Comparison of Memory Recall Performance for User Interface: Images vs. Words

Chloe Dimpas, Cedric Ngaheu Eken, Anupreet Kaur Gill, Jonathan Singh
Dept. of Electrical Engineering and Computer Science
York University
Toronto, Ontario, Canada M3J 1P3
{chloedim, cedric04, anu1512, jsingh22}@my.yorku.ca

ABSTRACT

The study evaluates the memory recall of users when interacting with image buttons as compared to text buttons. Participants played a text and picture matching game with four and seven pairs to match. Their performance results were collected and evaluated. It was found that participants took a shorter time and had better accuracy when matching images as opposed to matching texts. The mean image matching time was 13% faster than the mean text matching time for seven pair matching. Despite a negligible difference for four pair matching, the mean accuracy of text matching was 11% lower than that of image matching for seven pair matching.

Keywords

User Interface Design, Graphical User Interface, Spatial Memory.

INTRODUCTION

Since the genesis of mobile computing, the goal of software providers and tech companies to provide an enriching user experience to users through appealing mobile user interfaces has become one of major importance. The quality of the user interface (UI) design of software can determine its usability and how much profits or losses the company will make from the software. This all depends on how much effort, be it mental or physical, is required by the user to effectively and efficiently navigate through the UI and use the software.

Since the inception of its concept, mobile user interfaces have evolved and improved through the creation of various design techniques and principles that cater to the user's satisfaction. That said, there is still room for improvement. Taking into consideration that many important technical contributions given to society were accomplished after a series of evaluations, and trial and error procedures (including official UI design best practice principles), this user study will take the same path to add a brick to this edifice of accomplishments. In this case, the goal of this user study is to assess whether the use of visual imagery in UI design is better than written words for good and durable spatial memory recall by the user. Spatial memory is the ability to learn and remember the location of objects by interacting with them repeatedly. As Scarr et al. [5] noted, good spatial memory of the location of controls can promote fluid and efficient user interaction with the interface, hence it is an important aspect to be considered to enhance UI design.

The goal of this paper was to ask several users to play a game containing cards with pictures and cards with words. The gameplay data of each user was then collected and stored and a comparative evaluation was performed on it to derive insights and answers to the paper's original goal. The main idea was to determine whether the users had the best spatial memory recall with the cards containing images or the cards with written words. The findings of this experiment can be applied when making various UI design decisions. One example would be deciding whether buttons should be labeled with words or if they should be presented as an image or icon as shown in Figure 1 below. Software such as Microsoft *Word,* utilizes icons to distinguish the function of buttons, as shown in Figure 2. This experiment aims to provide insight as to why certain design decisions are made.

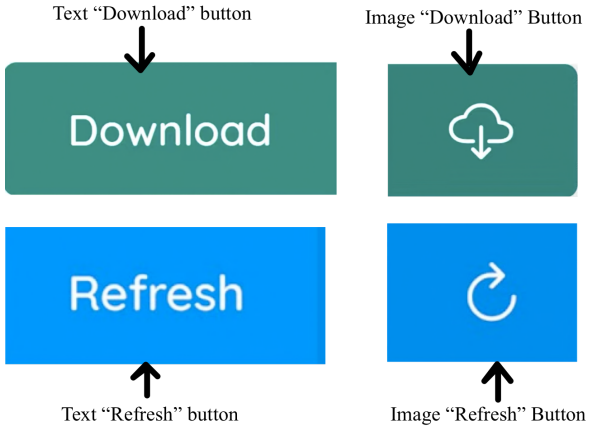

Figure 1. Screenshot of Image and Text buttons[1]

[1] https://images.app.goo.gl/taw2ggMMpuJtrJb49

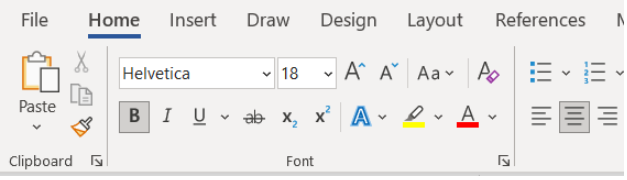

Figure 2. Screenshot of Microsoft *Word* ribbon[2]

RELATED WORK

Corsini et al. [2] examined preschool children's memory for words and pictures using a recognition-memory task. Children were asked to indicate whether an item has or has not been previously presented. Two recognition memory lists were created, one containing pictures (P) and one containing words (W). The two lists, W and P, were identical in content. A total of 40 preschool children were examined. The results indicated that pictures were better retained than words.

