COMP 3008 Project 2

Quantitative Usability Evaluation

Winter 2017

Dr. Robert Biddle

Due: April 7th, 2017

Submitted by: A1\_Team 01

Chantal Forget - 100886390

Shannon Stone - 100882544

Tineke Vanderhorst - 100879946

[**Part 1**](#_ygs2svx9mfsy) **2**

[Advantages and disadvantages of password schemes](#_lnv67znl9w80) 2

[Text28](#_9nlxrra8g2c0) 2

[Screen Images](#_h4p4zm91zd4) 3

[Blankpt28](#_pbdtfo7rxb4) 4

[Screen Images](#_7dpqu7h0szd7) 4

[Imagept28](#_k9a7rhn84y7d) 7

[Screen Images](#_bk4n96txp4sv) 8

[Log Processing Script](#_tyb96ec3ce3l) 10

[Descriptive Statistics](#_gxap8zlf73a5) 11

[Graphs](#_68qbec7uzir5) 12

[Discussion and Interpretation](#_h2qtf5fqn79b) 16

[**Part 2**](#_1n65nkksf9si) **17**

[Design Rationale for Password Scheme](#_vb2tbhlzxnyz) 17

[Password Creation](#_3wyuwv2p3xgv) 20

[Screen Images of the System in Use](#_rehvsp2nlx8q) 20

[Screen Images of the Password System](#_od9beuc3n26u) 22

[Training phase](#_c912g69yftpx) 22

[Testing phase](#_u6nxwa7x8rp3) 24

[LimeSurvey](#_u431kk95hc61) 25

[Usability Testing Data Analysis Compared to Text28](#_tmdnqlft490s) 25

[Descriptive Statistics](#_gkh73y9ae1x7) 25

[Graphs](#_svr8urojpe0h) 27

[Inferential Statistics](#_6yrbckcwshhr) 30

[Since our data is not normal we must run a Wilcoxon signed-rank test in order to test for statistical significance.](#_7wxpbb5iqoda) 30

[Discussion and Interpretation](#_65yxoan0i78y) 30

[Usability Testing Data Analysis Compared to User-Chosen Text Password Scheme](#_h6javg7voznu) 31

[Descriptive Statistics](#_hc7baet5o34g) 32

[Graphs](#_kaxcw333296e) 33

[Discussion and Interpretation](#_f1ikfyj85lde) 33

[Inferential Statistics](#_jmh5o24vh5ug) 33

[Interpretation](#_fdmcpsaz9z9d) 33

# Part 1

## Advantages and disadvantages of password schemes

### Text28

Advantages:

* Text can be quickly entered
* Text is random which makes it less likely to be victim of dictionary attacks. This is due to hackers expecting people to use passwords with meaning.
* May be possible to adapt for people with vision problems.

Disadvantages:

* This text is random and thus has no meaning to the user. This makes it harder for users to remember. As such, users are more likely write it down.
* When trying to memorize the password, we ended up saying it aloud to help remember the order. Such a practice means that anyone within hearing distance could write down the password.
* The password scheme only included lowercase letters, no capitals or special characters. If hackers think to use a random arrangement in a dictionary attack this would probably fail the easiest.

### Screen Images

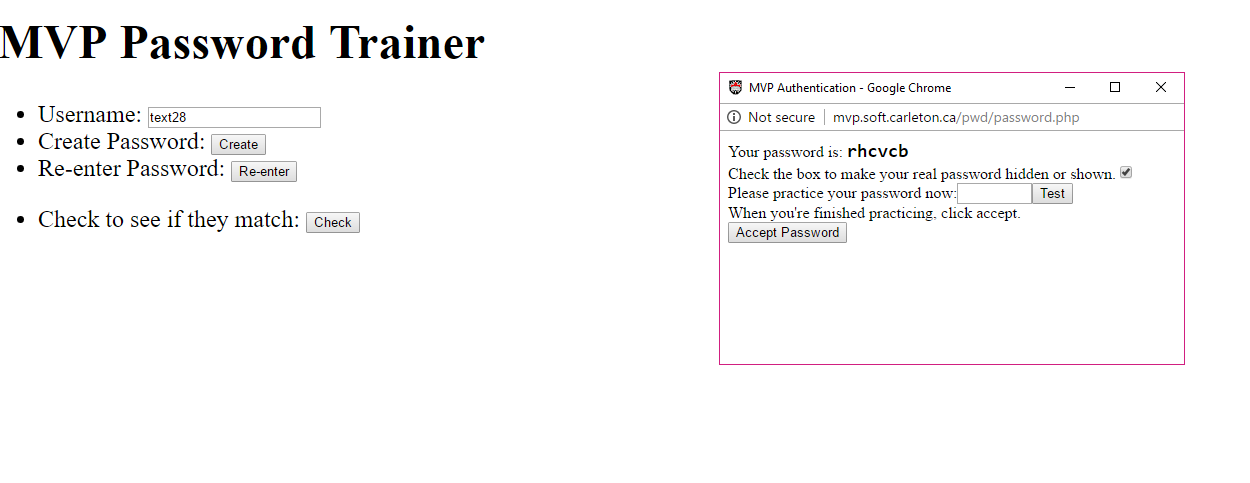


Figure 1. Practicing entering random text password

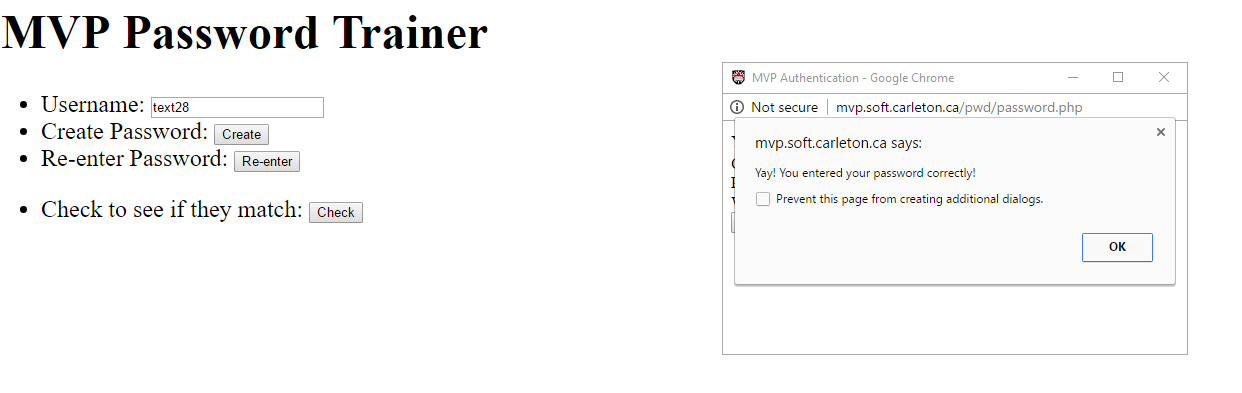


Figure 2. Correctly entering the password post-training

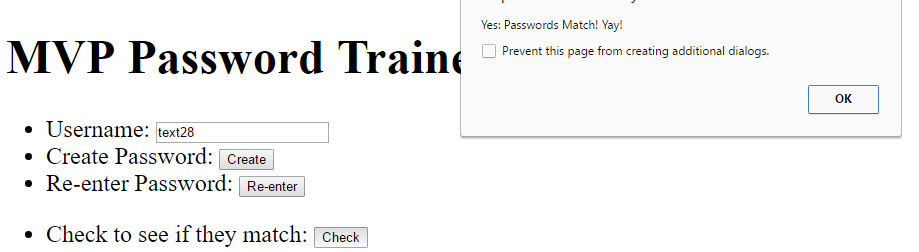


Figure 3. Random text passwords match.

