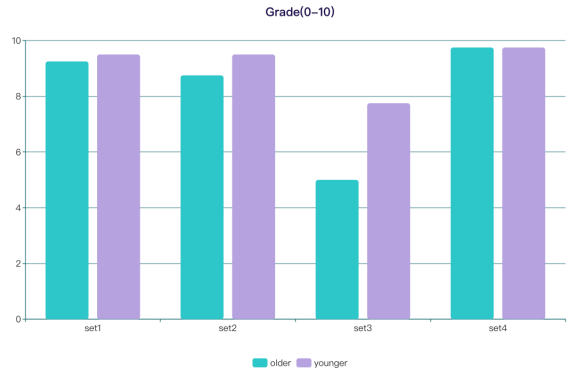

Figure 7: chart of mean grade

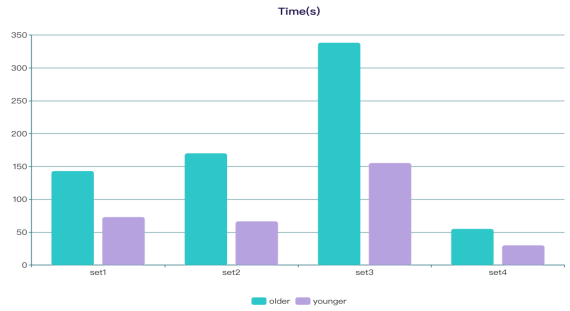

Figure 8: chart of mean time

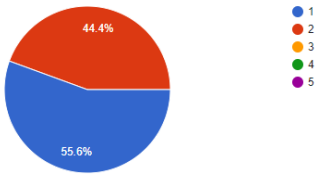

Figure 9: comfortable degree of set 1 (1 to 5 means increased discomfort)

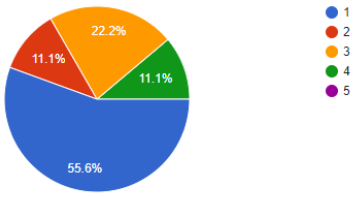

Figure 10: comfortable degree of set 2 (1 to 5 means increased discomfort)

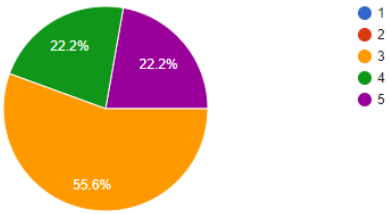

Figure 11: comfortable degree of set 3 (1 to 5 means increased discomfort)

6 THREATS AND VALIDITY

In this study, we had several limitations when implementing the laboratory study.

Internal Validity: Our study included within-subject tasks, participants may have gotten familiar with the format of the CAPTCHA

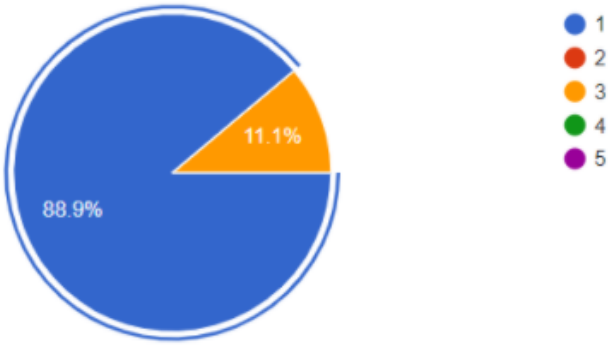

Figure 12: comfortable degree of set 4 (1 to 5 means increased discomfort)

test, which consumes less time. However, we randomized the order of the CAPTCHAs in each task and the order of four tasks to counterbalance the learning effect. The format of CAPTCHAs in each task was the same and participants suffered vision fatigue, especially in the older group. (In our study, participants were given the choice of a 2-minute break after they finished each task.) Furthermore, as we were observing participants' responses through cameras, they could be affected by the Hawthorne Effect [1]; for example, they could become nervous. Or, they are more concentrated since they have self-expectancy and want to gain high accuracy, which is not common in the real world. We are carefully conducting tutorials and don't speak much about time-consuming.

Construct Validity: When we did the questionnaire which answers the third research question, people could possibly lie because they may have forgotten some of their reactions. Because of this, we set some questions between each task.

Conclusion Validity: As we were experimenting, the refreshing time of the next CAPTCHA set was determined by the speed of the Internet, which could be wrongly added in the total time. So, we used a speed test application and required participants to set up their experiment environment with (enough network speed). Then, sometimes time-consumption may have had no impact in the different types of CAPTCHAS. And the better CAPTCHAS did not necessarily take the least time. In order to avoid some wrong conclusions determined by time data, we combined the factors we got from survey questions such as the recognition or comprehension load of different CAPTCHAs to people, when we were analyzing the results.