Borges et al. [1] compared recall and recognition performance under three different modes of presentation for 30 subjects in each of three grade levels. The modes of presentation were written words, black-and-white pictures of objects, and color pictures of objects and grade levels were college students, sixth graders, and fourth graders. The results showed that recall performance increased with age. Adult performance was significantly affected by mode of presentation; however, no difference was found across presentation modes for the fourth and sixth graders. Adult performance was best with color pictures. Black-and-white pictures yielded second best performance and the performance was worst with written words.

Powell et al. [4] investigated the organization and recall of pictures and words in children from the third, fourth, and sixth grades. It was concluded that both recall, and organization were greater among older children and for pictures than words. Powell et al. suggested that "pictures are easier to recall because the transformation of surface structure representations into organized deep structures is less influenced by the context of encoding than it is in the case of words" [4, p. 1].

Gehring et al. [3] did two experiments to analyze the recognition memory for words and pictures at short and long retention intervals . In both experiments, pictures were easier to recall than words at all retention intervals tested, and this superiority was constant over time. Gehring et al. [3] argued that recall memory is superior for pictures than for words up to retention intervals of about two weeks.

METHOD

This study compared how users retain information through pictures versus through text. The study itself was conducted under an observational lens, which simply means that the researchers did not influence the experiment in any way; they were just observing the participants as they progressed through the levels of the experiment. Each participant in the study completed four blocks of runs in the same trials, and the results of each of them were compared at the end of the study.

Participants

There were eight voluntary participants who went through the experiment, four male and four female, all ranging from ages 18 - 40. None of the participants were expected to have played memory games regularly, so as not to give them an advantage above the others. All of the participants regularly use smartphones or tablets. None of the participants were visually impaired. There was no compensation for participation in the experiment: monetary or otherwise.

Apparatus

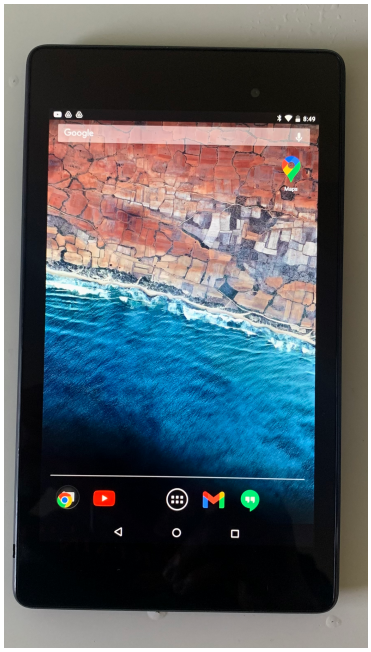

Figure 3. Samsung *Nexus 7*.

This experiment was conducted on a single device: a Samsung *Nexus 7* tablet as seen in Figure 3. The device was seven inches and had a resolution of 1920 x 1200 pixels, so the software layout was built with these specifications in mind.

As seen in Figure 4 below, the software that was used for the experiment was a "memory match" game, and was

[2] Picture retrieved from Microsoft Office Word Tools Board - Home Tab.

developed in Android Studio on the Microsoft *Windows 10* operating system.

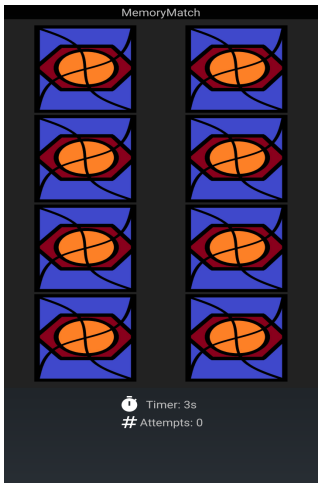

Figure 4. Layout of the card matching application.