### Blankpt28

Advantages

* The tiles do not have to be entered in a specified order. As they can be entered differently each time it protects against over the shoulder attacks.

Disadvantages

* Very hard to memorize because you can’t see the pattern when you’re entering the password. Unlike the password training section in which the tiles you click have borders around them. You could remember this better if you could see the pattern when entering it, but as such it's fairly impossible to remember unless you remember the exact square coordinates.
* Not accessible for people with vision problems.

### Screen Images

Figure 4. Incorrectly entering a password during the post-training phase.

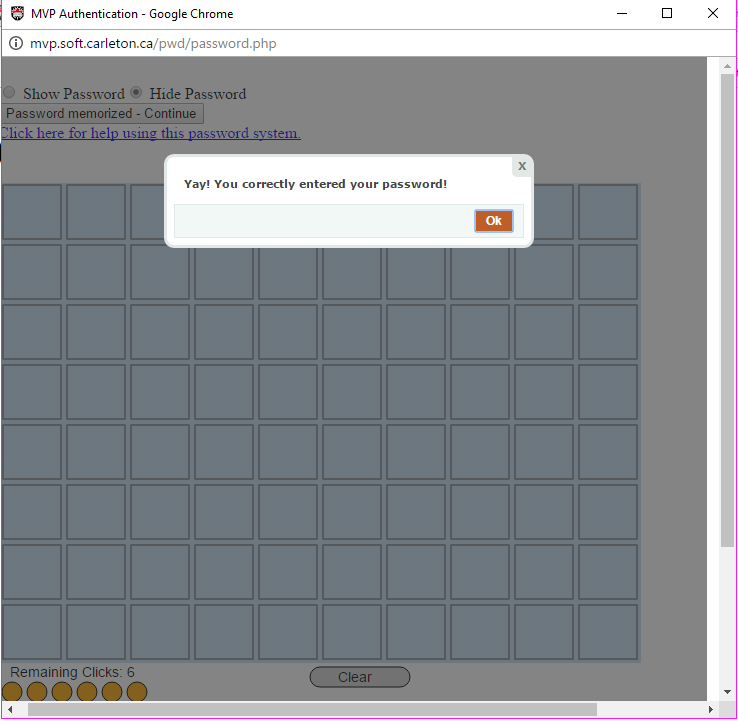


Figure 5. Correctly entering a password in the post-training phase.

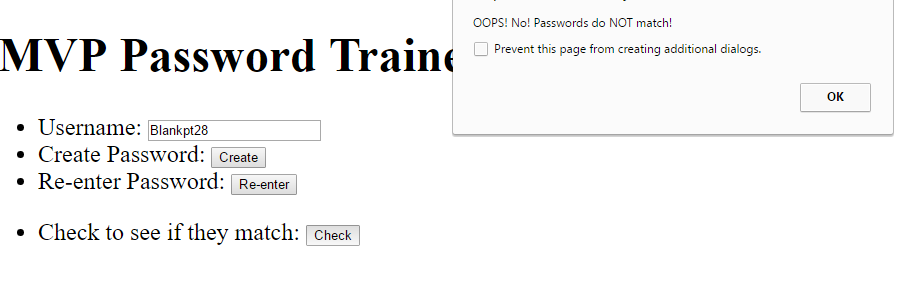


Figure 6. Passwords do not match after entering the passwords twice.

### Imagept28

Advantages

* Is more visual as the background is an image. This makes it easier to remember than Blankpt28
* No specified order, can enter differently each time - protects against over the shoulder attacks.

Disadvantages

* We noticed the same disadvantages as in Blankpt28
* Easier for over the shoulder attacks than a text based one.

### Screen Images



Figure 7. Training for Imagept28

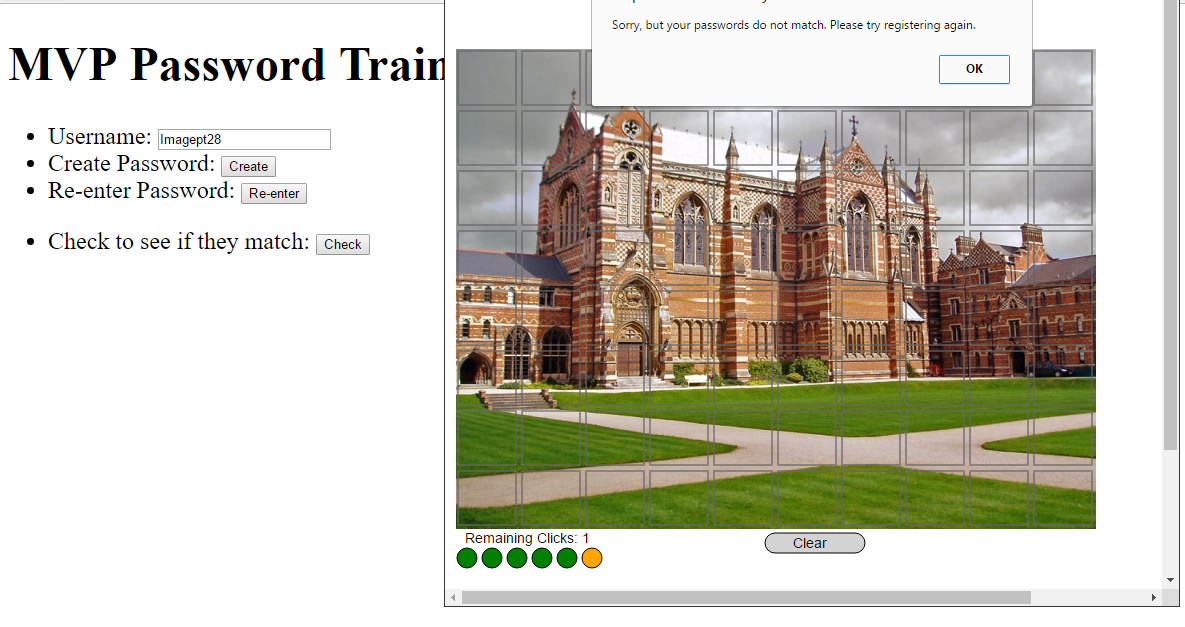


Figure 8. Incorrectly entering the password post training.



Figure 9. Correctly entering the password post training.

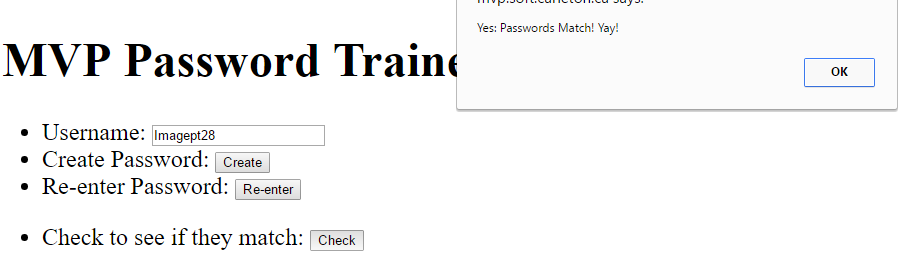


Figure 10. Passwords match

## Log Processing Script

We took a two-step approach to processing the original log data into data that can be used for statistical analysis. The first step was to organize the data into a csv file with a single row per login attempt. Then this data is run through again in order to create the final data for the analysis. Please see the files in the DataProcessing folder inside A2-T01-sourceFiles for full code, documentation, and resulting csv files.

The first step results in the logins.csv file. To create this file, the script iterates over every line in the original data. If a line is a start event, then it is added to a dictionary organized by user ids. When a line with login mode is encountered, the dictionary is searched to find the start event for the same user. If a start event is found, then the difference between the two times is calculated. A line containing the user, the event type (success or failure), the scheme, the site, and the time difference is added to the data to be printed to logins.csv. Some logins were ignored because there was not a corresponding start event for that user. This was a small percentage of the data and was likely caused by duplication in the logging of some steps.

