The Bird is the Word: A Usability Evaluation of Emojis inside Text Passwords

Tobias Seitz
LMU Munich
Germany
tobias.seitz@ifi.lmu.de

Florian Mathis
LMU Munich
Germany
mathis@cip.ifi.lmu.de

Heinrich Hussmann
LMU Munich
Germany
hussmann@ifi.lmu.de

ABSTRACT

Passwords still represent an annoying burden for millions of Internet users. Helping people create memorable and secure credentials is therefore an important goal for web-service providers to satisfy user needs. Due to the good memorability of pictures, emojis may be a suitable tool to create memorable and secure passwords. These small pictograms have seen an enormous rise in recent years, but their usage in regular passwords has not been explored for the Web. In a two-part user study with 40 participants we investigated if and how emojis are suitable in this context. We asked users to create passwords that contained both regular alphanumeric characters and emojis. The study shows that users' primary selection strategy was to create meaningful relationships between the emoji and the rest of the password. We also found that platform dependent renderings of emojis do not necessarily reduce usability, if the object represented by the emoji is distinctive enough. As websites are already starting to allow emojis in passwords, it is important to evaluate this step carefully. Our results can inform this decision and provide pointers to the usability implications.

CCS CONCEPTS

• Security and privacy → Usability in security and privacy;

KEYWORDS

Usable security, passwords, emojis, authentication

ACM Reference format:

T. Seitz, F. Mathis, and H. Hussmann. 2017. The Bird is the Word: A Usability Evaluation of Emojis inside Text Passwords. In Proceedings of the 29th Australian Conference on Human-Computer Interaction, Brisbane, QLD, Australia, November 2017 (OzCHI 2017), 9 pages.
https://doi.org/10.1145/3152771.3152773

OzCHI '17, November 28-December 1, 2017, Brisbane, QLD, Australia

ACM 978-1-4503-5379-3/17/11 $15.00
DOI: https://doi.org/10.1145/3152771.3152773

1 INTRODUCTION

Although researchers and practitioners try to find replacements for password-based authentication, it remains the most commonly used strategy on the Web. Until they are replaced, passwords will continue to cause usability troubles for Internet users. They have to manage a variety of accounts with multiple levels of importance. At the same time, users are advised to create unpredictable and unique secrets. This task is difficult, so users often simplify it by picking easy and memorable passwords, which they also re-use in many different places [7,14,16]. However, the downside of this behavior is that attackers can easily gain access to user accounts by centering their attacks on the predictability of user-chosen passwords. The consequences include financial damage or loss of private information that boosts social engineering attacks [41]. To alleviate this problem, we should make it easier for users to select passwords that are both secure and memorable.

In this paper, we take a look at how emojis could mitigate the problem and achieve this goal. Emojis are small pictograms to express emotions or visualize words. Now part of the Unicode standard, they are used ubiquitously in text-based communication in messenger apps and on the Web. Even the Oxford dictionary's "word of the year" was the "tears of joy" emoji (😂) in 20152, which highlights the cultural impact of these colorful symbols. It is therefore natural that emojis have made their way into authentication mechanisms [15]. The notion that image-based passwords could boost password usability was established decades ago. The omnipresence of emojis, however, gives rise to a new range of research questions in this context.

[2] http://blog.oxforddictionaries.com/2015/11/word-of-the-year-2015-emoji/

Until now, emojis were examined as alternatives to PINs, i.e. passcodes that only consist of emojis and no other characters[3]. However, while many services still disallow emoji characters inside passwords, some services can handle them already, e.g. Twitter, Slack, or StackOverflow. Although emojis are most commonly used on mobile apps, users still need to able to authenticate on desktops, if their password contains emojis. Currently, emoji input on physical keyboards is cumbersome, but not impossible. Alternative solutions to enter emojis on the desktop exist [2,8,28] and can enable the use of emojis in passwords across devices. Another issue that has not been addressed is the effect of platform-dependent emoji-images on password memorability: At the moment, all major platforms have different visual representations of the Unicode symbols [27], which makes the problem of identifying and interpreting them even more severe.

1.1 Research Questions

The potential usage of emojis in passwords led us to investigate the following questions:

RQ1: Are emojis suitable for creating both memorable and usable passwords?

RQ2: How do different renderings of the same emojis impact memorability?

RQ3: Which input options do users prefer when entering "emoji-passwords" on the desktop?

1.2 Summary of Findings and Contributions

We provide a first empirical dataset on users' selection strategies of emojis inside text-passwords, which suggests that users tend to relate emojis to the rest of the password and other things they like. We show how memorability is affected by different renderings of the same emojis. Some users were troubled by altered renderings, while the majority of the study participants succeeded to authenticate. Despite positive usability ratings, our participants were reserved towards adopting emojis in password authentication. Finally, the data allows us to discuss a specific set of implications to consider when Unicode and thus emojis are enabled inside passwords.

2 BACKGROUND AND RELATED WORK

We position our work in the field of usable security and the study of knowledge-based authentication mechanisms. To better motivate and differentiate our work from existing literature, we provide a brief overview of the problem space of password usability and how it was tackled with graphical authentication schemes.

2.1 Problem Space: Password Usability

A lot of academic research as well as data breaches from large online service providers have repeatedly demonstrated how predictable user-selected passwords are [4,18,25,40]. One of the reasons for this seemingly careless behavior is the large number of accounts that users have to manage. In a large-scale field study, Florêncio and Herley found that an average user logs into 25 different accounts on a regular basis [12]. Remembering access information for each of the accounts is demanding if not overwhelming for users, especially because they are advised to create unique secrets for different services. To reduce the burden, users rationally strive to create memorable passwords that serve as suitable credentials for a multitude of services. Consequently, users include information about themselves and "important others" in their passwords to facilitate memorizing the credentials [6,22,30]. Personal details often comprise meaningful dates or likings [21,40]. However, attackers can easily exploit this kind of knowledge about typical user behavior to reduce the number of guesses needed to compromise passwords [18].

Researchers and practitioners have thus investigated ways to help users create stronger passwords that are also usable. One dimension of password usability is how memorable they are, or how easy it is to memorize them. For instance, passphrases consisting of multiple words were proposed to facilitate memorization [17]. The theoretical password space of multiple-word passphrases is large enough to provide reasonable protection against guessing attacks [33]. The effective password space, however, may be a lot smaller [5], which is why suggesting suitable passphrases was studied. Such passphrase suggestions can have a positive impact on the strength of user-selected passwords [23,32], and they can also be more memorable than regular passwords [17]. However, Keith et al. as well as Shay et al. point out usability drawbacks that are mostly related to the increase of typographical errors with longer passphrases [17,33]. Mnemonic phrase-based passwords, where a user selects a short password based on the starting letters of a sentence, have shown benefits in terms of memorability. However, they are also vulnerable to sophisticated dictionary attacks [21,44]. Finally, feedback mechanisms have been studied to increase password strength and usability of password selection. Most notably, password meters providing real-time feedback on estimated strength can positively influence users while maintaining or even boosting usability [9,34,38,39].

In summary, real-time feedback and suggesting phrases as mnemonic device are useful approaches to make users select stronger passwords, but they might not be the best solution for creating very memorable and usable passwords.

[3] https://www.intelligentenvironments.com/now-you-can-log-into-your-bank-using-emoji/

2.1 Searchmetric Graphical Authentication

Another attempt to foster more memorable passwords tries to leverage the power of human visual perception. This often requires a step away from alphanumeric text-passwords and towards graphical authentication schemes. The general idea behind these schemes is to exploit people's abilities to recall images better than text, often termed the picture superiority effect [29]. However, sometimes the effect is void because there are too many distracting images. While there are a multitude of different graphical authentication mechanisms, our work is most related to recognition-based authentication through icons or pictograms, also known as searchmetric systems [29].

Wiedenbeck et al. focused a graphical password scheme for public spaces [43]. Their Convex-Hull-Click (CHC) system distributes random icons (e.g. from popular software at the time like Netscape Navigator, Adobe Reader, Mac OSX) across the screen. During enrolment, the user has to choose three or more icons as their "pass-icons". To authenticate, they click within the convex hull[4] multiple times in a row and not on the icons directly. Wiedenbeck et al.'s usability evaluation revealed long login times, while the pass-icon concept was memorable enough for the participants to log in after a week. Gao et al. proposed a similar, yet simpler scheme with their "ColorLogin" [13]. Here, the user also does not click icons directly, but indicates the row of a colored grid that contains the icons. Moving away from indirect selection techniques, Stobert and Biddle presented "Object PassTiles" [35], where users log in by clicking on images of objects inside a fixed grid. The image locations are randomized at each login to reduce susceptibility to shoulder-surfing attacks. In a user study, Object PassTiles showed memorability benefits, whereas log-in times degraded compared to other schemes. Object recognition appears to cause more mental effort, especially if object locations are randomized.

All the aforementioned projects followed a recognition-based authentication approach. Renaud and De Angeli argue that this particular way does not improve authentication usability, because the picture superiority effect cannot be directly applied here [29]. Indeed, authentication times are longer in most graphical authentication schemes. To reduce login times and while benefitting from the picture superiority effect, emojis have recently been considered a possible solution.

2.3 Emojis in Authentication

By direct translation from Japanese, emojis are "picture words" [37]. We often encounter these small pictograms in messenger apps and social networks when people use them to hint emotions. Sometimes they are also used to replace words within blocks of text, e.g. 🐦 instead of "bird". There are currently around 2600 different emojis as part of the Unicode standard. In 2015, more than half of the text on Instagram contained emojis, which is attributed to the introduction of easy input methods on both iOS and Android5. Since emojis are encoded as regular symbols, one of their benefits is the small size in storage, which make them also easy to transfer over networks. However, for users it is often difficult to agree on the specific meaning of some emojis [26]. Apart from that, they are a fairly universal language, as most users across different countries can make sense of their meanings, e.g. animals or objects.

Due to the growing adoption and benefits of emojis, researchers tried to incorporate them into authentication schemes. The Intelligent Environments company introduced an emoji-based passcode6. Users can select four emojis from a 9x5 grid to authenticate on mobile banking apps. Another system by Golla et al. [15,19] was recently proposed as an alternative unlock mechanism for mobile devices. Their EmojiAuth system has been evaluated with two scientific user studies. It is a replacement for unlock-PIN-pads, where each digit is replaced by an emoji. Both the security and usability evaluation showed conclusive evidence that the use of emojis is recommendable for this purpose. Apart from PIN alternatives, we only found one investigation into emojis inside text-based passwords: Al-Husainy and Malih proposed a prototype which allows users to include emojis in their password for an operating system account [1]. However, they did not report an evaluation of the concept.

2.4 Summary

In summary, we interpret related work as a promising direction for emoji-based authentication, because the results to date show user experience benefits. So far, emojis have mostly been used at scale for replacing PINs, and the field is open for research on emoji-passwords: It is ill-defined how users are to enter emojis when they log into websites using desktop computers, and how different renderings of the same emojis affect the process. The latter is a problem that does not occur with the aforementioned schemes. Our work aims to deliver first empirical insights into these issues to better understand the practicality of emoji-passwords.

[4] inner area of a polygon formed by the icons if connected by virtual lines

[5] https://engineering.instagram.com/emojineering-part-1-machine-learning-for-emoji-trendsmachine-learning-for-emoji-trends-7f5f9cb979ad

[6] Intelligent Environments' emoji passcode https://vimeo.com/130728753

3 USER STUDY

To answer our research questions and to make a recommendation regarding the suitability of emojis for password-based authentication in their current form, we ran a two-part user study.

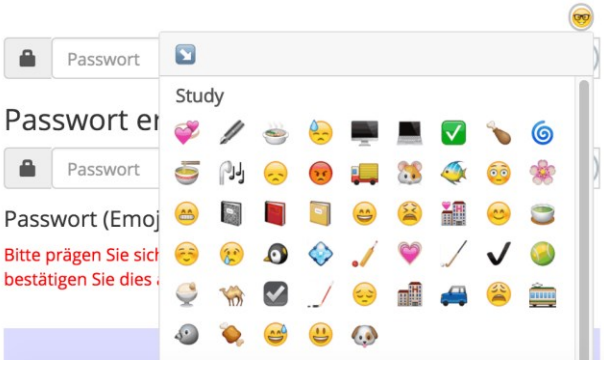

Figure 1: "Emoji-Picker", graphical user interface to select from the available emojis (n=50). The □-button on the top right brings up the dialog. The order is shuffled each time.

3.1 Study Design

Our investigation revolves around the idea that passwords could become more usable when they include emojis (RQ1). Thus, we divided our experimental study into two parts, namely the selection and the recall part. During selection, participants create a password that includes an emoji in a fictional scenario. After one week, they have to authenticate with the same password (recall). Here, we supply one group with different renderings of the same emojis to study the potential impact of switching platforms on authentication success (RQ2). To answer RQ3, we also provided different means to enter emojis on a desktop computer. Emojis could either be entered with a visual point-and-click interface7 (the "picker"), or by entering a "short-code", e.g. :dog:, which was inspired by the way emojis can be entered on Slack. The latter method is a recall-based mechanism rather than a recognition-based scheme, which distinguishes it from related work. To facilitate selection, we show the users a cheat-sheet with both the short code and the corresponding emoji.

The experiment was carried out in the lab of our institution for the first part, while the recall part was done remotely via the Internet after one week. The sessions in each part of the study were finished within two working days. Alongside quantitative measurements, we aimed to understand the metrics better by qualitative interviews and assessments. Moreover, the selection part required a back and forth between an emoji-password field and the questionnaire, which we could not fruitfully combine in common survey tools. This made us choose the lab setting and facilitates comparison to related work [15].

3.2 Conditions (Independent Variables)

The selection part was identical for all participants. The available emoji-renderings were the images as used by iOS 9.3. We expected that study participants would be familiar with them, because WhatsApp uses the exact same renderings on all platforms and it is the most used messenger service in the country of the study (Germany) [3]. We chose a subset of 50 emojis (14 from the "people" category, 7 from "nature", 6 "foods", 5 from "activity", 5 "places", 6 "objects", and 7 "symbols"), see Figure 1. We approached the selection of the 50 emojis like Golla et al. [15] and considered the most used emojis8. Moreover, we tried to balance the categories. While Golla et al. were looking to exclude emojis that were too similar to each other, we added more emoji from the "people" category. Those are popular and very similar to each other, and hence they allow us to study the effect of distinctiveness on memorability.

For the second part (recall), one independent variable "rendering" was used in a between groups design. The control group had to log in with the identical emoji set as in the selection part (see Figure 1), while the experimental group received the emojis as rendered by Android 7.0 (see Figure 3)9. The position of the emojis was shuffled in both conditions for each participant. As shown in Figure 3, some emojis are very similar across platforms while others are not. We hoped to detect a potential effect with this.

3.3 Measurements (Dependent Variables)

In the selection part, we measured which emojis were chosen and where they were positioned inside the password. Before we encrypted and stored the passwords, we analyzed them and stored key metrics like length, number of lower- and uppercase letters, digits, and symbols (LUDS metric [42]). Also, we collected subjective usability ratings about the use of emojis in the password. Since we did not want to shoulder-surf participants during password selection to avoid biasing them, we relied on self-report regarding the selection technique (point-and-click vs. short-code) and its perceived usability. In terms of qualitative aspects, we wanted to know how participants went about selecting their passwords. As control factors we recorded demographic data, and self-assessment of past emoji- and

7 http://ned.im/wdt-emoji-bundle/

8 http://emojitracker.com/

9 https://emojipedia.org/google/android-7.0/

password behavior. Usability assessments were made on 5-point scales, e.g. on the usefulness and perceived ease-of use of the system. Here, we took a part of the items of the questionnaire as shown in [15].

For part two, we measured login overall success rates, beside the number of attempts to succeed. We inquired how the participants went about memorizing and recalling their passwords. We again measured subjective usability ratings and asked for qualitative feedback on the concept.

3.4 Procedure and Questionnaires

After a short briefing, where the experimenter explained the general purpose of the study ("entering emojis and passwords") and that we will collect data passively, participants signed a consent form. They proceeded to guide themselves through a questionnaire on a PC. First, the questionnaire instructed them how to create their user-id by taking the first letter of their parents' names, birthplace and birth month, followed by demographic information. Afterwards, we gauged password coping strategies. Then the questionnaire highlighted the differences between emojis and emoticons to ensure a shared understanding for all participants. After self-assessing personal emoji usage and attitude, participants created a password. Here, we established situational awareness by introducing a scenario. The participants should imagine that WhatsApp requires all users to secure their account with a password that has at least one emoji and at least eight regular characters. While they were performing this part of the study, the experimenter encouraged them to think-aloud, while he stayed in the background to avoid shoulder-surfing. The password input boxes were accompanied by real time feedback on the policy requirements. Participants could also see their password in plain text below the password fields and had to check a box to indicate that they memorized their passwords (see Figure 2).

After selecting the password, the participants indicated and rated the usage of the short-codes and the picker, as well as the emoji-password concept in general. The questionnaire concluded with an interpretation task of two emojis (and) to see how well these would have been interpreted, if they had been included in the provided set of emojis.

Exactly one week after selection, we invited participants to come back by sending them a link to a survey. There were two different links, i.e. a separate link for each of the two experimental groups. Participants were randomly put into either the control or the experimental group. The survey asked them to recreate their user id using the same algorithm as before. Afterwards, they were asked to authenticate with their emoji password. If they failed, they could try two more times. After the first failure, we displayed the cheat-sheet to facilitate identifying the emoji through the short code. After the third failure, the participant proceeded to finish the questionnaire.

3.5 Hypotheses

We deduced the following set of hypotheses for our study:

Memorability:

H0a: password recall is independent of emoji rendering

H1: identical emoji rendering is beneficial for password recall

H2: short-codes facilitate recall in case of different renderings

Usability Perceptions:

H0b: Users perceive emoji-passwords just like regular text passwords.

H3: Users perceive benefits in creating passwords that contain emoji.

3.6 Sample and Demography

We invited people to participate in a "study on emojis" and spread the registration links via social networks. Also, the sign-up form for the study was sent out via an official university newsletter. We offered incentives of 5€ shopping vouchers (≈6 USD) for fifteen minutes of their time. In the first part of the study, 40 people participated. They were all students at our university, and between 19 and 44 years old (M=23). Users in this age range show the most intense use of emojis [10]. 39 participants came back for part two after one week (drop-out rate 2.5%). In the second part, 20 participants were in the control group, and 19 in the experimental group.

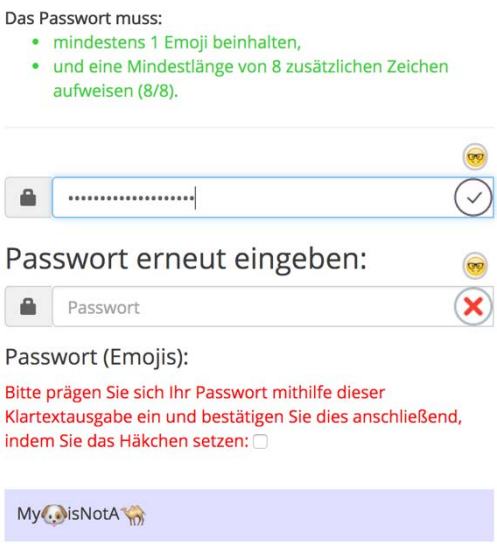

Figure 2: Password selection user interface. A policy with real time feedback informs the user which criteria have already been fulfilled (green text). The password needs to be re-entered to exclude errors. The red text asks the user to confirm that they have memorized their password.

4 RESULTS

Overall, the participants leaned towards a positive assessment of the usability of the emoji-password concept. On a five-point scale (1 = strongly disagree, 5 = strongly agree), they rated the statement "I liked adding an emoji to my password" with an average of 3.6 (SD = 1.19, Md = 4). However, at the end of the study they reported that they probably would not use emojis in their personal passwords, yet (M = 2.6, SD = 1.3, Md = 2). We interpret this as indication that emoji-passwords might not be practical, yet. The following sections shed more light on this finding.

4.1 Emoji Selection and Perception

From the given set of 50 emojis, 22 different emojis were chosen (see Figure 3). Most commonly, participants chose the camel (:camel:, 🐫, n=5), penguin (:penguin: 🐧, n=5), diamond (:diamond_shape_with_a_dot_inside: 💠, n=3), tea (:tea: 🍵, n=3), dog (:dog: 🐶, n=3), cherry blossom (:cherry_blossom: 🌸, n=3), and grinning face (:grin: 😁, n=3). All other emojis were chosen less than three times each. Six users selected two emojis in their password. Password selection took 52.9 seconds on average (SD = 55.00).

Emoji Positions

Regarding the position of emojis, 23 passwords ended with an emoji while 10 started with an emoji. Also, 9 passwords had an emoji between the first and last character. Those who picked two emojis inserted them in between and at the end (n=2), as the first and last characters (n=2), only in the middle as two consecutive characters, or as the first two characters.

Figure 3: Emojis as chosen by participants in the first part rendered by iOS 9.3 (left side) and Android 7.0 (right side). The experimental group was exposed to the Android version in the second part.

Selection Strategies

Prompted with a list of likely selection strategies we found that most (n=20) participants related the chosen emoji to the rest of the passwords. Four indicated using a particular emoji frequently, while another four said it seemed random enough to increase the strength of the password. The rest either chose it by association to a life event (n=3) or hobby (n=1). Eight participants provided an individual explanation (n=8). When asked whether this strategy to create a password generally matches their real-life password selection, average agreement levels pointed in this direction (Md = 4, M = 3.4, SD = 1.45).

We also used an approach similar to Grounded Theory to identify themes in the qualitative statements regarding the choice of emojis [36]. After open coding, we had 42 codes that we processed in axial coding. As a result, we found six general themes:

- **Internal consistency**: the emoji fits to the alphanumeric part of the password, e.g. a camel that comes after a mangled version of the Greek word for "heat"[10]
- **Context cues**: the user's location or the purpose of the password serve as a cue for the emoji, e.g. a computer ( ) because participants were sitting in front of one, or a check mark (✓) that represents a *correct* password.

[10] P26 broke down the components of their password in high detail.

- **Replacement**: parts of the selected password were replaced by an emoji, e.g. using a *p*enguin instead of the letter p, or replacing a dictionary word with the corresponding emoji.
- **Appeal**: emoji that somehow visually or emotionally appealed to participants (mentioned twice for ❁).
- **Liking**: emojis that represent something the participants like, e.g. a penguin 🐧, or a particular emoji that they feel connected to.
- **Usability and Security**: an emoji that either increases the perceived usability of the password or its security, because it particularly facilitates typing or one that may be a less frequently used emoji.

In the final (selective) coding stage, the core topic is substantiated: The overall storyline of the qualitative description of selection strategies can be read as a way to **improve memorability** of the selected password. Most participants focused on later retrieving the password from memory and thus selected an emoji that best fit this purpose.

4.2 Input Methods

Most (n=32) participants reportedly used the point-and-click tool to enter the emoji. Five relied on the short-codes and three tried both methods. Participants using the picker rated its usefulness with 3.81 (SD=1.09) out of five points (see Figure 4). Thus, these ratings show a fairly positive attitude towards entering emojis on the desktop with an on-screen user interface. The five people who entered the emoji with the short-code gave it an average rating of 4.2 (SD=1.3). Interestingly, those who used both methods started with the picker and then found it more convenient to use the short-codes.

A qualitative analysis of the explanations and think-aloud protocol reveals that participants found advantages and disadvantages for both methods. The picker received praise for its ease and speed of use, and for not being prone to typing errors. Also, it was perceived as an already familiar way to enter emojis (e.g. from the WhatsApp web interface). On the other hand, participants mentioned that it could be overlooked and that it is somewhat cumbersome to switch between typing and pointing. Those who used the short-codes generally liked the speed of entry and low effort. However, the learning and unfamiliarity of the codes as well as the more demanding selection process were the key drawbacks. Overall, participants found their chosen input method convenient across both parts of the study (M = 3.46, SD = 1.23, Md = 4).

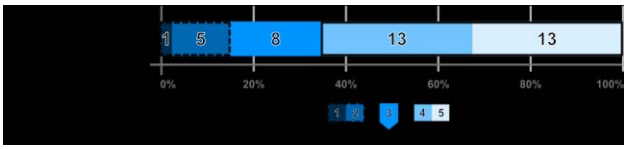

Figure 4: Subjective usability ratings of emoji picker. The majority found it useful. The short-codes were only utilized by five users.

We conclude that the participants deliberately chose their preferred input method and that they were happy and confident with the outcome of their choice.

4.3 Memorability and Recognition

The login success rates were very similar for both the control and the experimental group. Displaying a different emoji resulted in 13 successful logins and 6 failures, while the identical emoji set produced 15 successful and 5 failed login attempts. The small difference is not statistically significant on the 0.05 alpha level ($\chi 2(1) = 0.21$, $p = 0.65$). Twelve participants in the control group succeeded immediately, which was the case for eight users in the experimental group ($\chi 2(1) = 1.25$, $p = 0.26$). The medians of failed attempts were 0 in the control group and 1 in the experimental group. We also checked if there were emojis that were more likely to lead to a failed login. Most notably, the likelihood of login failures was approximately twice as high in the experimental group if the respondents had picked an emoji from the "smileys and people category" (see Figure 1).

We asked participants whether they noticed any differences in the renderings of the emoji. Although the emojis stayed the same in the control group, 7 out of the 20 participants reported to have noticed an unspecified difference. Although we remain unsure what caused this impression, we explain the responses with the randomly shuffled positions of the emojis inside the picker. In the experimental group - who actually saw different emoji renderings – 16 out of 19 said they noticed the difference. Eight of them remained certain which emoji they had picked (≈ 42 %), while another eight reported that this had made them unsure. There were three instances where the different rendering was an insurmountable hurdle: One participant said she was sure to remember her password, because she memorized it by the happy face to express a certain emotion. However, she failed to log in because she did not recognize the correct emoji that was supposed to represent this emotion. Another participant remarked that he had to rely on a Google search to figure out how his emoji looks on other platforms. Only then could he proceed to log in. Finally, one participant mentioned the different visual representation of the check-mark as a reason why he had to retry, although he could remember his password.

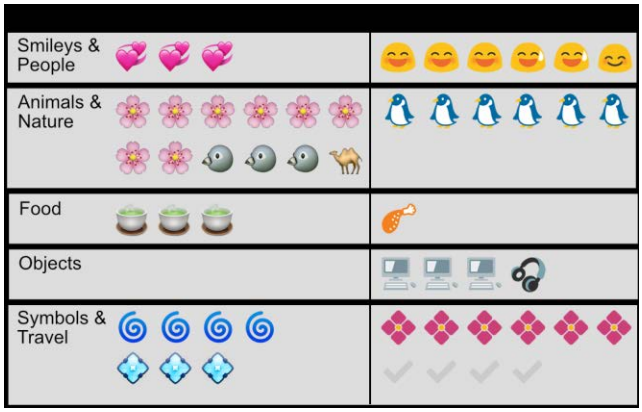

Figure 5: Emojis included in login failures, repeated by the number of attempts. The experimental group (right) made more errors when their password included smileys.

Other qualitative reports indicate that 18 participants reconstructed their password relying on the emoji-picker interface, while only one person noted that the recollection of the short-code was helpful. Moreover, we collected subjective a-priori memorability assessments on a five-point scale regarding the statement "Using emojis in my password would make it more memorable". The average agreement was 2.76 (SD = 1.38, Md = 2). We again probed this assessment after one week and found a small upward trend (M = 3.05, SD = 1.30, Md = 3).

While three comments pointed towards a different interpretation, the numbers suggest that the different renderings of emojis did not have a notable impact on the login success rates, i.e. reproducibility of the emoji-passwords. The small sample size increases the likelihood of type-2 for statistical tests, though. In that case, this would explain the low voluntary readiness to include emojis in their personal passwords. Nonetheless, qualitative evidence is conclusive that users try to leverage emojis for creating more memorable secrets, if the emojis are required by the password policy.

4.4 Interpretation Task

We prompted participants to name two words that they associate with the 🤷 (service woman) emoji, and two words for the 🙏 (praying hands) emoji. We coded the responses and found a large range of themes for the "service woman" emoji. There, 14 distinguished themes appeared in at least two of the participants' responses. The most prominent themes were "female" (n=6) and "pointing" (n=5). On the other hand, the participants were mostly concordant regarding the "praying hands" symbol. Here, we only found seven emerging themes. The most common themes here were "pray" (n=19) and "beg" (n=9). Thus, this anecdotal example highlights the ambiguity of some emoji, while others evoke clearer associations.

4.5 Limitations

Like many other password studies, our study has a few potential limitations. First and foremost, the size of our sample and its homogeneity (students, mostly technical background) lower the predictive power of the hypothesis tests, which makes false-negatives more likely. Thus, a more diverse sample might reveal more problems if the rendering of the emoji changes. However, the study participants are among the user group who is most familiar with emojis and can be seen as the most likely group to adopt emoji passwords.

Moreover, ecological validity might be reduced because the passwords were purely created for the purpose of the study and participants knew they would not really be using them later. However, we introduced a plausible scenario, as the participants' self-assessment of their selection strategies and behavior shows. Research on the ecological validity of password studies recommends using such specific scenarios and finds that study behavior then often matches real-world behavior fairly well [11,20]. Nevertheless, we can only draw conclusions about short-term usage of emoji-passwords at this point.

Finally, our study focused on usability and general attitudes. Thus, we did not perform guessing attacks on the passwords created in our study. We also reduced the number of emojis to 50. In theory, this small number would already increase the password-space by approximately three orders of magnitude if one emoji is used in an 8-character password ($log(50 * 8) \approx 2.6$). However, in reality, only a subset of the available emojis at fairly predictable positions was chosen. Hence, while the passwords may be memorable, they might not show the full spectrum of security benefits.

5 DISCUSSION

In the following, we put the results into a larger context and show implications on the use of emojis within text passwords.

5.1 Distinctiveness is Key

We observed that only one quarter of the participants picked an emoji that visualizes "emotion", i.e. smileys or hearts. Perhaps, it is more difficult to distinguish emotions in smileys from one another, while it seems effortless to tell animals, food and objects apart. Note that this problem exists even if the emoji renderings are consistent, i.e. only one style is used. Thus, to make emojis more usable for passwords (RQ1), service providers would need to preselect the available options just as we and Golla et al. did [15]. If smileys are enabled, it is important to choose only very distinctive ones. Unrestricted in-the-wild usage of emoji-passwords might be more problematic if there are different variations of a particular object, e.g. different dogs or even branded emojis as currently promoted by market analysts [10]. Users might then remember that they used a beer-emoji, but not if it was a Becks or Budweiser. We suggest using whitelists of emojis that are distinctive enough to tell them apart easily.

Tough ones like the "service woman" in our interpretation tasks could be blacklisted. However big the number of available emojis, the theoretic password space increases with each one.

5.2 Selection Strategies are Based on Relationships

We found out that our study participants tried to relate their emoji to either the rest of the password, the context, or to themselves. The former is particularly interesting, because this strategy enriches a simple password and is not possible with other graphical authentication schemes. Users already tend to use autobiographic memories in their alphanumeric passwords, and emojis complement this strategy. Apparently, this works best with objects or animals because their meaning is specific and often they are associated with something that users like. In turn, this makes them memorable (RQ1). We assume that this association is more important than the used rendering. The picture superiority effect might be negligible if not void in this context. Thus, we assume H0a is true and there is no true effect of rendering on recall rates – provided that there are no different types of the same emoji, e.g. multiple dogs (RQ2). We therefore also reject H2 and conclude the short-codes are not necessarily helpful for password recall per se, but only a means to enter emojis quickly once users have learned them (e.g. through Slack). Short-codes evoke the same memory burden as regular words. Hence, any memorability benefit through emojis primarily arises from the associations that can be cued even with different renderings. Consequently, we propose to focus on relatable emoji-passwords to boost memorability. This implies that emoji-picker user interfaces should be modified for password-fields by narrowing down the choice to animals, objects and a distinctive set of meaningful symbols.

5.3 Use Pickers as Input Method for Emoji Passwords

We found that a majority of our study participants perceived the graphical point-and-click user interface as suitable for entering emojis on the desktop (RQ3). Thus, if adoption of emoji-passwords becomes more widespread, e.g. because an increasing number of accounts is created on mobile devices (cf. [24]), websites need to ensure users can still authenticate on desktop computers. The picker has certain disadvantages, e.g. the need to download images instead of just using Unicode characters or the loss of motor memory. Yet, we conclude it still provides the best usability in this context, as participants gave it good ratings and were generally more open to using emoji-passwords, if entering them is fast and easy. Thus, H3 could be accepted, if frequent usage of emoji-passwords produces reinforcement effects.

6 FUTURE WORK

There are a few open questions that we plan to look into. First of all, which kind of memory aids or support tools would users employ for emoji passwords? We expect it is trickier to write down an emoji-password than regular PINs or passwords. Also, we try to specify most suitable emojis for passwords. EmojiAuth also tries to maximize security by generating an individual set of emojis for each user. However, this is more difficult on the web and therefore calls for novel solutions. To improve our understanding of password authentication, diary studies in which the participants change some of their current passwords to emoji-passwords could reveal the long-term effects.

7 CONCLUSION

In this paper, we evaluated the use of emojis inside text-based passwords. In our two-part user study, where participants created an emoji-password, we focused on the usability and memorability of this twist on the most common authentication method. We wanted to know if emojis are a suitable addition to make passwords more memorable (RQ1), if memorability is affected by platform-dependent renderings (RQ2), and which method users prefer to enter their emoji-passwords (RQ3). We found that picture-word associations are a unique feature that distinguishes emoji-passwords from similar authentication methods. It was especially these associations that were leveraged by our study participants to help them memorize credentials – a task which most of them accomplished. Also, the associations trump potential troubles to recall the password in case the rendering changes - people could still login even with a different visual rendering of their chosen emoji.

By now, any website that supports Unicode passwords (cf. [31]) already allows emojis, unless they are explicitly blacklisted. There is reason to believe that some users already authenticate this way [8]. The HCI community should therefore reduce pitfalls and facilitate entering emoji-passwords for those eager enough to try them. At this point, we discourage introducing password composition policies that require emojis in order to sign up, because some drawbacks are still not fully solved.

This work paves the way for evaluating the feasibility of emoji-passwords in a broader scope of authentication. While we were able to deliver insights on human behavior and attitudes, work on the guessability of emoji-passwords is necessary. Since it is only a matter of time that malicious password attacks will incorporate emojis in password lists, now is the right time for research in HCI and IT security to make quick progress and mitigate attacks on emoji-passwords.

ACKNOWLEDGMENTS

We would kindly like to thank Christina Schneegaß, Malin Eiband, Axel Hösl, and Franziska Schwamb for proof reading. Also, we really appreciated the OzCHI reviewers' constructive comments and the organizers' support for the final manuscript.

REFERENCES

[1] Mohammed A Fadhil Al-husainy and Raghda Ahmed Malih. 2015. Using Emoji Pictures To Strengthen the Immunity of Passwords Against Attackers. *European Scientific Journal* 11, 30: 153–165.

[2] Apple. 2016. How to use the Touch Bar on your MacBook Pro. Retrieved June 14, 2017 from https://support.apple.com/en-us/HT207055.

[3] Bitkom. 2017. Messenger Report 2016. Retrieved June 11, 2017 from https://www.bitkom.org/Presse/Presseinformation/Zwei-von-drei-Internetnutzern-verwenden-Messenger.html.

[4] Joseph Bonneau. 2012. The Science of Guessing: Analyzing an Anonymized Corpus of 70 Million Passwords. *Proceedings - IEEE Symposium on Security and Privacy*, IEEE Comput. Soc, 538–552.

[5] Joseph Bonneau and Ekaterina Shutova. 2012. Linguistic properties of multi-word passphrases. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 7398 LNCS: 1–12.

[6] Alan S. Brown, Elisabeth Bracken, Sandy Zoccoli, and King Douglas. 2004. Generating and remembering passwords. *Applied Cognitive Psychology* 18, 6: 641–651.

[7] Anupam Das, Joseph Bonneau, Matthew Caesar, Nikita Borisov, and Xf Wang. 2014. The Tangled Web of Password Reuse. *Proceedings of Network and Distributed System Security Symposium (NDSS 14)*, Internet Society, 23–26.

[8] Artiom Dashinsky. 2015. Why you should (not) use Emoji in your passwords. *Medium*. Retrieved June 14, 2017 from https://medium.com/@hvost/why-you-should-not-use-emojis-in-your-passwords-b8db0607e169.

[9] Serge Egelman, Andreas Sotirakopoulos, Ildar Muslukhov, Konstantin Beznosov, and Cormac Herley. 2013. Does My Password Go Up to Eleven?: The Impact of Password Meters on Password Selection. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*, 2379–2388.

[10] emogi Research. 2016. 2016 Emoji Report. Retrieved June 10, 2017 from http://cdn.emogi.com/docs/reports/2016_emoji_report.pdf.

[11] Sascha Fahl, Marian Harbach, Yasemin Acar, and Matthew Smith. 2013. On The Ecological Validity of a Password Study. *Proceedings of the Ninth Symposium on Usable Privacy and Security (SOUPS '13)*, 1–15.

[12] Dinei Florêncio and Cormac Herley. 2007. A Large-Scale Study of Web Password Habits. *Proceedings of the 16th international conference on World Wide Web (WWW '07)*, ACM, 657–665.

[13] Haichang Gao, Xiyang Liu, Ruyi Dai, Sidong Wang, and Xiuling Chang. 2010. Analysis and evaluation of the colorlogin graphical password scheme. *Proceedings of the 5th International Conference on Image and Graphics, ICIG 2009*, 722–727.

[14] Shirley Gaw and Edward Felten. 2005. Reuse and Recycle : Online Password Management. *Extended Abstracts of the Symposium on Usable Privacy and Security (SOUPS '05)*, CMU Usable Privacy and Security Laboratory, 42–43.

[15] Maximilian Golla, Dennis Detering, and Markus Dürmuth. 2017. EmojiAuth : Quantifying the Security of Emoji-based Authentication. *USEC 2017*, Internet Society, 1–13.

[16] Blake Ives, Kenneth R. Walsh, and Helmut Schneider. 2004. The domino effect of password reuse. *Communications of the ACM* 47, 4: 75–78.

[17] Mark Keith, Benjamin Shao, and Paul Steinbart. 2009. A Behavioral Analysis of Passphrase Design and Effectiveness. *Journal of the Association for Information Systems* 10, 2: 63–89.

[18] Patrick Gage Kelley, Saranga Komanduri, Michelle L. Mazurek, et al. 2012. Guess again (and again and again): Measuring password strength by simulating password-cracking algorithms. *Proceedings - IEEE Symposium on Security and Privacy*, 523–537.

[19] Lydia Kraus, Robert Schmidt, Marcel Walch, Florian Schaub, and Sebastian Möller. 2017. On the Use of Emojis in Mobile Authentication. *IFIP Advances in Information and Communication Technology* 502: 1–15.

[20] Kat Krol, Jonathan M Spring, Simon Parkin, and M Angela Sasse. 2016. Towards robust experimental design for user studies in security and privacy. *Learning from Authoritative Security Experiment Results (LASER '16)*, USENIX Association, 21–32.

[21] Cynthia Kuo, Sasha Romanosky, and Lorrie Faith Cranor. 2006. Human Selection of Mnemonic Phrase-Based Passwords. *Proceedings of the second Symposium on Usable Privacy and Security (SOUPS '06)*, 67–78.

[22] Yue Li, Haining Wang, and Kun Sun. 2017. Personal Information in Passwords and Its Security Implications. *IEEE Transactions on Information Forensics and Security* 6013, c: 1–1.

[23] Sanam Ghorbani Lyastani, Sascha Fahl, and Michael Backes. 2016. Improving Password Memorability and Strength Using Mangling Rules. *Symposium on Usable Privacy and Security - Extended Abstracts (SOUPS '16)*, USENIX Association.

[24] William Melicher, Darya Kurilova, Sean M Segreti, et al. 2016. Usability and Security of Text Passwords on Mobile Devices. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 527–539.

[25] William Melicher, Blase Ur, Sean M Segreti, et al. 2016. Fast, Lean, and Accurate: Modeling Password Guessability Using Neural Networks. *Usenix Security*.

[26] Hannah Miller, Daniel Kluver, Jacob Thebault-Spieker, Loren Terveen, and Brent Hecht. 2017. Understanding Emoji Ambiguity in Context: The Role of Text in Emoji-Related Miscommunication. *Icwsm*.

[27] Henning Pohl, Christian Domin, and Michael Rohs. 2017. Beyond Just Text: Semantic Emoji Similarity Modeling to Support Expressive Communication *ACM Transactions on Computer-Human Interaction* 24, 1: 1–42.

[28] Henning Pohl, Dennis Stanke, and Michael Rohs. 2016. EmojiZoom: Emoji Entry via Large Overview Maps *Proceedings of the 18th international conference on Human-computer interaction with mobile devices and services companion - MobileHCI '16*, ACM, 510–517.

[29] Karen Renaud and Antonella De Angeli. 2009. Visual Passwords: Cure-All or Snake Oil? *Communications of the ACM* 52, 12: 135–140.

[30] Shannon Riley. 2006. Password security: what users know and what they actually do. *Usability News* 8, 1: 2833–2836.

[31] Tobias Seitz, Manuel Hartmann, Jakob Pfab, and Samuel Souque. 2017. Do Differences in Password Policies Prevent Password Reuse ? *CHI '17 Extended Abstracts on Human Factors in Computing Systems*, ACM.

[32] Tobias Seitz, Emanuel von Zezschwitz, Stefanie Meitner, and Heinrich Hussmann. 2016. Influencing Self-Selected Passwords Through Suggestions and the Decoy Effect. *Proceedings of the 1st European Workshop on Usable Security*, Internet Society, 2:1-2:7.

[33] Richard Shay, Patrick Gage Kelley, Saranga Komanduri, et al. 2012. Correct Horse Battery Staple. *Proceedings of the Eighth Symposium on Usable Privacy and Security (SOUPS '12)*, ACM, 1–20.

[34] Richard Shay, Blase Ur, Lujo Bauer, et al. 2015. A Spoonful of Sugar? The Impact of Guidance and Feedback on Password-Creation Behavior. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '15)*, ACM, 2903–2912.

[35] Elizabeth Stobert and Robert Biddle. 2013. Memory retrieval and graphical passwords. *Proceedings of the Ninth Symposium on Usable Privacy and Security - SOUPS '13*, ACM, 1.

[36] Anselm Strauss and Juliet M. Corbin. 1990. *Basics of qualitative research: grounded theory procedure and techniques*. SAGE Publications.

[37] Caroline Taggart. 2015. *New Words for Old: Recycling Our Language for the Modern World - Caroline Taggart - Google Books*. Michael O'Mara Books.

[38] Blase Ur, Felicia Alfieri, Maung Aung, et al. 2017. Design and Evaluation of a Data-Driven Password Meter. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*, ACM, 3775–3786.

[39] Blase Ur, Patrick Gage Kelley, Saranga Komanduri, et al. 2012. How Does Your Password Measure Up? The Effect of Strength Meters on Password Creation. *Security'12 Proceedings of the 21st USENIX conference on Security symposium*, 5–16.

[40] Blase Ur, Sean M Segreti, Lujo Bauer, et al. 2015. Measuring Real-World Accuracies and Biases in Modeling Password Guessability. *24th USENIX Security Symposium (USENIX Security 15)*, USENIX Association, 463--481.

[41] Ding Wang. 2016. Targeted Online Password Guessing: An Underestimated Threat. *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security (CCS '16)*, ACM, 1242–1254.

[42] Daniel Lowe Wheeler. 2016. zxcvbn: Low-Budget Password Strength Estimation. *25th USENIX Security Symposium (USENIX Security 16)*, USENIX Association, 157–173.

[43] Susan Wiedenbeck, Jim Waters, Leonardo Sobrado, and Jean-Camille Birget. 2006. Design and evaluation of a shoulder-surfing resistant graphical password scheme. *Proceedings of the working conference on Advanced visual interfaces - AVI '06*: 177.

[44] Weining Yang, Ninghui Li, Omar Chowdhury, Aiping Xiong, and Robert W Proctor. 2016. An Empirical Study of Mnemonic Sentence-based Password Generation Strategies. .