External Validity: Since our research included participants from the US and China, it had bias against Chinese since all the materials are in English. So, we designed the task and the survey using both English and Chinese.

7 DISCUSSION

According to the data which we collected, we summarized the more suitable for people of age this study covered is calculation question CAPTCHA.

Combing the result from the last section and our questionnaire, in this part, we will discuss the reason why they made mistakes and what may improve on the design of the CAPTCHA format in the future. When people do task1 (set1) and task2 (set2). Older people may have presbyopia or color weakness problems, so we may improve the contrast of the font color or embed voice verification into the webpage. In some cases, the letters may partly be overlapped due to the parameter setting of CAPTCHA generators. Such as m and n, once they are overlapped with other letters, it is hard to recognize whether it is m or n. Chinese people may not be familiar with the upper and lower cases. Letters such as"k, p, c". And the numbers 0 and the letter O have similar shapes and they lead to mistakes. Thus in the future design, designers may overlook the cases of letters.

When people do task 3 (set 3), the reason why they made mistakes are the target in a picture is too small and the picture has a lower resolution rate. So designers may improve the resolution of the picture and improve the recognition of target objects.

One of the reasons cause the difference between the younger group and the older group is their typing speed. Younger people may type and switch cases fast and be more familiar with using the keyboard.

When people do task4(set4), both the younger group and older group have a high grade and lower mental load. When deploying this type of CAPTCHA, designers may set low complexity of math problems when they are able to secure the website.

In this study, we found that CAPTCHA's type of selecting picture the title required spent a lot of time and the accuracy was lower than others. The reason why people always used a lot of time to pick pictures was, for example, the same color and composition confused people to choose the wrong answer. Those reasons can help AI's image recognition more accurately. At the same time, the choice from people can be a database to help AI's image recognition to learn more experiments.

Furthermore, in the study, we just compared two different age groups. It is not enough to figure out all people's reactions. In the future, we plan to add more age groups and collect more data to make our results have more details.

8 CONCLUSION

In this study, in order to find the CAPTCHA more suitable for the two age group people, we have three research questions about time and grades between those different types of CAPTCHA and ages.

According to the data we collected, from set 1 to set 4, the older people generally needed more time to answer tasks than the younger group.

For both groups(older and younger), set 4 (calculation problems) has the highest correctness and cost the least amount of time.

For set 3 selecting picture the title required cost the most time to finish. The older people have a significantly lower accuracy rate and cost more time than younger people.

In summation, we based our results on the statistical analysis and the answers to research questions. Firstly, set 4 has less response time to cost. Secondly, the correctness rate of set 4 between the two groups is the highest. Furthermore, the largest proportion of all tasks, thought set 4 didn't make them feel uncomfortable when

they are completing it. Thus, the type of set 4 is the most suitable between the four different formats in this study for different ages people.

REFERENCES

[1] Rob McCarney et al. "The Hawthorne Effect: a randomised, controlled trial". In: *BMC medical research methodology* 7.1 (2007), pp. 1–8.

9 APPENDIX

Tutorial

Introduction

Hi, my name is Jiawei, and this is Zehao. Thanks for being the participants and helping us to complete this study. We already talked about some content of our study when we invited you. But today we still have a simple introduction to the study. Our goal is to evaluate how the type of captcha affects different age groups. Today we will do some tasks and a survey. Don't worry, there are just some simple quizzes about captchas. We will be walking you through our study. Then we will mute off all of you now. If you have a question, please click the "hands up" button, then we will unmute you and answer your question.

"[Wait.]"

I'll be reading from this script to be consistent in the information I provide you and the other people taking part in this study. Please don't discuss this study with anyone, even in the chatbox, as we don't want other participants to receive any advance information. The goal of our research is to evaluate how the type of captcha affects different age groups. You'll be using Quizlet to finish our task. Let's get started with today's study. Let me know if you have any questions as we go along.

"[Wait for questions.]"

Firstly, log into your own account of Quizlet, if you don't have one, press register to create an account. This website is free and we don't need the primer membership. If you fail to log in or register, please give me a "hands up". If you log in successfully, please give me a "thumbs up". (this will be held in zoom). We will show you how to log into Quizlet.

"[Show the process of logging into Quizlet, Participants log in Quizlet. Wait until everyone logs in successfully.]"