For the second stage, the logins.csv file is read in and a file called userData.csv is created with a line for each user. First, a dictionary is created from the data where user ids serve as the keys and the values are a list of the logins by that user. This dictionary is then iterated over and the total logins, successes, failures, mean success time, and mean failure time is calculated for each user. Line with this information as while as the user id and the scheme that was used is added to the data to be output to userData.csv.

## Descriptive Statistics

Total Logins

|  |  |  |  |
| --- | --- | --- | --- |
|  | Mean | Median | Standard Deviation |
| Text28 | 16.778 | 16 | 4.177 |
| Blankpt28 | 16.556 | 17 | 3.745 |
| Imagept28 | 18 | 17.5 | 3.284 |

Table 1. Data for total logins.

Successful Logins

|  |  |  |  |
| --- | --- | --- | --- |
|  | Mean | Median | Standard Deviation |
| Text28 | 14.667 | 15 | 1 |
| Blankpt28 | 13.778 | 15 | 2.279 |
| Imagept28 | 15.25 | 15 | 1.780 |

Table 2. Data for successful logins

Unsuccessful Logins

|  |  |  |  |
| --- | --- | --- | --- |
|  | Mean | Median | Standard Deviation |
| Text28 | 2.111 | 1 | 3.855 |
| Blankpt28 | 2.778 | 2 | 2.635 |
| Imagept28 | 2.75 | 2 | 2.890 |

Table 3. Data for failed logins

Successful Login Times

|  |  |  |  |
| --- | --- | --- | --- |
|  | Mean | Median | Standard Deviation |
| Text28 | 11.714 | 6 | 9.987 |
| Blankpt28 | 26.145 | 27.417 | 10.721 |
| Imagept28 | 30.940 | 29 | 10.201 |

Table 4. Data for successful login times

Unsuccessful Login Times

|  |  |  |  |
| --- | --- | --- | --- |
|  | Mean | Median | Standard Deviation |
| Text28 | 31.370 | 3 | 87.271 |
| Blankpt28 | 25.769 | 25.285 | 20.586 |
| Imagept28 | 37.553 | 33.429 | 41.438 |

Table 5. Data for unsuccessful login times

### Graphs

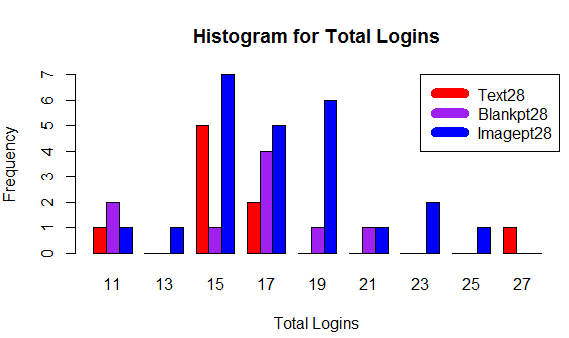


Figure 11. Histogram of the total logins per user for each password scheme.

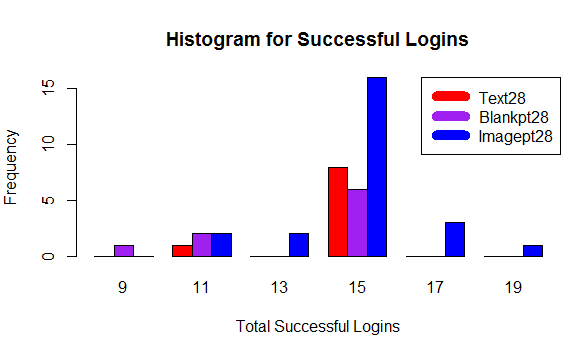


Figure 12. Histogram of the successful logins per user for each password scheme.

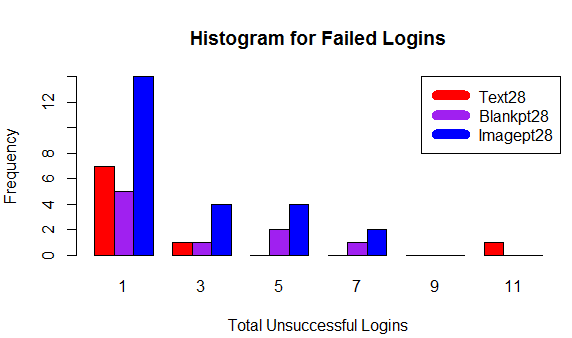


Figure 13. Histogram for the failed logins per user for each password scheme.

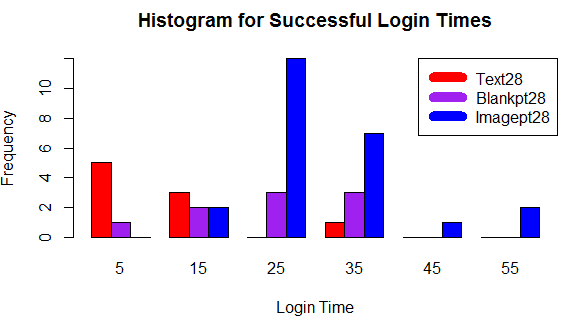


Figure 14. Amount of time it took users to successfully login for each password scheme.

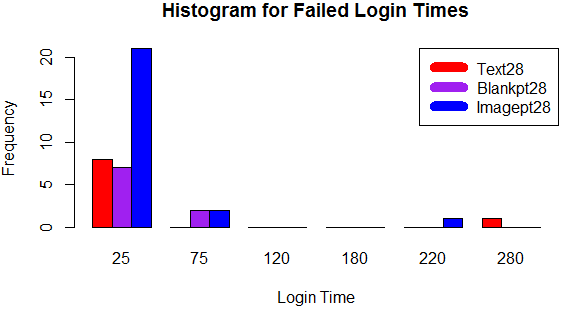


Figure 15. Amount of time it took users to unsuccessfully login for each password scheme.

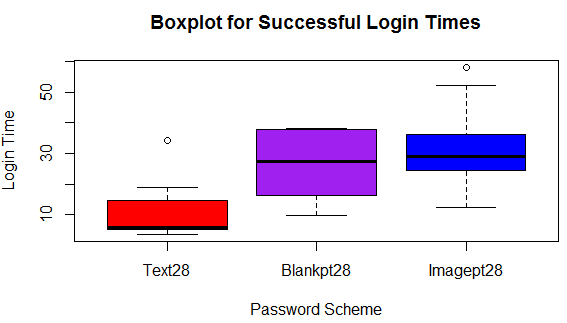


Figure 16. Comparison of the amount of time it took users to successfully login for each password scheme.

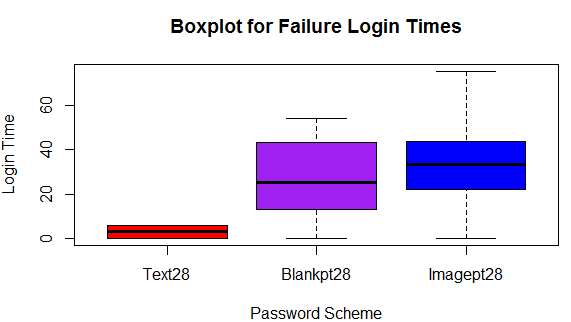


Figure 17. Comparison of the amount of time it took users to unsuccessfully login for each password scheme.

### Discussion and Interpretation

Data is fairly similar across Total Logins, Unsuccessful Logins and Successful Logins (Table 1, Table 2, & Table 3). Therefore our results should not be determined by the number of participants across the password schemes.