Figure 5. Example of two face-up matching cards.

The software was used to see if participants can better recall images or written text. As far as the game itself goes, the cards were all face-down and randomly ordered at the beginning of each round. The player needed to match each of the pairs in as little time and with as few attempts as possible. An attempt was counted when the player flipped two cards over in a pair, and the count occurred regardless of whether or not the pair matched. Once a player tapped to reveal the first face-down card, they had to pick another face-down card. If the newly revealed card matched the previous one, both cards remained revealed and one pair would have been found, an example of which is shown in Figure 5. If the newly revealed card did not match the previous one, both cards would be flipped back over and the process would continue until all cards were revealed.

Procedure

The participants were informed of the purpose of the experiment. Participants were shown a short demo of how the card matching game works. There were no practice trials. Participants were seated and asked to hold the tablet in the hand that they are most comfortable with, as seen in Figure 6. Participants were asked to play four rounds of the card matching game. At the end of all four blocks, the participants were asked to complete a questionnaire. The purpose of the questionnaire was to gather information about the demographics of the participants and to collect feedback about the experiment. The questionnaire asked the participants to answer questions about themselves, (i.e., “What is your age and sex?”, “How many hours a day do you use a touch screen device?”, “Do you play memory games regularly?”) as well as their experience with the task. Their performance was logged in the tablet’s memory and later transferred to a computer for analysis.

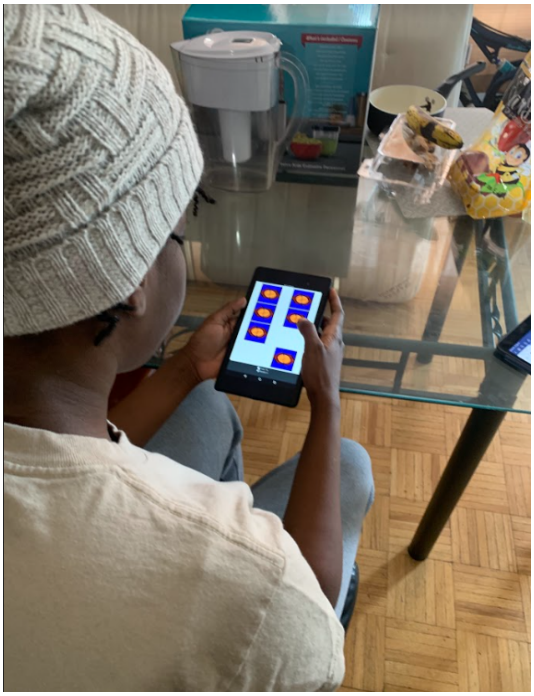

Figure 6. A participant performing the card matching task with four pairs of images

Design

The study employed a 2 × 2 within-subjects design. The independent variables and levels are as follows:

- Number of pairs to match (four, seven)
- Type of matching (words, pictures)

The total number of trials was 8 participants × 2 pairs to match × 2 types of matching = 32 trials.

To offset learning effects, half of the participants matched words in their first block, while the other half started with matching pictures. The type of matching alternated

between blocks to offset order effects by counterbalancing. Participants who belonged to group A did word-matching for the first and third blocks, and picture-matching for the second and fourth blocks. Participants who belonged to group B did picture matching for the first and third blocks and word-matching for the second and fourth blocks. For all participants, the first two blocks had four pairs of cards to match, while the last two blocks had seven pairs to match.

The dependent variables were speed and accuracy. The speed was the completion time per trial, measured in seconds. The number of attempts it took to match all the pairs of cards was used to measure accuracy. The accuracy was inversely proportional to the number of attempts. The more attempts, the lower the accuracy.

RESULTS AND DISCUSSION

Speed Analysis

The "matching time" was measured as the time between the selection of the second card and the selection of the final pair. It was measured in seconds. The grand mean of the matching time was 49.00 seconds. As seen in Figure 7, the mean matching time for seven pairs is 1.65 times greater than the mean matching time for four pairs. The mean matching time for four pairs was 37.06 while the mean matching time for seven pairs was 61.13.