Then, let's find out our captcha set. Everyone can see my screen? So the first step, we will input "cs567 user" into the input form. This is my account and we have created the captcha set already. Then you will see four sets: Captcha set1, Captcha set2, Captcha set3, and Captcha set4. As you can see, each captcha set will include one type of captcha: Captcha set1: selecting picture the title required. Captcha set2: input the irregular words and numbers in the picture. Captcha set3:answering the question in the picture. Captcha set4:Type of captchas: input the regular words in the picture. So, you just need to follow the requirements and input the answer you see. Just do what you will do when you see captchas in normal life. After you answer the questions, if your screen doesn't go to the next question, just click 'next' or if you have other problems, give us a "hands up". We will help you.

Now, you can click 'learn this collection', then you will save the learning collection to your account. We will show you how to complete those operations. Now I leave the time to you, please add four study sets to your account, if you have any questions, just 'hands up' to let us know.

On the left side of the webpage, there are messages with their subjects and dates listed.

"[sharing the screen to let participants know what need to do]"

Then, let's test a demo set. Click on this demo set, there are10 questions(captchas), which contain four different captchas. Now we will do an example. By the way, pay attention to some uppercase and lowercase letters. If you answer with an uppercase letter but the letter in the picture is a lowercase letter. The answer will be wrong.
"[After showing the whole process]"
Ok, now you will do those stuff by yourself, of course, the questions will be different from the example
End of Tutorial
After Main Task
"[Participants will finish see the total time of the task. Wait 10 minutes, or until everyone is finished.]"
Alright, we're going to move on. Now you could see the total time you spent on your task, we will collect this information also. One thing that we will let you know is please turn on the screen recorder when you are going to start the task, and end the recorder when you are done. After that, please upload the video to google drive, we will send you an email including the instructions you may need. Please click the "log out" button if you haven't already done so.
After finishing the whole four tasks, we will send you a survey, it will have some simple questions about the feeling and the problems you meet when you complete the tasks before. This survey is more about the experience when you finish different types of captchas. After you finish the survey, just click the 'finish' bottom. Then, we will get what you thought after those tasks.
The screenshots of the example we will show to the participants
Steps
Search the CAPTCHA set:

Figure 13

Choose set (tutorial set)

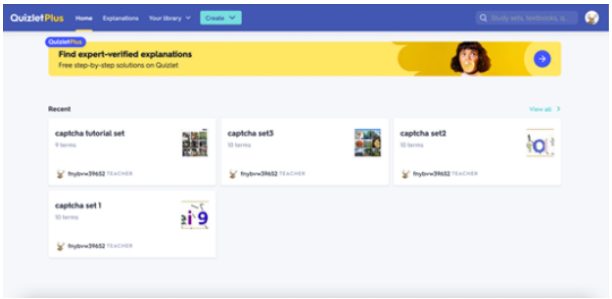

Figure 14

Press Write" button

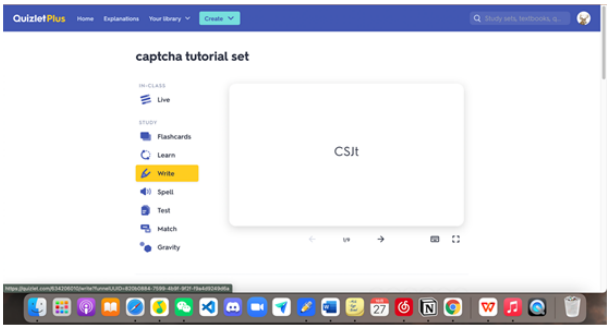

Figure 15

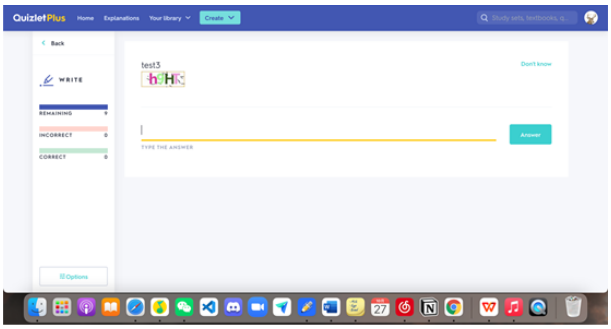

Figure 16

Input the answer, then press the "answer" button
When answering the image-based CAPTCHA, input the number of pictures in order. Then press the "answer" button. You can click the image to see it.