The standard deviation for the unsuccessful login times for the Text28 password scheme is very large due to the presence of an outlier (Table 5). Due to the presence of this outlier the mean is skewed and may not be reliable, making the median a better estimate of the central tendency in this case. For all three password schemes the medians for successful and unsuccessful logins appear to be similar (Table 4 & 5).

**Histograms and Boxplots**

Figure 11 shows us that the Imagept28 password scheme has the most variance and number of total logins out of all three password schemes while on a logins per user basis, the Text28 password scheme was used the least. When examining the distribution of successful logins we can see that for all three schemes most users had 15 successful logins (Figure 12), this is consistent with the results of the medians (Table 2).

The histogram for failed logins shows that the Imagept28 password scheme has the largest number of users with a failed login (Figure 13). There also appears to be a huge variance for failed logins for users using the Text28 password scheme. This indicates that users either knew their password very well or did not with no in-between (Figure 13).

When examining the times for a successful login we can see that the Text28 password scheme has the fastest login times while the Imagept28 password scheme has the slowest (Figure 14 &16). In addition the Imagept28 scheme has the most even distribution while the Text28 data is highly skewed, as the median and second quartile overlap, and the Blankpt28 password scheme has less variation above the median compared to below (Figure 16). Both the Text28 and Imagept28 schemes contain an outlier above the upper whisker (Figure 16).

The distribution of the times for the failed logins tells us that the majority of users for all three password schemes took 25 seconds to enter their passwords incorrectly (Figure 15). In addition, when looking at the histogram we can assume that the 280 second and 220 second times are outliers where the user most likely abandoned their attempt (Figure 15), these have been left out of Figure 17 in order to better see the characteristics of the boxplots. When looking at the boxplots we can see that the interquartile range for the Text28 password scheme is very small, showing little variation in the data (Figure 17). Once again the Imagept28 scheme has the most even distribution of the three password schemes with the data appearing almost symmetrical (Figure 17). Finally the data for the Blankpt28 scheme is slightly skewed with more data above the median (Figure 17).

Of the three password schemes that we were provided data for, we believe that the Text28 scheme has the best usability. We came to this conclusion based on its fast login times and the smaller percentage of failed logins.

Please see the file a2\_Part1\_Q3.R in the RFiles folder inside A2-T01-sourceFiles for full code and documentation.

# Part 2

## Design Rationale for Password Scheme

For our password scheme, we decided to have 3 different components comprised of 3 landscape images, 1 randomly assigned word, and 3 emojis. We did so as to use the chunking principle. Just as phone numbers use chunking to separate the numbers into three different components, we are chunking into three different components to make the password scheme more secure. As well, we hope that using the chunking principle will allow users to remember each section easier. For both the landscape and emoji portion, we will not be shuffling the images each time a user arrives at the page. This is due to the fact that it dramatically increases the time taken to enter a password as users must scan all possible images.

The three sections of our login include 3 pictures of landscapes to choose from, a randomly assigned word of 6 letters, as well as 3 emojis to choose. We chose to put them in this order because we felt it has the highest memorability. The primary and recency effect should give the landscape image and emoji image components higher memorability. The text component was chosen to be in the middle because we felt that since we were using words instead of randomly assigned text it would naturally have a higher memorability that could counteract the serial position effect. We also wanted the emojis to go after the text because most people will put emojis at the end of a sentence.

We feel that by combining the text password scheme and using a variation of the image password scheme could leverage the strengths of both schemes. In order to keep the entry time to a minimum we chose to scale down the image scheme by having the user select three images instead of six.

The three different components of our password scheme are described in further detail below.

**Landscape**

We used landscape images as these images can differ quite spectacularly. We used photos that are distinct enough that users will hopefully remember them quickly when they return to the page. The landscape images were retrieved from <https://pixabay.com/en> by searching for “landscape”. These images are available for free non-commercial use under the creative commons license.

**Randomly assigned word**

We wanted the user to enter a randomly assigned word of a certain length. We wanted to incorporate this as words are easy for users to remember. As we noticed during the random text password entry in Part 1, it was difficult to remember a randomly generated text, as it had no meaning to the user and was simply a random assortment of words. Though we are still assigning the user a word, the word is an English one that will be easier for users to remember. Words that are six letter in length were chosen because they offer some complexity to increase the difficulty of dictionary attack but there are still a large number of commonly known words of this length which most users would have seen in their everyday lives.

**Emojis**

We wanted to use emojis as these are common images people are used to seeing in their everyday life. We thought this distinction was important when considering the differences between this and the password scheme shown in class that uses images of objects. This is due to the fact the emojis are not novel images, and the ones we chose to use are ones most Smartphones users in our generation use or see on a regular basis. We will randomly assign which 3 emojis the user will have to choose from. The graphical representation of the emojis we used are part of the open sourced EmojiOne collection and were taken from: <http://emojipedia.org/emoji-one/> .

**Password space**:

In order to evenly distribute the memory load, we wanted to distribute the desired password space as evenly as possible between the three components. The desired password space was approximately 228.

This number cannot be split in 3 perfectly evenly, so we decided to start with a space of for the random word component. We then figured out how we could have a space as close to as possible for each of the image and emoji components. We decided to have grids of 16 images or emojis where the password consists of any 3 unique tiles in each section. This leads us to a password space of 16 choose 3 for each section. 16 choose 3 is slightly greater than . All three components of our password are independent of one another, so the password space of each of the components is multiplied together to get the overall password space. . Therefore, our total password space is slightly greater than the desired .

## Implementation of Our Password Scheme

All of the source code for our implemented password scheme and testing framework can be found in the PasswordTestingApplication folder in the submitted A2-T01-sourceFiles folder.

### Password Creation

The password created for each site happens in the Password model. Three unique images are randomly selected from a set of 16 landscape images and three unique emojis are randomly selected from 16 emoji images. The random word is chosen from a file of 1024 common 6 letter words. Each randomly chosen portion is then assigned to the password.

### Screen Images of the Scheme in Use

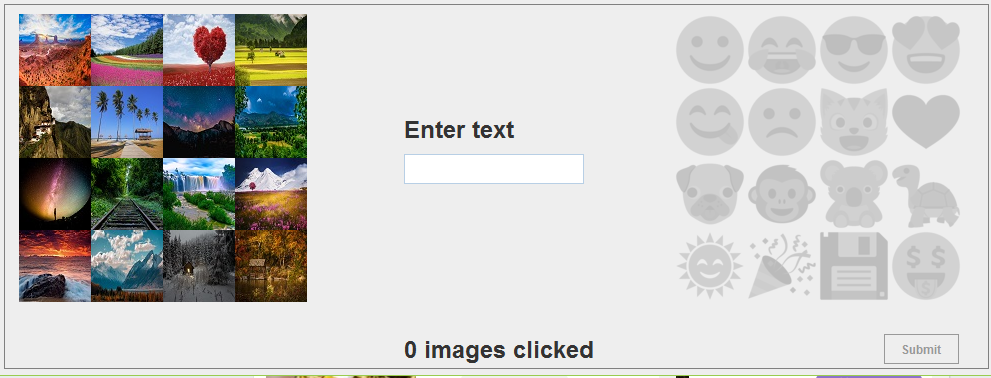


Figure 18. A user is prompted to click the 3 landscape images in their password. All other portions of the scheme are unavailable. The text at the bottom of this screenshot lets the user know how many images they have clicked so far, since images do not remain selected. This is to minimize over-the-shoulder attacks.