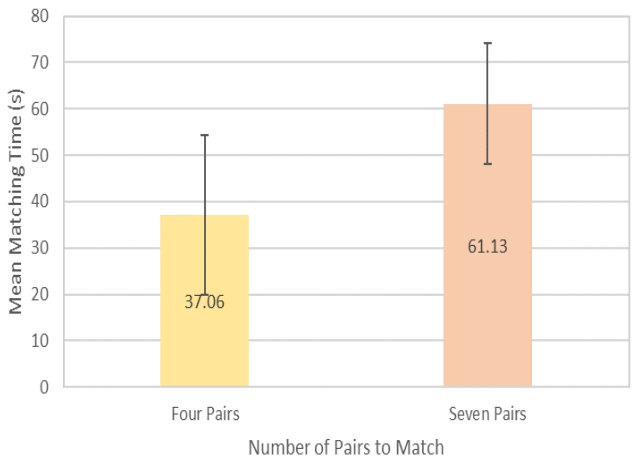

Figure 7. The Mean Matching Time based on the number of pairs to match

The data is further broken down in Figure 8. as it shows the mean matching time for four pairs, based on the type of matching. The mean image matching time was 39.13 seconds while the mean text matching time was 35 seconds. It should be noted that one of the participants had an abnormally high image matching time of 102 seconds. This affected the mean of the image matching time. If the data from this outlier is removed, the mean text matching time becomes greater than the mean image matching time. The mean image matching time is 30.14 seconds and the mean text matching time is 32.57 seconds.

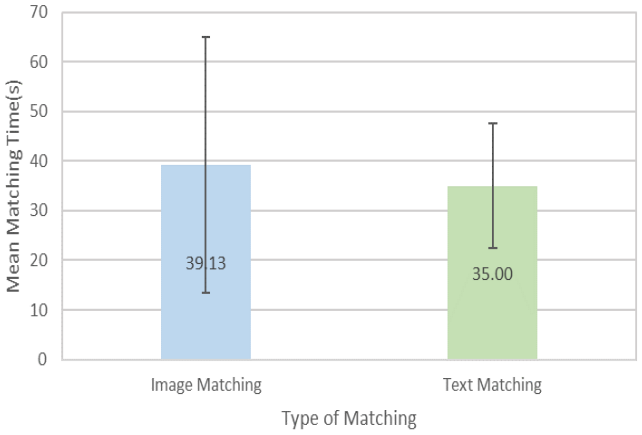

Figure 8. Four-Pair Matching Time

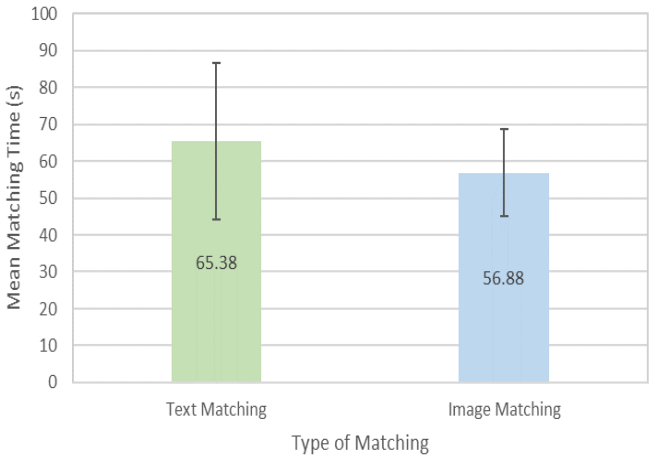

Figure 9. Seven-Pair Matching Time

For seven-pair matching, the mean time was 65.38 seconds for text matching and 56.88 seconds for image matching. See Figure 9.

While the difference in mean matching time was small for four-pair matching, for seven-pair matching the mean time for matching images was 13% faster than seven-pair text matching.

From the data and the graphs above, the observed trend is that participants take shorter time to match images than words.

Accuracy Analysis

In this experiment, accuracy was measured as the number of attempts that the user made in order to find all the matching pairs. The fewer the attempts, the greater the accuracy.

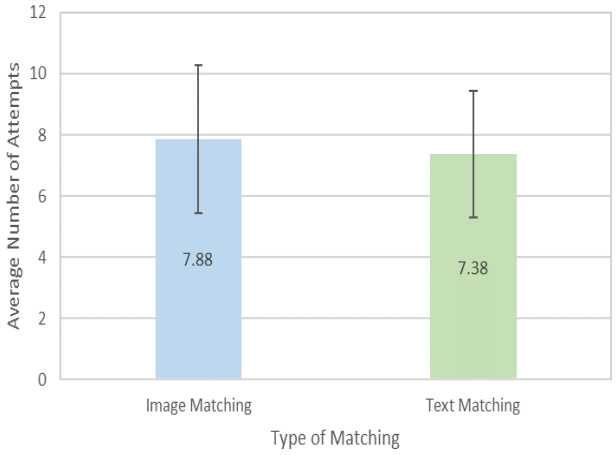

Figure 10. Four-Pair Matching Accuracy

For four-pair matching, the average number of attempts did not differ by much depending on the type of matching. As seen in Figure 10, both types of matching had an average of 7 to 8 attempts.

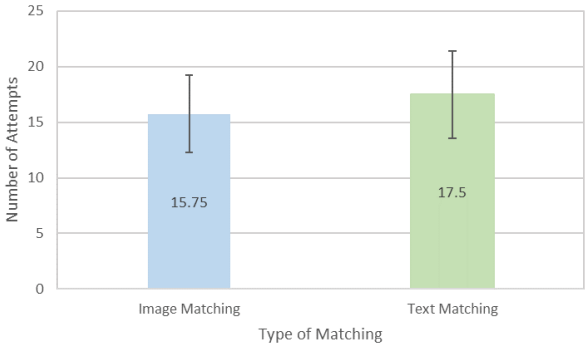

Figure 11: Seven-Pair Matching Accuracy

For seven-pair matching, the average number of attempts differed by about 11%. As seen in Figure 11, image matching had an average between 15 and 16 attempts while text matching had an average between 17 and 18 attempts.

From the data collected from the experiment, it can possibly be concluded that text matching requires more attempts than image matching.

Participant Feedback

Based on the questionnaire given at the end of the experiment, participants expressed positive feedback regarding text matching and picture matching games. Most of the participants indicated a strong preference for the picture matching game as opposed to the text matching game. When asked for a rating on a Likert scale from 1 (least likely) to 10 (most likely) on how likely they prefer the picture matching game more than the text matching game, the average rating was 7.9. Most participants did not experience eye fatigue and frustration while doing the experiment. Overall, participants enjoyed both text matching and picture matching games when asked about their experience.

CONCLUSION

Overall, the results showed that the image matching task required less time than the text matching task. For four-pair matching, image matching took 7.46% less time than text matching. For seven-pair matching, image matching took less time as well.

The accuracy did not have much difference between the two types of matching when only four pairs of cards were matched. However, for matching seven pairs of cards, image matching took fewer attempts than text matching.

The results from this experiment can provide insight on how to make UI design decisions when choosing between text and picture representation. The improved speed and accuracy with image matching may indicate that users have better memory retention when it comes to images as opposed to text. This is definitely something that UI designers should keep in mind when designing applications.

REFERENCES

1. Borges, M. A., Stepnowsky, M. A., and Holt, L. H.,Recall and recognition of words and pictures by adults and children, *Bulletin of the Psychonomic Society*, *9*, 1977, 113-114.
2. Corsini, D. A., Jacobus, K. A., and Leonard, S. D., Recognition memory of preschool children for pictures and words, *Psychonomic Science*, *16*, 1969, 192-193.
3. Gehring, R. E., Toglia, M. P., and Kimble, G. A., Recognition memory for words and pictures at short and long retention intervals, *Memory & Cognition*, *4*, 1976, 256-260.
4. Powell, A., Havnaer, J., and Wiggins, W., Organization and recall of pictures and words in children, *Psychonomic Science*, *29*, 1972, 385-387.
5. Scarr, J., Cockburn, A., and Gutwin, C., Supporting and exploiting spatial memory in user interfaces. *Foundations and Trends in Human-Computer Interaction*, *6*, 2013, 1-84.