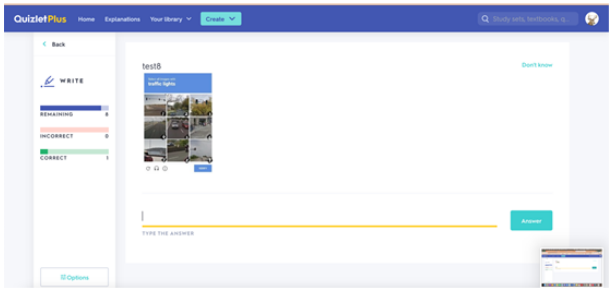

Figure 17

If you input the wrong answer, don't worry J, press any key to continue, then you will reinput the answer during the next round!

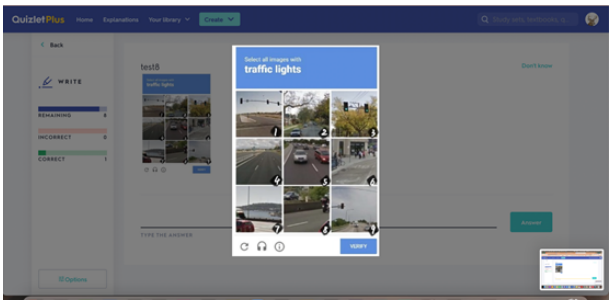

Figure 18

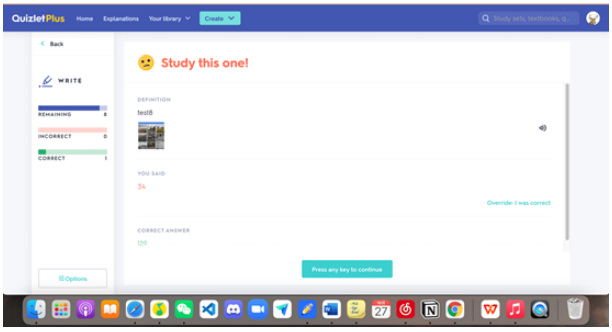

Figure 19

The left side menu is an overview of your progress.

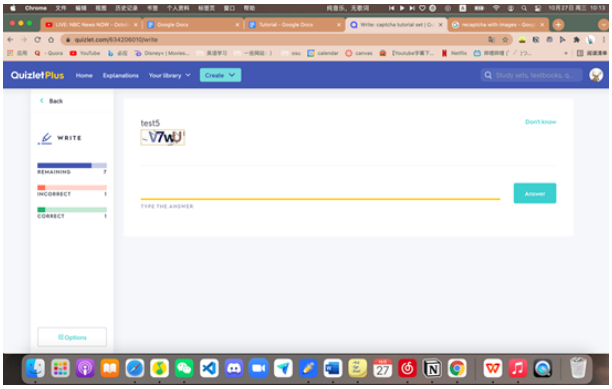

Figure 20

In the end, you will receive a report and then you're done! Good Job!
End of Session

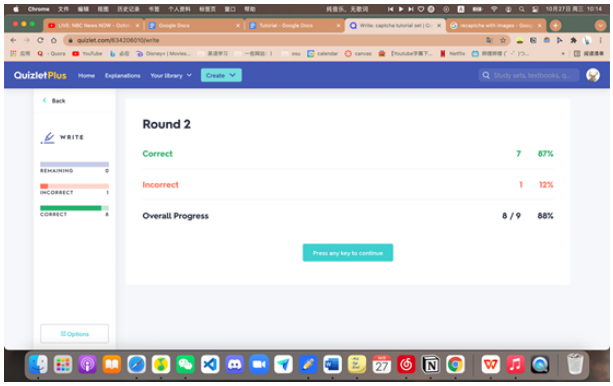

Figure 21

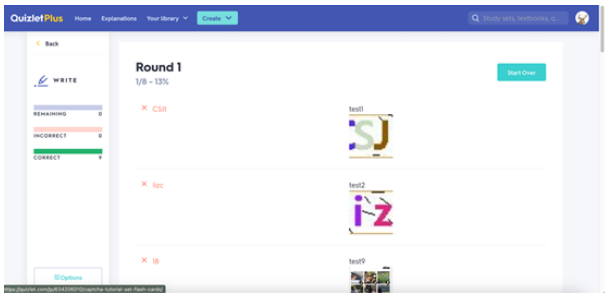

Figure 22

Alright, that completes the experiment. Thank you so much for all of your help. Please stay seated and we will be coming round to give you your incentives.

Questionnaire: https://docs.google.com/forms/d/e/1FAIpQLScxG_7rOTA2iXvtg0ZxgwPvfOoECLblOARKvymMvwzwUbSrbA/viewform?usp=sf_link