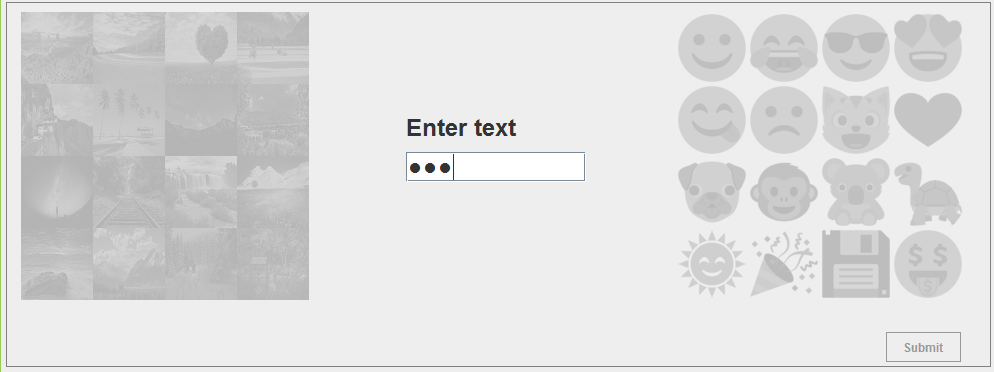


Figure 19. A user is prompted to enter the random word portion of their password. All other portions of the password are greyed out.

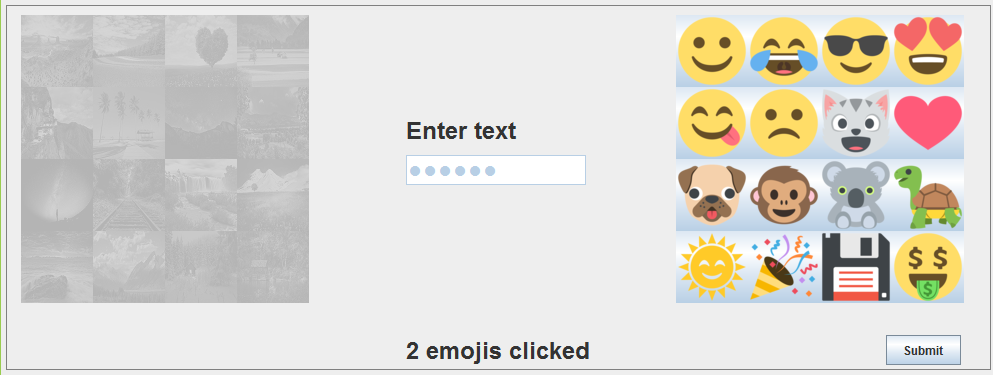


Figure 20. A user is prompted to enter the 3 emojis in their password. All other portions of the password are greyed out. The text at the bottom lets the user know how many emojis they have clicked so far since emojis do not remain selected to minimize over-the-shoulder attacks.

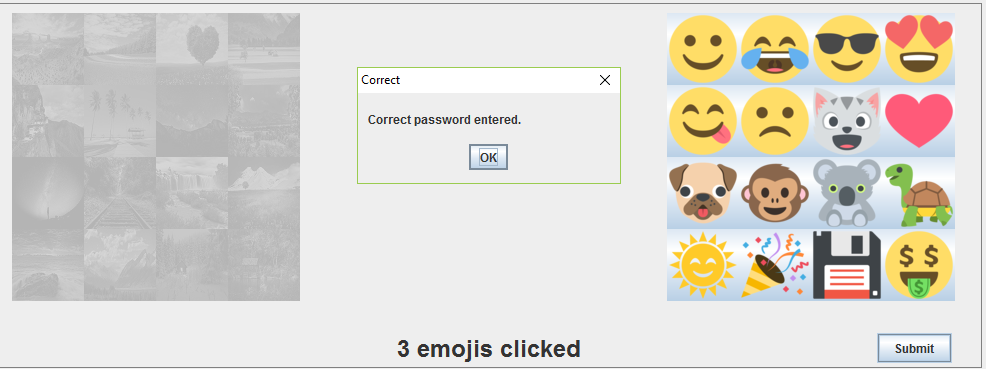


Figure 21. When a user enters their password correctly, they are notified.

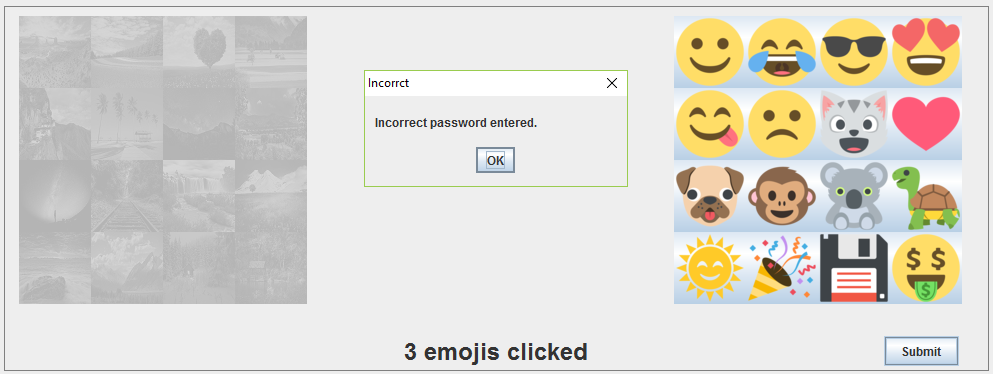


Figure 22. When a user enters their password incorrectly, they are notified.

## Screen Images of the Password Training and Testing System

### Training phase

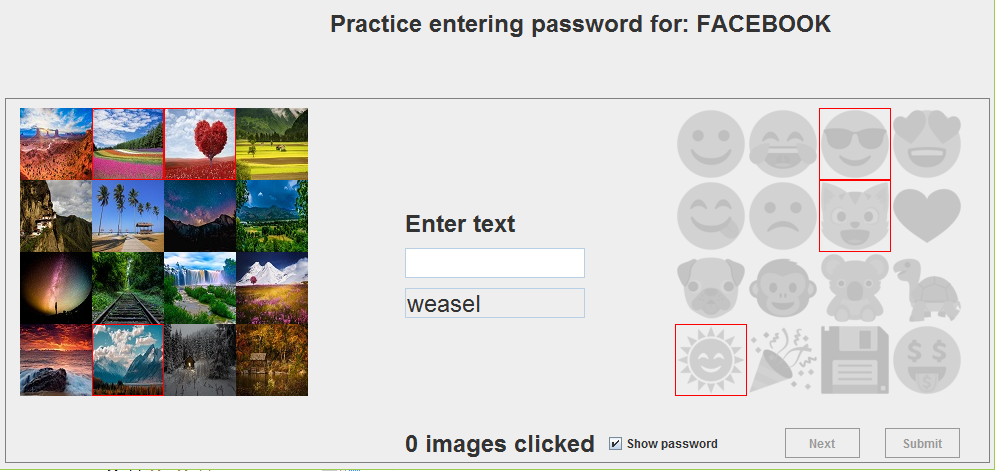


Figure 23. A user is given a new password for a website, and asked to train with it. Default state is to have the password outlined in red. Images clicked indicates how many images they have clicked so far.

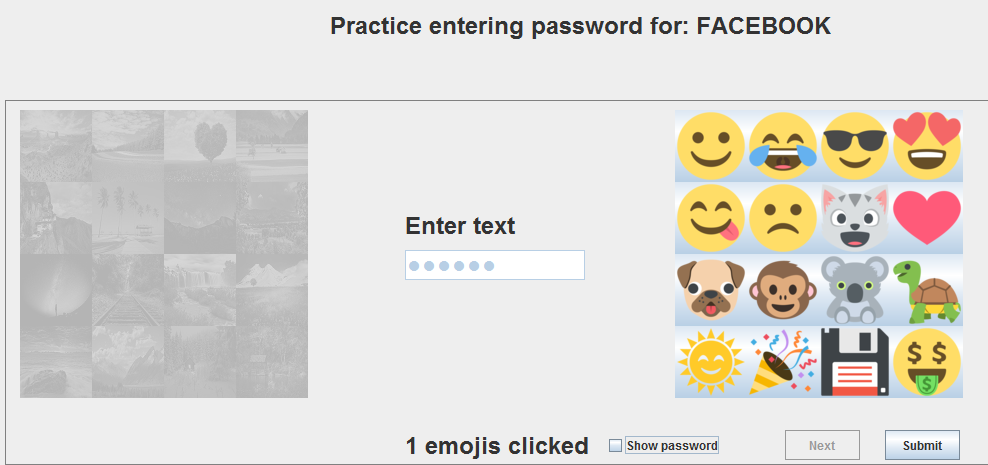


Figure 24. A user can train without the system showing their password. This allows users to try entering their password without help from the system before proceeding to the testing phase. After the user has entered the password correctly 2 times, the next button becomes active and the user can click it to proceed to the next password or to testing if they have trained all three passwords.

### Testing phase

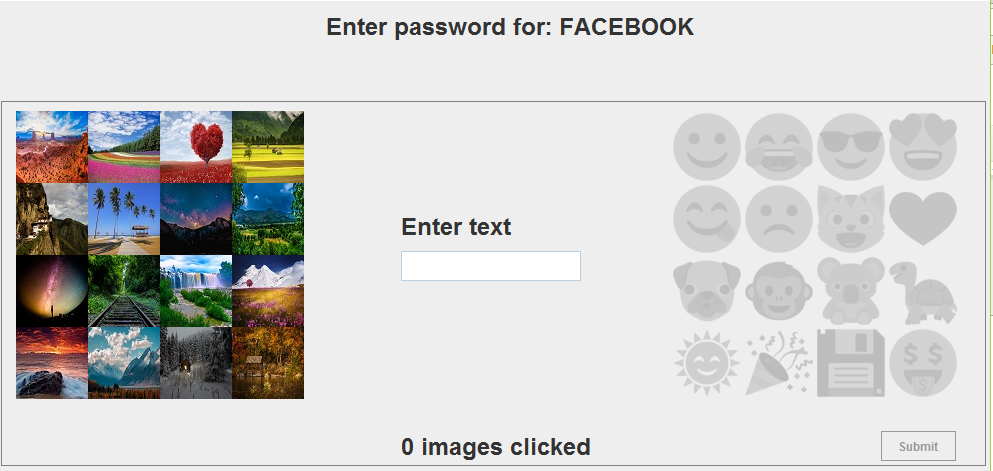


Figure 25. A user is prompted to enter the password for each site they trained on. The system only allows users to enter the passwords in order (ie. enter 3 landscape images, then text, then 3 emojis).

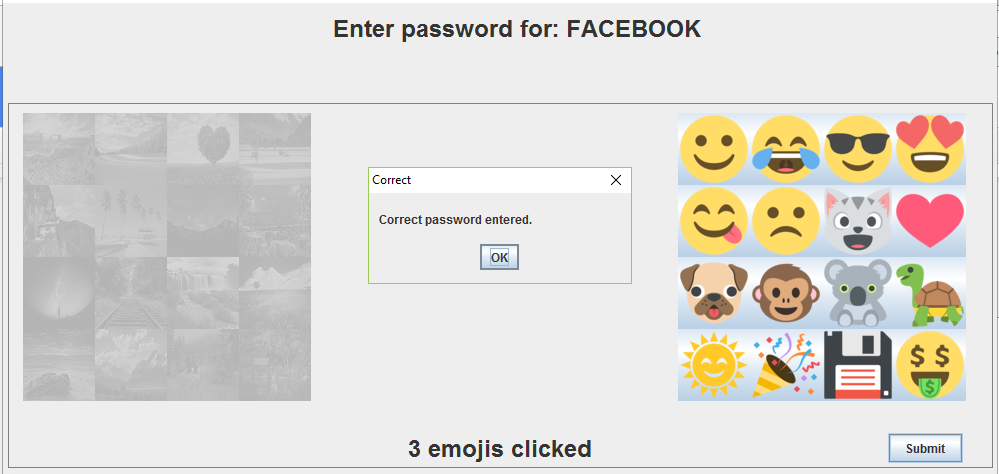


Figure 26. When a user has correctly entered their password, they are notified by the system of this fact. They are then moved onto the next password to enter.

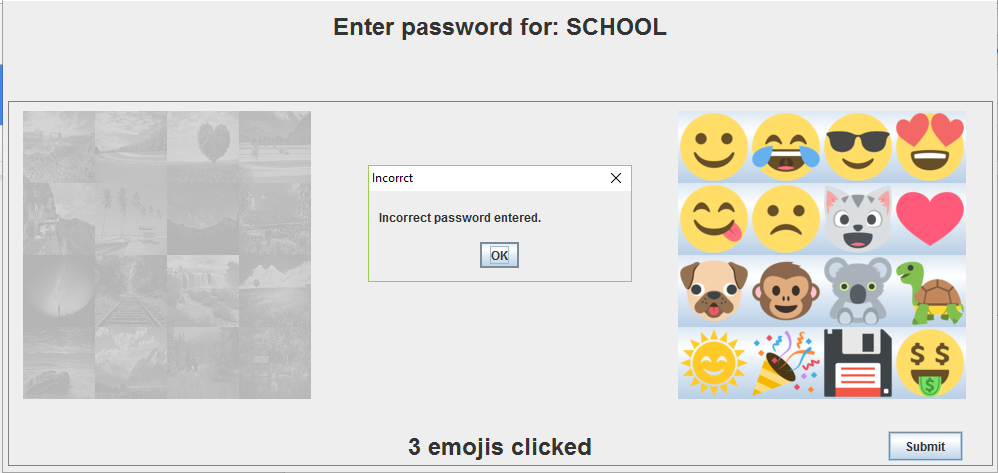


Figure 27. A user is allowed 3 attempts to enter a password incorrectly. For each incorrect attempt a user is notified of the fact that the passwords do not match. When they have had 3 attempts the system moves the user onto the next password to enter.

## Survey on Participant Perceptions of Our Scheme

Our live survey can be found at the following link. <https://hotsoft.carleton.ca/comp3008limesurvey/index.php/629513?lang=en>

The questions are also available in the LimeSurvey Questionnaire.pdf which is found in the A2-T01-sourceFiles folder.

## Usability Testing Data Analysis Compared to Text28

### Descriptive Statistics

Total Logins

|  |  |  |  |
| --- | --- | --- | --- |
|  | Mean | Median | Standard Deviation |
| Text28 | 16.778 | 16 | 4.177 |
| ThreeChunk | 5.222 | 6 | 1.309 |

Table 6. Data for total logins.

Successful Logins

|  |  |  |  |
| --- | --- | --- | --- |
|  | Mean | Median | Standard Deviation |
| Text28 | 14.667 | 15 | 1 |
| ThreeChunk | 3.333 | 4 | 2.473 |

Table 7. Data for successful logins

Unsuccessful Logins

|  |  |  |  |
| --- | --- | --- | --- |
|  | Mean | Median | Standard Deviation |
| Text28 | 2.111 | 1 | 3.855 |
| ThreeChunk | 1.889 | 2 | 1.779 |

Table 8. Data for failed logins

Successful Login Times

|  |  |  |  |
| --- | --- | --- | --- |
|  | Mean | Median | Standard Deviation |
| Text28 | 11.714 | 6 | 9.987 |
| ThreeChunk | 12.654 | 13.75 | 8.309 |

Table 9. Data for successful login times

Unsuccessful Login Times

|  |  |  |  |
| --- | --- | --- | --- |
|  | Mean | Median | Standard Deviation |
| Text28 | 31.370 | 3 | 87.271 |
| ThreeChunk | 9.544 | 10.917 | 8.698 |

Table 10. Data for unsuccessful login times

### Graphs

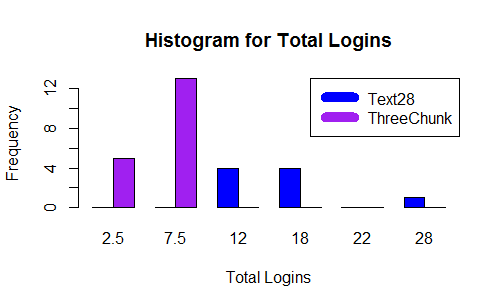


Figure 28. Histogram of the total logins per user for each password scheme.

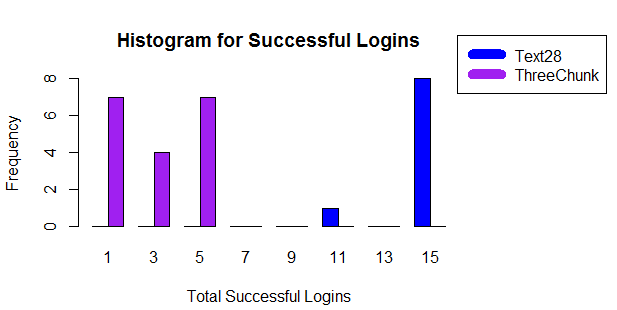


Figure 29. Histogram of the successful logins per user for each password scheme.

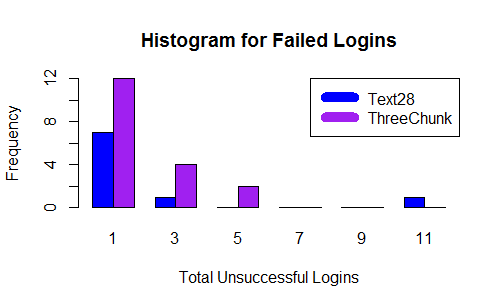


Figure 30. Histogram for the failed logins per user for each password scheme.

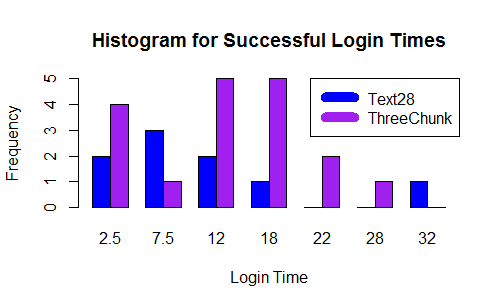


Figure 31. Amount of time it took users to successfully login for each password scheme.

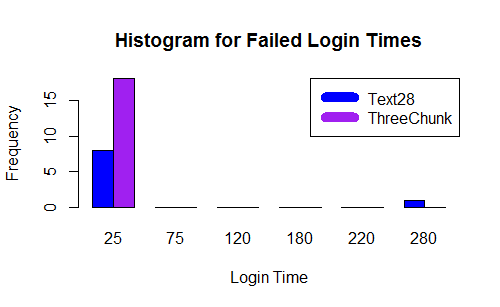


Figure 32. Amount of time it took users to unsuccessfully login for each password scheme.

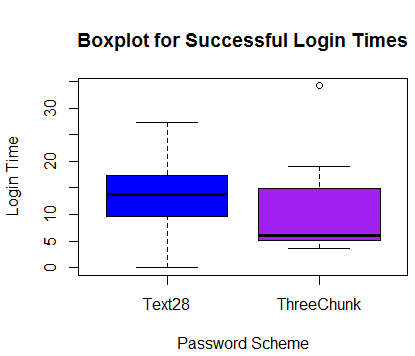


Figure 33. Comparison of the amount of time it took users to successfully login for each password scheme.

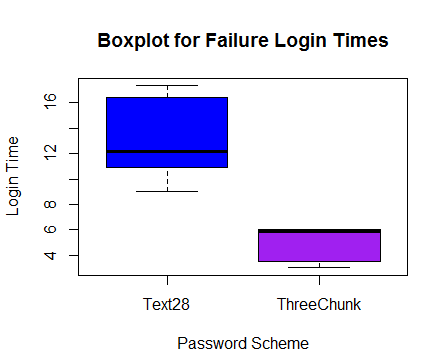


Figure 34. Comparison of the amount of time it took users to unsuccessfully login for each password scheme.

### Inferential Statistics

### Since our data is not normal we must run a Mann-Whitney U test in order to test for statistical significance.

The number of successful and unsuccessful logins were changed into proportions of the total number of logins in order to properly compare the data as Text28 had a greater number of total logins than ThreeChunk (Table 6, Figure 28).

|  |  |  |
| --- | --- | --- |
|  | p-value | Significant? |
| Proportion of Successful Logins | 0.072 | No |
| Proportion of Failed Logins | 0.072 | No |
| Successful Login Times | 0.5885 | No |
| Unsuccessful Login Times | 0.1618 | No |

Table 11. p-values from a Mann-Whitney U test looking for significance.

### Discussion and Interpretation

The histogram for the number of failed logins demonstrates that for both password schemes the majority of users had 1 - 2 failed logins (Figure 30).

Figure 32 shows us that almost all users failed their login in under 25 seconds. The one Text28 user who took 280 seconds is assumed to have abandoned the login attempt earlier and timed out. The histogram for successful login times shows us that the majority of users for the Text28 password scheme took 7.5 seconds to log in while the majority of users for the ThreeChunk password scheme took 12 to 18 seconds to login (Figure 31).

In both boxplots 0 values were removed as these were cases where no success or failure events took place . The boxplot for the successful login times (Figure 33) shows that Text28 has an even distribution while ThreeChunk is skewed, with the majority of the data falling in the third quartile and an outlier occurring above the whiskers. However for unsuccessful logins both distributions are skewed, with the majority of the data points for Text28 occurring in the third quartile (Figure 34) and the majority of data point for ThreeChunks occurring in the second quartile (Figure 34). Once again outliers are removed from the unsuccessful login boxplot due to the large outlier caused by a user timing out.

Although Text28 had a greater number of successful and unsuccessful logins (Tables 7 & 8, Figure 29) once the data was converted into proportions of the total number of logins and a Mann-Whitney U test was run these results were found to be not significant (p = 0.072) (Table 11). Likewise the differences in login times for both successful and unsuccessful logins for both password schemes were found to be not significant (p > 0.05) (Table 11). This could be due to the sample sample size.

Please see the file a2\_Part12\_Q6\_Part1.R in the RFiles folder inside A2-T01-sourceFiles for full code and documentation.

## Usability Testing Data Analysis Compared to User-Chosen Text Password Scheme

Unfortunately, due to a misreading of the instructions the LimeSurvey questionnaire did not include any questions that directly compared a user-chosen text password to the password scheme created for this assignment. Therefore, with permission from Professor Biddle, we artificially created the data for this. We did this by assuming that participants were answering our questions by comparing our password scheme to the common text password that requires 8 characters, with at least one upper and lower case letter, a number and a symbol.

Firstly, for the memorability of a user-chosen text password we opted to record 13 participants saying these passwords are ‘Very Easy to Remember’ and 6 participants saying they are ‘Easy to Remember.’ We feel that these numbers are realistic because the majority of our participants rated the text component of our password scheme as ‘Very Easy to Remember,’ we skewed the results towards more memorable because: a user-chosen text password would not have any others parts that had to be remembered along with the text, a user-chosen password often includes something that user likes, such as a song lyrics or name (Ur et al., 2016) which makes that password more memorable for that user, and users often reuse the same passwords, or similar passwords with minor changes, for more than one account, which can also help a user remember their password (Stobert & Biddle, 2014, Ur et al., 2016).

Next, for the entry time of a user-chosen text password we decided to document 15 participants saying these passwords are ‘Very Fast’ to enter, 3 saying they are ‘Fast’ to enter, and 1 saying they are ‘Neither Fast nor Slow’ to enter. We feel these numbers are realistic because: the majority of our participants rated the text component of our password scheme as ‘Very Fast’ or ‘Fast’ to enter, similarly, from the previous section the Text28 password scheme was found to have the fasted login times (Figure 12), and user-chosen text passwords have been found to have the fastest login times over modified text or graphical passwords(Forget, Chiasson, Oorschot, & Biddle, 2008; Stobert & Biddle, 2013).

Finally, for the security of a user-chosen text password we chose to record 5 participants saying these passwords are ‘Very Secure,’ 11 saying they are ‘Secure,’ and 3 participants saying they are ‘Neither Secure nor Unsecure.’ We feel that these numbers are realistic because this follows the pattern that we found in our survey results for the text portion of our password scheme. We skewed the results towards higher security since we find it unlikely a user would chose to use a text password they felt was unsecure – they would try to make it more secure. Many websites most websites include some sort of visual aid on password creation indicating how secure it believes a user’s password to be, and Ur et al. (2006) found that although users had several serious misconceptions about for the most part they understood which characteristics would make their passwords more resistant or susceptible to attacks, both of which would increase a user’s confidence in a password of their choice.

The final set of questions asking the participants if they would consider using our password scheme were assumed to mean “instead of a regular chosen text password.”

During testing the log files for a participant were lost and were therefore not in the previous section, however their LimeSurvey responses are still included.

### Descriptive Statistics

User Memorability Ratings

|  |  |  |
| --- | --- | --- |
|  | Mode | Median |
| ThreeChunk | 2 | 2 |
| User-Chosen Text | 5 | 5 |

Table 12. User’s ratings of password scheme memorability, 2: Hard to Remember. 5: Very Easy to Remember

User Entry Time Ratings

|  |  |  |
| --- | --- | --- |
|  | Mode | Median |
| ThreeChunk | 4 | 3 |
| User-Chosen Text | 5 | 5 |

Table 13. User’s ratings of password scheme entry time. 2: Slow. 3: Neither Fast nor Slow. 5: Very Fast.

User Security Ratings

|  |  |  |
| --- | --- | --- |
|  | Mode | Median |
| ThreeChunk | 4 | 4 |
| User-Chosen Text | 4 | 4 |

Table 14. User’s ratings of password scheme security, 4: Secure.

User Usage Consideration Ratings

|  |  |  |
| --- | --- | --- |
|  | Mode | Median |
| Bank | 1 | 2 |
| Facebook | 1 | 3 |
| School | 1 | 3 |

Table 15: User ratings of if they would consider using the ThreeChunk password scheme for their personal accounts. 1: Never. 2: Between Never and Maybe. 3: Maybe.

### Graphs

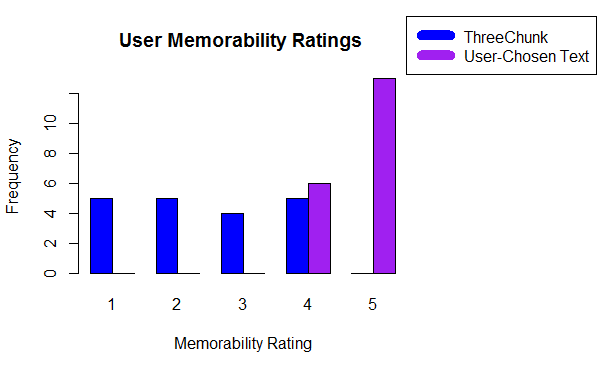


Figure 35. Users ratings of memorability (1: Very Hard to Remember, 5: Very Easy to Remember)



Figure 36. User ratings of password entry time (3: Neither Slow nor Fast, 5: Very Fast).



Figure 37. User ratings of password scheme security (3: Neither Secure nor Unsecure, 5: Very Secure)

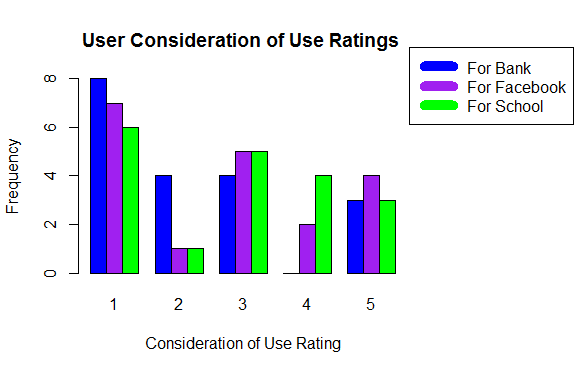


Figure 38. User ratings of if they would consider using the ThreeChunk password scheme for their personal accounts. (1: Never, 3: Maybe, 5: Definitely)

### Inferential Statistics

### Since our data is not normal we must run a Mann-Whitney U test in order to test for statistical significance. For the consideration of use we compared the number of users who answered 1 (Never) or 2 (Between Never and Maybe) with the number of users who responded 5 (Definitely) or 4 (Between Maybe and Definitely) to see if there was a significant difference between how many users would be interested in using the password scheme and how many would not be.

|  |  |  |
| --- | --- | --- |
|  | p-value | Significant? |
| User Memorability Ratings | 5.931-7 | Yes |
| User Password Entry Time Ratings | 0.002 | Yes |
| User Security Ratings | 0.8327 | No |
| User Consideration of Use Ratings for Bank Account | 0.005 | Yes |
| User Consideration of Use Ratings for Facebook Account | 0.114 | No |
| User Consideration of Use Ratings for School Account | 0.009 | No |

Table 16. p-values from a Mann-Whitney U test looking for significance.

### Discussion and Interpretation

Figure 35 demonstrates that the majority of users find text passwords that they choose very easy to remember while the passwords create by the ThreeChunk scheme had a more even distribution, although was no users rated them “Very Easy to Remember” (5). This result was found to be significant (p < 0.05) (Table 16). Figure 36 shows us that the majority of users find chosen text passwords very fast to enter and passwords created by the ThreeChunk scheme fast to enter, which was also found to be significant (p = 0.002) (Table 16). The histogram for user security ratings and p-value shows us that there was no significant difference in how users perceived the security of both password schemes (Figure 37, Table 16). These results are consistent with Tables 12, 13, and 14. Finally Figure 38 shows that the majority of users would not want to use the ThreeChunk password scheme for any of their personal accounts. However, only for a bank account were users significantly more likely to not want to use the ThreeChunks scheme (p = 0.005)

Please see the file a2\_Part12\_Q6\_Part2.R in the RFiles folder inside A2-T01-sourceFiles for full code and documentation

## Interpretation

Since there was no significant difference in the proportion of successful or failed logins or the login times for a successful or failed login we believe our password scheme not better than the Text28 password scheme. However, we believe that with some graphical changes, such as putting some space between the landscape images and finding images that compress to a smaller size better, the usability of our password scheme could be increased.

Nevertheless, as can be seen in Figure 38, it is unlikely that users would prefer it over a chosen-text password as the many perceived it as less memorable and taking longer time to enter than a chosen-text password (Figures 35 and 36).

References

Forget, A., Chiasson, S., van Oorschot, P.C., and Biddle,R. (2008). Improving Text Passwords Through Persuasion. ACM SOUPS 2008.

Stobert, E., & Biddle, R.(2013). Memory retrieval and graphical passwords. Proceedings of the Ninth Symposium on Usable Privacy and Security.

Stobert, E., &Biddle, R. (2014). The Password Life Cycle: User Behaviour in Managing Passwords. In Proc. SOUPS, 2014

Ur. B., Bees, J., Segreti. S.M., Bauer, L., Christin, N., Cranor, L.F. (2016). Do Users' Perceptions of Password Security Match Reality?, Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems.