Discussing present-day password solutions: Are graphical passwords the way?

Bianca Burtoiu

Undergraduate Dissertation
Computer Science
School of Informatics
University of Edinburgh

2020

Abstract

Even many years after the introduction of personal computers, passwords still pose a great challenge to the common user. Many potential security options are available, but none offer a perfect universal solution. Primarily designed with security in mind, most password recommendations and policies fail to consider human factors; the weakest link in the online security system is the human himself.

Out of a thorough online security literature review, two widely explored areas of interest resurface: the visually stimulating graphical passwords and the non-memory demanding password managers. It is claimed that images make more memorable passwords, but may not be as secure as text passwords. Password managers relieve users from having to memorize all their passwords, but also force them to give up control over them. The account of advantages and drawbacks of each system hint that one could be used to improve the other.

All points are carefully considered in building a proof of concept for a graphical password tool, which is initially proposed to function alongside password managers to give users more control of their passwords. The concept is evaluated within a group of subject experts.

Critical analysis calls for a new, standalone implementation of the graphical password tool, which is then evaluated in a 139-participant online user study. The results of the study lead to a proposal for a graphical password tool particularly suitable for smartphone devices.

This report confirms previous findings and extends new suggestions for the improvement of password authentication systems, adding to the vast set of studies in the security and usability literature.

Acknowledgements

I would like to thank my supervisor, Robin L. Hill, for all his much valued support and encouragement throughout the project.

I am also very grateful to the 5 subject experts that accepted to take part in my interviews and to the 139 people who filled in the online survey I conducted for this project. Their contributions helped me reach insightful conclusions in the field of password-based authentication.

Table of Contents

Chapter 1

Introduction

1.1 Motivation and objectives

Digital passwords were introduced in the 1960s, alongside the development of the first computer systems. Back then, passwords were simple, used by the few specialists that had access to computers and stringent protection methods were not high-priority concerns. However, since the explosive growth of the Internet in the 1990s, password security has become a key research interest due to the pervasiveness of modern web services and their increasingly critical nature [43]. Nowadays, passwords protect most (if not all) online accounts, ranging from personal communications to financial, corporate or governmental key services.

A lot of password policies and advice is based on theoretical security measures [49] that fail to consider usability requirements. Users are overwhelmed by complex password policies (e.g. at least one uppercase character, one symbol, one digit etc.) and by the increasing number of accounts they need passwords for. Most cope with this effort by reusing the same password across multiple accounts - but experts strongly advise against that [41]. When failing to enforce appropriate security measures, companies lose millions of dollars in data breaches, while individuals are at risk of identity theft. Only in 2019, a total of 15.1 billion records were compromised [15].

Usability then becomes a key factor towards bettering human behaviour with respect to passwords. The official ISO 9241-11 definition of usability is "the extent to which a system, product or service can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use" [13]. This report focuses on the particular usability aspects of *enjoyment*, *ease* of use and *perceived memorability* in password systems; these are some of the key underlying factors behind long-term adoption and sustained security.

In the recent years, the problem of balancing security and usability in user-generated passwords has been discussed in hundreds of pieces of research by experts all over the world. Multiple alternative authentication methods have been proposed; none successful enough to replace the text password. Among the alternatives showing the most promising results in terms of user enjoyment and adoption we find graphical passwords

and password managers, which became key directions of my research.

This report aims to:

- synthesize research conducted so far, exploring the main security and usability challenges and requirements towards the development of password solutions,
- contrast between advantages and disadvantages of existing password solutions, seeking whether there exist aggregate solutions which could perform better than the standalone parts,
- produce a proof of concept for a password solution that meets a certain use case,
- assess its performance on a user population, and
- analyse behaviour patterns in the user study results to confirm or disprove usability claims.

Completing the above aims will lead to the conclusion of which password practices are likely to bring the highest usability benefits, while being realistic to ask of people.

1.2 Report structure

The remainder of the paper is structured as follows: Chapter 2 introduces the topic of password security and discusses key features of text passwords, the most widely used authentication method. Chapters 3 and 4 document my research on graphical passwords and password managers, giving an in-depth account of the relevant previous work. Taking all findings into consideration, I develop a proof of concept for a graphical passwords authentication solution in Chapter 5. A first implementation of the tool, as an overlay on top of a password manager, as well as an initial expert user evaluation follow in Chapter 6. Chapter 7 presents a second, standalone implementation of the tool, as well as its evaluation via a 139-participant online survey. Chapter 8 highlights the main conclusions of the report and proposes future research directions.

Chapter 2

Background

2.1 Introduction

Typical credentials employed for user authentication fall into three categories: *"Something You Know"*, such as passwords or PINs, *"Something You Have"*, such as a token or a card, and *"Something You Are"*, such as biometrics; or combinations thereof [55]. Along with the growth of the Internet, there has been a surge in the number of online services requiring authentication. Passwords, most commonly in the form of text passwords, have since then been the primary barrier to protect users' personal information. Even though alternative authentication methods have been widely explored in the recent years, the deployability and versatility of text passwords kept them on the leading position.

This chapter introduces the ubiquitous text password, examining its fundamental characteristics.

2.2 Text passwords

Text passwords are the most common authentication method available. They are convenient to use, simple to implement and can provide reliable levels of security at low costs.

The more websites people started having accounts for, the more difficult it became for the common user to keep track of their passwords. The simplest solution people found was to use the same password across all their accounts. However, that exposed them to a dangerous vulnerability: a password used on a low-security site - easily compromised by an attacker - may subsequently allow access to a higher-security site [41].

Many of the deficiencies of password authentication systems arise from human memory limitations [80]. In the past two decades, research efforts were concentrated towards educating users to create and use secure passwords.

2.2.1 Common coping strategies

Password creation policies are among the most common forms of guidance provided. The policies have greatly evolved over the past years: in 2004, the United States National Institutes of Standards and Technology (NIST) guideline suggested that passwords must consist of at least eight characters including one uppercase character, one lowercase character, one number and one symbol [66]. Years along the line, in 2011, the famous XKCD comic (see Figure 2.1) persuaded people to use long, rather than complex passwords. The advice formalized much later: NIST only updated their recommendations accordingly in 2017 [6]. In 2020, the latest advice from Google is to use a long series of meaningful yet personal words, difficult to guess by others [51].

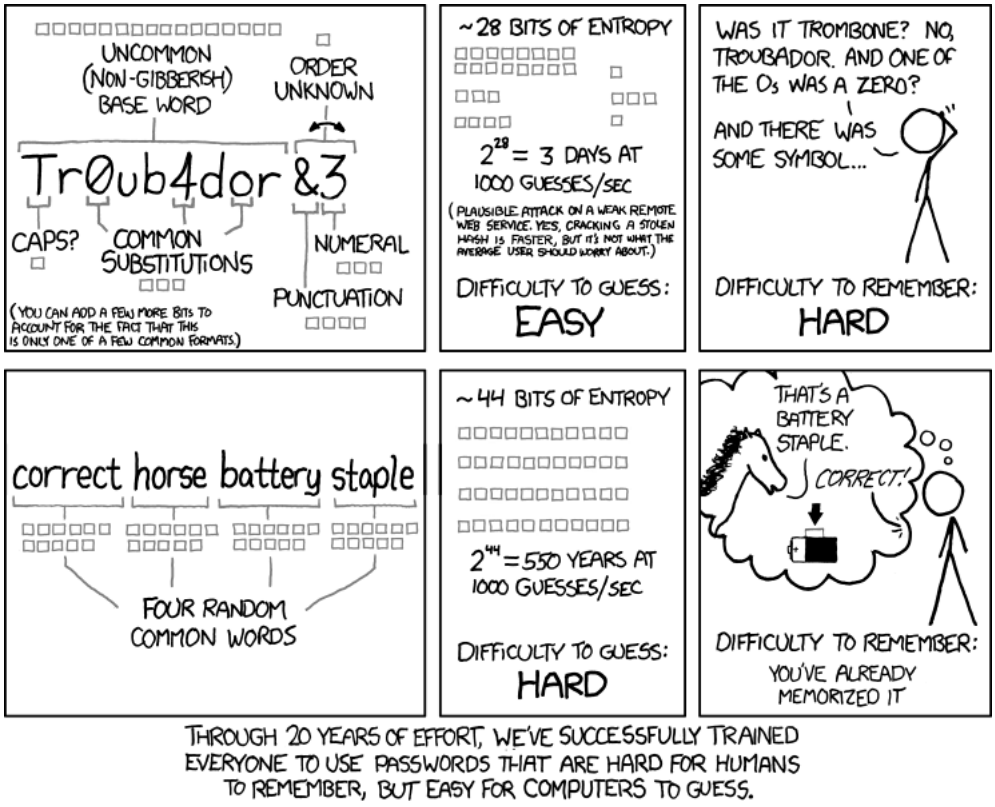


Figure 2.1: XKCD comic (Source: [7])

One study [68] confirms the prevalence of length over complexity by having used 8000 participants' passwords to assess the quality of 8 different password policies against a automated guess-generating algorithm. The two strongest policies imposed passwords at least 16 and, respectively, 20 characters long, with lowercase characters and digits sufficient in term of character space.

Another study [69] evaluates users' reactions ahead of increasingly complex password creation policies: users tend to create new passwords by modifying old ones, and the modifications done are of the least amount of effort required to comply with the new policy, usually by **appending several characters** to an old password. Another study [40] finds that most commonly, these modifications are easily predictable, making a new password fairly easy to deduce from an old one.

2.3 Requirements

The characteristics of text passwords, as well as their effects on users' behaviours can be detailed far beyond the points noted above. Researchers have proposed innumerable alternative solutions, each one of them approaching a subset of the faults of text passwords. Based entirely on previous experience regarding text passwords, we extract the following set of considerations towards the proposal of a new password solution:

1. People learn to use new authentication methods by relating them to their existing knowledge. Any solution must be naturally linked with people's existing habits and behaviours.
2. It should be investigated whether any negative user practices, such as the previously mentioned coping strategies, could be remodeled into positive behaviours. Whenever possible, best security practices should be made default, without any additional user effort.
3. It appears that memorability is tightly coupled with predictability; solutions should aid people enforce true randomness in their password choices while still being able to build meaningful memorization connections.
4. As technology evolves quickly, any solution should be versatile, flexible and not locked into any soon-to-be-obsolete requirements.

These are the key points the reader will find references for throughout my research. The immense popularity of the field of passwords and online security ensured significant breadth of sources to consider in my analysis towards an alternative password solution.

The two most promising research areas I discovered were graphical passwords and password managers; the next chapters will investigate these in depth, looking for ways to meet the stated requirements.

Chapter 3

Graphical passwords

3.1 Introduction

Graphical password schemes most generally refer to an authentication scheme that involves one or more images or sketches, with users having to successfully complete a task related to them to achieve successful authentication. They appeared as an attempt to approach the problem of poor memorability that text passwords posed. Based on the *Picture superiority effect* [65], stating that images are easier to recall than words, it was believed that graphical passwords will be more memorable and therefore, more usable for people.

Multiple psychological studies, some as old as from 1968 [65, 70, 34] support the claim that humans have a significant capability to recognize and recall visual information, and that the brain handles image-based and verbal information with different mechanisms. In graphical passwords, this predisposition is hoped to reduce the memory burden. In other words, the hope is that people would find it easier to create and memorize a *graphical* password than a *text* password.

In a graphical password scheme, instead of having to enter in typical text characters, users must correctly select, for example, an image or a specific location within an image, or correctly reproduce a drawing. The specific selection or sketch then represents their authentication secret. This will be further elaborated on shortly.

Graphical passwords first appeared around 1999, promising improved memorability and strength against guessing attacks [32]. Since then, multiple schemes have been elaborated and evaluated in the research community.

3.2 Existing schemes

Visual memory tasks vary in difficulty depending on the properties of the retrieval process. Graphical passwords can be split into several different categories, based on how they leverage the human memory into producing and retrieving the authentication secret.

3.2.1 Recall-based schemes

In recall-based schemes, users typically produce a password by fully recalling it from memory, with no external support. They are also called "drawmetric" schemes [45] as the password must be sketched on a canvas (either blank or an image). The memory process of recall is known to be a difficult task as it has no supporting prompts, having the heaviest cognitive load among all methods.

Intuitive parallel Text passwords fit in this category, since users are generally presented with a blank password field to fill in.

The Draw-a-Secret scheme [53] was among the first recall-based techniques proposed, where users drew a "squiggle" of their choice using a mouse or stylus. One study [61] has shown that users were likely to use short, predictable sketches. An alteration to the scheme, namely Background Draw-a-Secret [47], aimed to reduce predictability by adding an underlying background image; users were shown to create more complex passwords, but became prone to the risk of choosing image-specific patterns. Both schemes have memorability rates lower than 80%.

3.2.2 Recognition-based schemes

In recognition-based schemes, users create a password by selecting several images from a large set. Then, at the authentication stage, they must correctly select those same images from among additional image decoys to successfully authenticate. Existing recognition-based schemes use several different types of images, including images of faces or objects.

PassFaces [25] is one of the most popular commercially-deployed schemes. Users are shown several portfolios of faces (e.g. four portfolios of nine faces each), with one face per portfolio serving as the authentication secret; correct authentication implies choosing the correct image across the four portfolios [28]. One field study [44] found that participants selected predictable faces, biased on racial and gender preferences; the latest PassFaces software strongly suggests the use of system-assigned portfolios instead of user-chosen.

3.2.3 Cued recall-based schemes

Cued-recall borrows aspects from both recognition and recall into a scheme where users must select specific points or areas from within one or more images to authenticate. In these also called "locimetric" schemes [45], the secret consists of the ordered selection of locations within the image. The image is intended to play an important role in selecting these locations; multiple studies [32, 71] show that people can accurately remember particular details in images. In an ideal system, the details selected would be particularly meaningful to the user, but not to an attacker trying to impersonate them.

In the PassPoints [79] scheme, users create a password by selecting five pixels on a single image to create their password. At the time of retrieval, the points must be selected

(within an acceptable tolerance area for each pixel) in the same order to permit authentication. *Cued Click-Points* [42] is a similar system, where users create a password by selecting one pixel over five sequential images, while *Persuasive Cued Click-Points* [39] extends *Cued-Click Points* to restrict users' pixel selection within a randomly-determined viewport to decrease the predictability of the users' pixel choices. Positive feedback in [28] claims very high memorability levels (up to 94%) for PCCP. A significant risk for these schemes, however, is the predictability of users' choices in password cell selection [37, 44, 75].

3.2.4 Discussion

The three types of schemes differ in terms of security and perceived usability.

While the memory task of recall involves actively reconstructing the information required to authenticate, recognition and cued-recognition only require the decision as to whether that same piece of information has been used before or not [14]. Other previous work [76] similarly supports that recognition is an easier memory task than recall.

In a study [72] contrasting the three different types of memory retrieval on system-assigned graphical passwords, researchers found recognition-based passwords to be more memorable than free-recall passwords. Between cued-recall and recognition, however, cued-recall schemes had a much faster login time.

In terms of security requirements (will be discussed in detail in Section 3.3), the different memory retrieval techniques constrain the total number of possible passwords that could be produced by a certain scheme (in Section 3.3 this notion will be defined as the *password space*). Recognition-based schemes tend to have much smaller password spaces than the other types of schemes: there are much fewer ways to choose one correct image out of multiple decoys, than to draw a sketch with a stylus or to select a single pixel out of an 720x480 image [74].

3.3 Security

Graphical passwords face a varied pool of vulnerabilities due to their visual nature.

3.3.1 Capture attacks

Capture attacks involve the attacker directly obtaining the password (or part of it) by capturing the screen of the device used at the moment of authentication. While in text passwords characters are masked (e.g. with a * or •), in graphical password schemes important visual information must be revealed for authentication. Then, the system is vulnerable to *shoulder-surfing attacks*, where credentials can be captured by direct observation, as well as to *input recording attacks*, where the input of the devices used during the login process (keyboard, screen, mouse) can be recorded by malware installed on the users' machines. In recognition-based schemes, for example, the images

to be selected as part of the password usually are large, discrete units, easy to track in an attack [32]

In most cases, removing the visual feedback comes at a high usability cost, making the tool more difficult (or even impossible) to use. In recall-based schemes, users draw their password sketch on a canvas; if the trace left by the mouse or stylus was transparent, although more secure, users may be less confident and more prone to mistakes from drawing their input blindly. These drawbacks can greatly impact adoption and usability of such tools.

3.3.2 Guessing attacks

A guessing attack implies that the attacker would repeatedly attempt different passwords for a specific account, or against a database of hashed passwords, until gaining access or encountering a match. The guesses can be generated exhaustively or produced in a calculated manner, as we will see below.

3.3.2.1 Password space

In terms of the total number of passwords that can be produced following a certain rule set or policy, password schemes can be evaluated in terms of their *theoretical* and *effective* password space.

The full, theoretical password space comprises *all* possible passwords that can be produced, while the effective password space is the subset of the full theoretical space that users of the system are more likely to choose their passwords from.

Usually, the theoretical password space can easily be calculated mathematically; the effective password space, however, is much more difficult to correctly estimate. The goal of password schemes is to have the size of effective password space as close as possible to the one of the theoretical password space [37].

Intuitive parallel The theoretical password space of 8-character passwords with at least one lowercase, one uppercase, one digit and one symbol is a set of approx. 3 quadrillion (10^{15}) strings [11]. The effective password space, however, is much more limited since some strings, e.g. “7jfJgb5*”, are less memorable than others.

3.3.2.2 Offline and online security

For guessing attacks, the most important aspect is the type of platform the repeated password guessing attempts are conducted on.

In **offline attacks**, the attacker has:

- Access to a database of encrypted (hashed) passwords they try to decrypt, usually obtained via security breaches. Such attacks are not targeted, as usually the goal of the attacker is to break into *any* of the accounts; or

- a very large (or unlimited) number of times they are allowed to try a password. This would happen in the scenario of a poorly-implemented web service, which does not have any lock-out policy in place (lock-out policies will be discussed shortly, in the context of online attacks).

To run a password-guessing hashing algorithm against a set of hashes, an attacker needs sufficient computational resources to approach the task; however, it is known that even with less sophisticated resources and funds available, a very large number of guesses can quickly be produced. For the 3 quadrillion strings mentioned above, [12] states that an Nvidia GTX 1080 8-GPU system can crack a password by iterating through the *entire 8-character password space* in less than 4.2 hours.

To fend against such attacks, password schemes aim to have a password space larger than any attacker would have computational resources to realistically explore. For the same password policy, due to the exponential growth of password spaces as a function of the length, a 16-character password would instead take the same GPU setup *5.7 trillion years* to crack [12]. Moreover, most companies nowadays prioritize investments in top secure storage of their password data.

In an **online attack**, an attacker repeatedly attempts to log into an account for an online service, for which he must know the username beforehand (!). In recent years, countermeasures in the form of **lock-out policies** have been put in place to deter against illegal access; a user is allowed to enter the wrong password a small number of times (usually less than 5) before additional security measures are enforced, such as increasing time between repeated attempts, enforcing an identity check via another platform (e.g. e-mail or SMS) or locking the account temporarily.

Such policies are a very powerful protective measure: they are able to defend the security of accounts even if the password scheme used in the system would not withstand an offline attack. [50] claims that in the case of a 6-digit PIN, if the lock-out policy locks the account for 24 hours after 3 failed attempts, it would take 10 years to cover just 1% of the 1 000 000 possible PINs, and if the policy locks the account even for only 2 hours, it would still take 270 days. Consequently, they vouch against the need for strong passwords so long as a suitable lock-out policy is in place. In 2016, [60] also claimed that well-implemented systems and the use of slow hash functions make offline attacks less likely in practice.

The above lead to the strong conclusion that **in the case of online attacks, highly complex password requirements can be relaxed**. This is a key point of my research, to be further discussed in Chapter 5.

3.3.2.3 Password assignment

Some password schemes impose equality between the theoretical and effective password space by using **system-assigned** passwords: assigning passwords from the entire password space, at random, for users to memorize. Since randomly-assigned passwords are all equally probable, the time required to break one is a function of the total password space, irrespective of whether the attack is online or offline.

The main drawback of this method, however, is that users find it difficult to remember passwords given to them. This is an intuitive remark, also confirmed by a study [67] on the memorability of system-assigned passphrases. Previous work in this field, however, shows promising results; [28] analysed the impact of cues in supporting memorization of system-assigned graphical passwords, finding week-old memorability rates of up to 98%. Assigning passwords to users could, therefore, be considered a viable option if the security of a password system would be insufficient if users created their own passwords.

Yet still, most password schemes available permit **user-chosen** passwords, shaped under a set of policies. Each user has the opportunity to personalize a secret tailored to their own preferences and security needs. Having complete control of the password choices comes with the expectation of good memorability and ease of use. However, it is commonly known that human nature is inherently non-random. This means that users will be more likely to exhibit predictable behaviour, drastically reducing the realistic number of possible passwords and thus the effective password space. This is particularly concerning for offline attacks, but less impactful for the more realistic online attacks, as long as the effective subset of passwords is varied enough.

3.4 Implementation features

3.4.1 Input methods

Platform Graphical passwords are theoretically suitable for use on various screen sizes (e.g. smartphone/tablet/desktop); the images used as part of the authentication process are discrete units that can tolerate resizing to different screen resolutions. Graphical passwords are a topic of interest for authentication on mobile devices, where typing commonly poses difficulties; such alternatives to keyboard entries are becoming increasingly popular [32].

User input Graphical passwords are suitable for touchscreen, as well as non-touchscreen devices: on touchscreen devices, passwords can be input via finger or stylus taps, while on non-touchscreen devices passwords can be input using mouse clicks, keyboard navigation (arrow/tab/enter keys) or keyboard text mappings (each image maps to a keyboard character to be input in a text password field, as seen in [54]).

3.4.2 Accessibility

Graphical and text passwords alike are subject to a number of accessibility issues. This is a topic complex enough to discuss in an entire research paper, however, I shall briefly touch on the key considerations.

Users with *visual* impairments are generally not able to use graphical password tools due to the essential visual element. Furthermore, most recognition- and cued recall-based schemes require reasonable color vision for users to be able to create and identify memorization cues in the image(s) used as part of the authentication mechanism.

Users with *mobility* impairments are also at a disadvantage if they are not able to

control a computer mouse or the full span of a regular keyboard. To adjust for these users, password schemes must enable the users to cover the entire password input space with a limited number of hand or arm movements.

3.4.3 Storage requirements

Another important aspect of graphical password authentication is the discretization and encoding process used to store graphical passwords on the host system and compare them against user entries to provide authentication.

Discretization and encoding In graphical password schemes, the images (or portions of images) representing the user secret are discretized and encoded in text form, then hashed and stored in secure databases, with identifiers to allow the system to easily validate an incoming password against the correct, stored version [32].

Additional information For text password authentication, databases usually simply store key-value credentials pairs *(username, password)*. For a graphical password authentication process, however, additional information must be stored by the system alongside the actual password. In recognition-based schemes, the system must provide a library of images (including decoys) upon every authentication, while also keeping track of the correct images to be chosen (i.e. the actual user password). In cued-recall schemes, the system must have the knowledge of the correct image(s) (i.e. the correct memorization cues) to show each user.

It is important to note that this increased amount of information to be encoded can have costly implications in terms of memory and performance requirements for graphical password schemes.

Chapter 4

Password managers

4.1 Introduction

Another increasingly popular solution to the password administration problem is using a password manager. Although they have only become more widespread in the 2010s, password managers have existed since late 1990s; the Web Confidential [24] password management tool, among the first of its kind, has been released on the Macintosh ecosystem in 1998.

Password managers have a strong advantage of being able to handle an increasing number of accounts. In the past years, the number of accounts per user has been constantly growing: from 7-8 per person in 2006, to 18 in 2013 and over 25 in 2018 [77, 58].

Password managers primarily assist in password storage and retrieval. That implies storing passwords in a secure database and generating them on demand [22]. Their key feature is the *master password* which is used to allow access to the vault and to encrypt the data. They support password generation by producing and storing complex, randomized strings of characters to be used as passwords.

Password managers are widely advertised and recommended by security experts [52, 35]. Despite these efforts, adoption among the wide public has historically stayed low [64]. Password managers have promising capabilities to provide increased security, but cannot perform up to standards if people use them incorrectly. This chapter discusses the key features of password managers, as well as the problems they face in practice, primarily regarding adoption and usability.

4.2 Platform

Password managers exist in various forms, depending on the data storage methods used, the service architecture and the management features available.

4.2.1 Standalone password managers

Standalone password managers are dedicated software programs that act as a database for users' passwords. They can be split in two different categories, depending on the data storage method used:

Local tools store password data exclusively on the device they are installed on. The most common examples are KeePass [21] and Password Safe [23], which provide a user interface to an encrypted database only accessible from the device in question. Even though they were very popular in the 2000s, with the recent increase in the number of personal digital devices owned by each individual, such systems are at a disadvantage in terms of portability.

Intuitive parallel A local password manager is, intuitively, very similar to a physical locked vault.

Cloud-based tools, by contrast, permit access to password data from multiple devices, which makes passwords easily recoverable if access to one of the devices is lost. The companies developing these tools host grand, central servers which store the password data, making it available from all devices a user may own. The benefit of portability, however, comes with the demand for impenetrable security; users must trust the provider with all their passwords. Companies developing the most popular tools, such as Dashlane [17], LastPass [18] and 1Password [16] promise to operate at the latest security standards. Although some password management tools are free to use, the majority require a paid subscription (ranging between £25 and £50 per year [20]).

4.2.2 Built-in password managers

Password managers can also come built into host pieces of software, most commonly, into Internet browsers. Nowadays, the majority, if not all browsers support password management by storing passwords in the cloud.

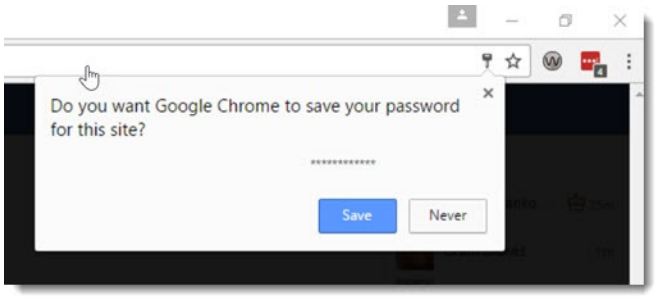


Figure 4.1: Google Chrome password save prompt (Source: [19])

This service provides a consistent user experience throughout all devices, as well as convenience and ease of use. Browsers usually auto-fill personal data fields and prompt to save new passwords (See Figure 4.1), requiring the lowest level of effort from the user. On the other hand, such tools face privacy and security concerns because, once logged in, they have no additional periodic credential verification, and e.g. if a malevolent user gained access to someone's active browsing session, any website for which

credentials were previously saved could be accessed. For Google Chrome, one of the most popular browsers, once the browser is authenticated with a valid Google account, the full list of passwords is accessible in plaintext (!) at passwords.google.com.

4.3 Analysis

Multiple password management tools of the types described above have been commercialized in the past decade; even though they can theoretically solve the security concerns of a substantial number of online users, they have not always been positively received in the user community. Previous work has analysed into the usage patterns of password managers, as well as into users' decision to adopt or reject them. The key points of research are discussed in detail below.

4.3.1 Adoption metrics

Multiple studies investigated into the reasons and motivations behind the low adoption levels password management tools record in practice. Password managers suffer from low adoption rates, especially among non-expert users. In fact, one of the most notable aspects about password managers is how much their success varies between these distinct user populations: users with fair security knowledge appear to be up to three times more likely to use a password manager than less-experienced users; this has been proven twice in a study evaluating users' security behaviours, completed in 2015 [52] and replicated in 2019 [35].

Factors supporting adoption Most studies assessing adoption levels of password managers found *increased security* and *effort amelioration* as the key factors supporting adoption [29]. Users were also motivated by their *convenience* (e.g. in the use of auto-fill) and *usefulness* in handling a large number of passwords [48].

Factors opposing adoption One smartphone password manager adoption study [29] listed *time commitment* as one of the most common reasons users chose against using a password manager. Users complained that password managers require significant training effort, especially at the setup stage. Other common inhibitors were *threat apathy* and *lack of immediacy*: a considerable number of users are unaware of the security risks their accounts are exposed to. These reasons have naturally lead to users' poor motivations towards using password managers.

Another significant category of users noted *trust and control concerns* as significant drawbacks in adoption. This has been confirmed by several reports [30, 48, 81], where users of all types of password managers disliked the idea of having a single service controlling over their passwords, and being vulnerable in a single point of failure in case of a security breach.

4.3.2 Usability metrics

A number of research papers analysed commercial password management tools in terms of their usability. In their published papers, password manager companies make

a number of usability claims which were often invalidated in subsequent user studies. If used correctly, password managers have been proven to increase the security of users' accounts [59]. A system with multiple usability flaws, however, will not be used correctly; in the case of password managers, if users make incorrect use of them, they become vulnerable to significant security risks. The most notable usability flaws of password managers are detailed below.

In existing literature [41, 55], the usability of password management tools was measured by assessing users' behaviour in completing actions related to the primary use cases of password managers:

- correct authentication using a password stored in the password manager,
- saving a new set of credentials into the password manager, and
- modifying an already existing set of credentials in the password manager.

Mental model A user study comparing between two browser password managers (PwdHash and Password Multiplier) [41] found the most significant problems rooting from users having an erroneous mental model of the password manager. Mental models describe "users' reasoning process in interacting with any given system" - more precisely, the actions users believe must to be done to accomplish their goals when using a system.

Norman's Gulf of Evaluation [63] is a measure of the correspondence between the state of the system and users' interpretation of that state, frequently used in Human-Computer Interaction research. In an ideal case, the gulf is small, indicating that the system provides information in a form that is easy to understand and matches the way the user thinks about the system.

In this case, there existed discrepancies between the actual state of the system and users' understanding of it. Users were not able to correctly understand when and how to activate the password managers, nor how long they remained functional once activated. Participants incorrectly assumed the systems would generate and save passwords for all the accounts accessed in one computer session, when in fact, only the credentials for the first accessed account were safely stored. Other users believed the password managers would automatically generate a new password upon each login - in clear contradiction with the fact that only the correct password should grant access to an account.

Inconsistent feedback The aforementioned problems were aggravated by the lack of feedback the tools returned upon user interaction. If users did not receive any confirmation cue for their actions, they wrongly assumed the operation was completed with no error ("I guess that's what's needed to be done" [41]). If the password manager did not give them any cue over whether it was active or not, users tried to apply their previous experience, assigning meanings to unrelated interface elements; one user wrongly considered the "lock" icon next to the website URL in the browser bar signaled that the password manager was running - when in fact, that symbol usually indicates whether the connection is secure.

Lack of trust and control Another frequently encountered problem in password man-

ager usability evaluations [41, 55] was the perceived lack of control; users preferred to enforce their own security measures instead of relinquishing control to an external password management tool. [55] compared between two portable and one online password managers, finding the same predisposition: respondents favoured local solutions specifically because they gave users a better sense of authority over their passwords. On the same note, other previous work investigating into usability properties of password managers [36] states that password manager companies should pay special attention towards increasing users' trust in their security architecture and encryption procedures.

Chapter 5

Graphical passwords tool

5.1 Introduction

Graphical authentication methods and password managers - the two leading areas of my research - were discussed in depth in Chapters 3 and 4. A common aspect of the two is that although they both theoretically have strong security and practicality advantages, their commercial applications suffer in popularity and adoption among common users.

The drawbacks these solutions have encountered in practice are typically related to poor usability, increased complexity and lack of customization in the authentication process. My research contributions focused on proposing solutions to these most prominent issues.

Most attention was paid to improvements over graphical passwords; by carefully considering the strengths and weaknesses of current schemes, I developed a proof of concept for a graphical passwords tool. Led by the overall goal to simplify authentication for Internet users, the solution favors simplicity and flexibility and can adapt to a multitude of contexts and platforms. Password managers are revisited as one of the potential applications of the tool, proposing user effort rationing towards improved online security.

The full set of features and capabilities of the graphical passwords tool, as well as two diverse use case adaptations and their evaluation against two distinct user populations are detailed in the following chapters.

5.2 Background

As discussed in Chapter 3, Section 3.2.4, cued-recall graphical password schemes scored best in usability studies. The behaviour of this proposed solution is inspired from PassPoints [79], a cued-recall authentication mechanism which has shown positive usability results. Some of the practical features of the solution are based on GPEX [31], a password generation plug-in that converts graphical passwords into complex

text passwords, also adapted from PassPoints.

The proposed solution follows a similar direction, exercising a cued-recall memory task: upon a background image of choice, the tool overlays a discrete grid and requires users to select a number of cells to create a visual password.

The tool improves over PassPoints by relaxing security requirements to improve usability metrics. By limiting the attack vector to online attacks only, the system can safely replace the task of selecting pixels to create a password (a usability disadvantage of PassPoints) with the task of selecting clearly delimited cells on an image. The authors of PassPoints also found that users learned how to authenticate using PassPoints passwords fairly quickly and even without any previous experience [79].

While the PassPoints scheme limited its password length to five points (pixels) only, the tool can provide flexible security: the easily customizable password space offers varying levels of security with little modifications to the user interface.

5.3 Description

The authentication scheme uses images and geometrical shapes into a cued-recall scheme, where users select a specific number of cells on top of a grid (See Figure 5.1a) to create a visual password (See Figure 5.1b).

(a) Grid of visual password cells (b) A visual password: selection of 6 cells

Figure 5.1

5.3.1 Background image

The scheme is flexible in the sense that any image could be used under the password cell grid. The purpose of the image is to act as a memory cue in the retrieval of the correct cells making up the password. Previous work [47] shows that the addition of background images in graphical schemes improves the security of the passwords created, as opposed to drawing them on a blank canvas.

Hotspot attacks The images must suit a certain validity requirement. Multiple studies [46, 75] discuss the risk of images whose salient areas are used as cues, leading up to predictable passwords. Rightly so, when creating a new password, people are more

likely to select the eye-catching parts of an image to ease the cognitive load; salience can occur in multiple forms: color, size, placement etc. [79]. Such images are vulnerable to *hotspot* attacks, where malicious actors exploit the most attractive areas of the image first when attempting their illegal guesses.

Intuitive parallel The graphical *hotspot* attack is equivalent to when text password attackers exploit the most frequently used passwords first, such as "password", "123456", "1Monkey!" etc.

Therefore, images should have a uniform probability distribution across the surface, i.e. no parts of the image should be more attractive or likely to be selected than others. In practice, that translates to a picture with an overwhelming number of points of interest.

To illustrate this concept, take the example of Figure 5.2a, where the plate and the food are expected to attract more attention than the grey background. On the other hand, in Figure 5.2b shows overwhelmingly many points of interest through the variety of fruits, with no point clearly more attractive than another.

(a) Image with few salient areas (Image from [1])

(b) Image with many salient areas

Figure 5.2: Non-uniform and uniform image comparison

Memory interference Another practical requirement is that if the tool is used in multiple instances i.e. multiple graphical passwords are creating using this tool, the images chosen as backgrounds for different passwords should differ from each other. Otherwise, the picture superiority effect would be less successful in stimulating the memory to retrieve the corresponding password [71, 62]. The scheme appears superior to text passwords in this respect: if different background images were considered, [38] found that users coped significantly better in remembering multiple PassPoints passwords than in remembering multiple text passwords.

Intuitive parallel If a user has the passwords "alabama42" and "alabama56" for two different websites, they may find it hard to remember which password belonged to which website, as they only differ very little from each other.

The tool should provide a library of uniform, stock images available to use underneath the password grid. To improve usability, the tool should also allow users to upload their own, personal images as background grids, as long as they fit the uniformity requirement described above.

5.3.2 Space discretization

The background image is then discretized through a grid of predefined, clearly delimited click regions. The overlay grid cells can vary in shape and size across the image. Users can choose from a selection of geometrical shapes for the cells, to produce a grid of their liking. The masks can adjust in size to suit various security requirements, by determining a smaller or larger theoretical password space (see Section 5.4.2).

Previous work [79, 39] uses the image pixels as password cells; since it is almost impossible for users to select pixels with such precision, they have also proposed the notion of an adjustable *tolerance* area, and algorithms [79] to ensure that different legal entries of the same password are hashed identically (see Figure 5.3). Their claim against using a visible grid discretization is that it limits the variety of password choice.

Figure 5.3: PassPoints points with tolerance area (Image from [79])

In my work, I decided to follow a different direction: the tool requires the images to be varied, with the more salient areas across the image, the better. This requirement may cause a *paradox of choice*, a Human-Computer Interaction concept [2] which states that having too many choices deteriorates the quality of the decisions made, or even stops the users from making a choice at all. Clearly discretizing the image and imposing clear bounds, therefore, is likely to help towards easier decision-making in password creation. Password cells of various geometrical forms could serve as memorization cues to further aid memorization.

GPEX (the most similar existing work) [31] implemented a 20x20 grid discretization over the original PassPoints functionality; users would have to select five cells atop the grid instead of five pixels to create a password. They compared user behaviours in creating and memorizing graphical passwords using PassPoints vs. their own implementation and returned two key findings. They found that drawing visible grid lines over the image:

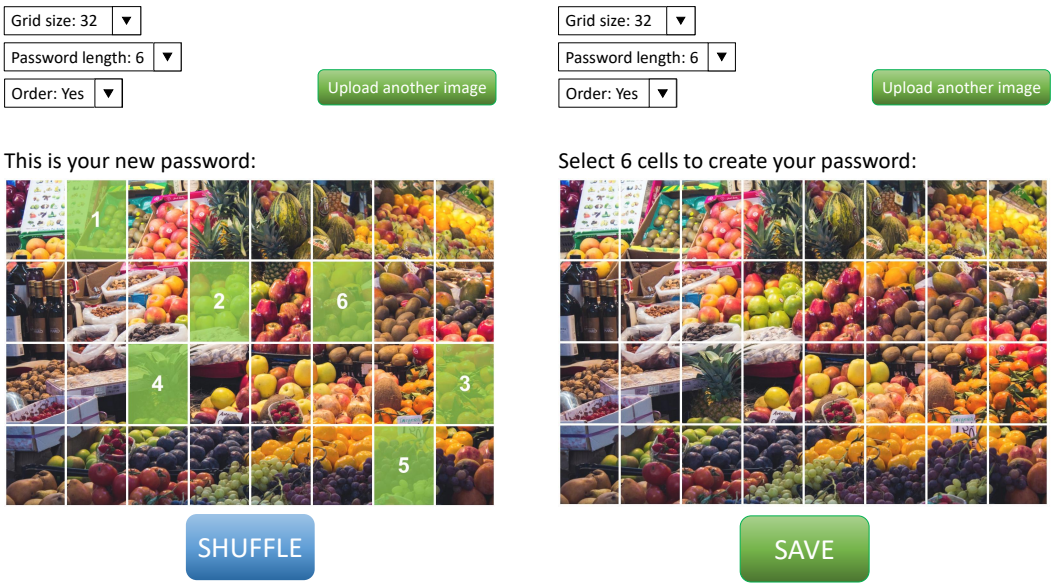
- does not affect usability and user satisfaction, and

- can alleviate the risk of hotspot attacks by enriching the pool of choices at password creation.

These results further support the potential positive outcomes of using a clear discretization of the password space.

5.3.3 Password assignment

The tool could easily support the use of **system-assigned passwords**: for a given image, grid size, ordering and length, a simple randomizing algorithm could be implemented to produce arbitrary passwords. Users could shuffle between combinations, the system continuously rendering random combinations until they find one the user is ready to commit to. Presenting the random password to the users is very simple, by highlighting in green the selected cells, with numbers indicating their ordering (see Figure 5.4a). This step is much facilitated by the clear discretization of space discussed in Section 5.3.2; in previous work [79] communicating system-assigned passwords with pixel cells and a tolerance area was not immediately straightforward to understand by users.



(a) System-assigned password generation (b) User-chosen password generation

Figure 5.4: Comparison: system-assigned and user-chosen password generation

The tool can also allow for **user-created passwords**. In this case, the users select the image of choice, the grid size, the ordering and the combination themselves (see Figure 5.4b.

5.4 Security

5.4.1 Capture attacks

A number of vulnerabilities graphical password tools are generally exposed to have been thoroughly discussed in Section 3.3.1. Some particularities of this tool allow for further attack vectors described below.

Shoulder-surfing attacks As described in Section 3.3.1, visible password feedback can be illicitly obtained through direct observation. In terms of practical implementation, cell selection feedback should normally be hidden when inputting passwords. Moreover, the number of cells expected to be input should not be mentioned either.

Intuitive parallel Just like how text passwords appear as "*****" in the password field, when inputting a password in the graphical tool, the cells would show no feedback as they are being clicked on.

However, this can make it more difficult for the users to be aware of their progress through the password input.

To fend against shoulder-surfing attacks, a scheme combining graphical passwords with text mappings has been proposed [54] where each cell would correspond to a keyboard symbol which would be input in a password field text box, instead of clicking on the password click cells themselves. Although this converts the attack vector into a text password attack vector, the system changes the mappings at every login, resulting in different passwords every time. This could be a potential solution to the visual feedback problem discussed above.

5.4.2 Guessing attacks

As noted in Chapter 3, Section 3.3.2.1, the security of this graphical password tool can be discussed in terms of its *theoretical* and *effective* password space. Each of these corresponds to different decisions over password assignment supported by the tool: *system-assigned* passwords or *user-created* passwords.

For *system-assigned* passwords, the theoretical password space is identical to the effective password space, which is the best security guarantee the tool can provide. For *user-created* passwords, the effective password space is expected to be smaller than the theoretical password space, and, in the best case, as large as the theoretical password space.

This tool can be adjusted to suit for various levels of security through its flexible configuration. The *theoretical* password space of the scheme has several customizable configuration items:

- the total number of password click cells spanning the image (the grid size),
- the length of the password: the number of cells to be chosen, and
- whether, at retrieval time, the password must be input in a particular order or not.

Grid size	Password length	Ordering	Theoretical password space
32	6	Any order	$9.06 * 10^5$
32	6	In order	$6.52 * 10^8$
36	6	Any order	$1.95 * 10^6$
36	6	In order	$1.40 * 10^9$
54	8	Any order	$1.04 * 10^9$
54	8	In order	$4.20 * 10^{13}$
60	8	Any order	$2.56 * 10^9$
60	8	In order	$1.03 * 10^{14}$
100	10	Any order	$1.73 * 10^{13}$

Table 5.1: Password space customizations comparison

From Table 5.1, the reader will note that the larger the password space, the more cognitively demanding the scheme is. This indicates that the scheme is unsuitable against offline attacks, as it may encounter adoption issues due to the considerable memorization effort required.

The problem of building a large password space was, throughout my research, a constant limiting factor towards the proposal of password scheme improvements. However, as discussed in Chapter 3, Section 3.3.2.2, this requirement can be relaxed, which allowed me to investigate more into usability improvements instead.

Important note Therefore, this tool is not designed to withstand offline attacks, yet can be safe against online attacks. The use case for this graphical passwords tool requires the existence of a **suitable lock-out policy in place** on the host system.

5.4.2.1 Cell ordering

The majority, if not all password schemes enforce strict ordering. Most commonly, for successful authentication, passwords require each element (e.g. character) of the secret to be retrieved in the same order as when the password was created. For example, if a text password was “abcdef”, authentication would not be permitted if a string with the same characters, but in a different order (such as “abdcfe”) was input.

In this scheme, however, this condition is relaxed, allowing users to create a password by selecting a set of cells which, for successful authentication,

- must be retrieved in the *same order* they were initially chosen in (serial recall), or
- could be retrieved in *any* order (free recall).

I chose to apply this condition since the idea of unordered recall in password memorization is not very widely explored. The concepts of free and serial recall are widely discussed in psychology studies [56, 78]; although user memorization trials show better results for serial recall, this conclusion is not necessarily transferable to our case. The experimental contexts used in those studies were very different to ours (less than 10 participants memorizing lists of words).

Hypothesis Compared to ordered retrieval, unordered graphical password input can improve ease of use by decreasing the cognitive load.

Although permitting the password cells to be input in *any* order reduces the password space by several orders of magnitude (see Table 5.1), I decided to experiment whether this security reduction would lead to any notable improvements in ease of use and enjoyment.

5.5 Implementation features

5.5.1 Input methods

Platform This tool can be used on multiple different platforms (e.g. smartphone/tablet/desktop) by adjusting the background image and password cell grid to suitable dimensions. The only requirement is for the image to be large enough to permit the overlay of a password cell grid, without the resulting password cells being too small to be clicked or pressed on.

User input This tool supports all common graphical password input methods previously discussed in Chapter 3, Section 3.4.1. Furthermore, since this tool features a clearly discretized password input space, users can navigate through the password grid using their keyboard.

Intuitive parallel Users could navigate through the password cell grid using the keyboard arrow keys and select cells using the "enter" key. Then, a password could be delivered as a set of instructions: "three moves to the right, enter, two moves down, enter etc."

5.5.2 Accessibility

This tool is subject to the common accessibility issues of graphical passwords, previously discussed in Chapter 3, Section 3.4.2.

Users are required to have reasonably good vision (and color vision) to be able to create and identify memorization cues in the background image of the password grid.

Users with poorer motor skills can use the tool due to its clear password space discretization feature. Since the tool permits keyboard input, users could input their passwords by navigating through the password grid using only the arrow / tab / enter keys or a specially adjusted keyboard.

5.5.3 Storage requirements

As previously discussed in Chapter 3, Section 3.4.3, graphical password schemes must be able to encode, store and validate user passwords. The discretization and encoding process transforms a graphical password into a text representation, which is then encrypted for safe storage in a password database.

Discretization and encoding Due to the clear discretization of the input space, each graphical password cell on the grid becomes an atomic piece of information to be encoded. The text representation of the graphical password should correctly indicate the selected and un-selected password grid cells. Figure 5.5 shows a very simple example of such an encoding function, only for demonstration purposes; in practice, a function similar to a cryptographic hash function [5] should be used to produce a complex output. Similarly to common cryptographic hash functions, the encoding function required in this case should be *deterministic* (the same grid and image always result in the same hash) and it should be *infeasible to generate* a grid and image from a given hash.

The text representation is then salted and hashed (an identical process as for regular text passwords [10]), then the output is stored in the database or returned as a text password. Figure 5.5 exemplifies the graphical password encoding, salting and hashing process. The random salt - a string unique for each password generated - could be procured, for example, from the domain name of the website the password is generated for (as in [31]).

Figure 5.5: Graphical password encoding and encryption

Additional information Alongside the encoded representation of the user password, this graphical password tool must store the following additional information:

- the underlying background image
- the correct grid configuration to be placed over the background image: the grid size and shape (e.g. grid of 48 triangular cells, grid of 20 rectangular cells)

Both these items are required for each set of credentials stored; i.e. for each username on the system (e.g. website) using this authentication process. Note that these items must be provided at every login attempt and thus must be quickly accessible by the host system.

5.6 Target users

The characteristics of the target population play a significant role in the design of any password system [32]. It is unfeasible to demand a password solution to fit the requirements of all Internet users, as each population group can have different (and even contradicting) functionality and usability requirements.

The digital proficiency of the users should match with the complexity of the training required to use the tool [32]. Cued-recall graphical password schemes were evaluated as easy to learn [33], with the mental model of "select the correct item to authenticate" being easily comprehensible in practice. The practical implementations discussed in future chapters, however, may require more training effort.

The long-term memorability of the graphical password depends on the quality of the mental cues and associations built when the user is constructing the password; passwords built using very memorable associations should be suitable for infrequent use, while passwords used very often are usually easier to memorize regardless of their complexity.

Chapter 6

Initial application

6.1 Introduction

The graphical passwords tool described in Chapter 5 can be implemented in various contexts to fulfill various functionality requirements. This first application uses the graphical passwords tool to approach some of the most notable usability problems password managers face in practice, previously discussed in Chapter 4, Section 4.3.2.

In this chapter, I elaborate on a proof of concept of a joint graphical password/password manager tool, specifying its key purposes, features and implementation details.

6.2 Description

A problem raised by a significant part of the participants of password manager usability studies [41, 55] was the **lack of trust and control** in the system; a notable part of the population is hesitant to relinquishing all their passwords to a third-party system, in spite of the strong security and privacy guarantees those providers offer.

The solution proposes leveraging memory for increased control: while traditional password managers store the entire password for a certain account, in this case, the password manager would only store a *substring* of the password (an incomplete password), requiring the user to provide the missing substring upon each authentication to produce the complete password (see Figure 6.1).

The immediate question the reader may ask is "*If I still have to memorize part of the password to log in, how does the password manager help me then?*" Indeed, password managers reduce the cognitive load to the maximum by fully relieving users from any memorization effort (except the master password). For this to be a realistic proposal, the remaining substring (that users that must still memorize) must be much easier to memorize than the original password.

This proposed solution brings back **user control** in the sense that if one's password manager gets breached into, the attacker would not be able to gain access to any of the victim's accounts, since the password manager only stores *incomplete* passwords.

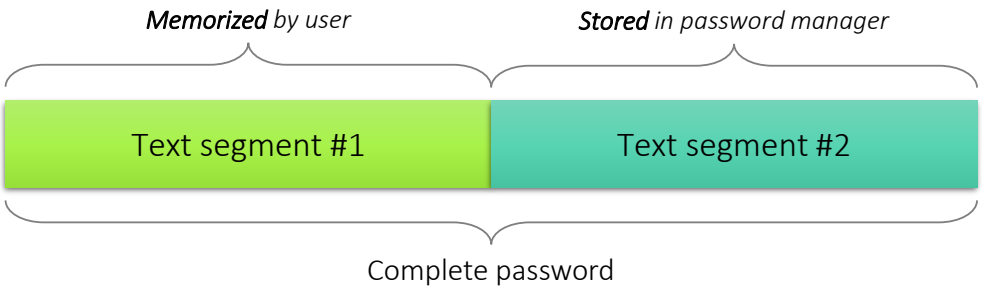


Figure 6.1: User password: segment breakdown. Note: the proportion of the segments is for demonstration purposes, they do not have to be equal in length.

6.2.1 Password segmentation

Memorized password segment The part the user needs to memorize must require a (much) lower memorization effort than the original password. [3] proposes the idea of adding a simple, memorable string (e.g. "bacon") to one's password, *after* the password manager has already auto-filled the text password field. In practice, this is not a feasible solution: if this string is too simple, it can be quickly guessed, resulting in a brittle layer of security. On the other hand, if the string is too complicated, it becomes difficult to memorize and contradicts the purpose of using a password manager in the first place.

We return to the idea of visual information being easier to memorize than words, and use the tool described in Chapter 5 to provide the segment of the password to be memorized (Segment 1 in Figure 6.1). The user has to create one graphical password and memorize that *instead of* the original text password. For each login, the user would input the correct visual password corresponding to that online service to correctly authenticate.

Stored password segment The remaining password segment (Segment 2 in Figure 6.1) has two key characteristics:

- it is stored in the password manager, and
- it is auto-generated.

Although most password managers provide password-generation capabilities, research has shown that the feature is not widely used in practice [81]. In this implementation, when a user registers a new password in the manager, the tool automatically creates a fully randomized, secure text segment, unique for *each* account. We recall that the goal of the system is to offer the user the least memorization effort; this segment is stored in the password manager as it is unfeasible to memorize otherwise.

This feature improves *convenience of use* by implementing a good password manager security practice **by default**, with no additional effort from the user.

Existing password managers were also shown to induce *poor mental models*, as users could not get a clear understanding of how and when their passwords were stored [41]. In this case, however, the concept of concatenating password segments is very

common among users: “appending a couple of characters” to an existing password is a frequently used coping strategy when users want to create new passwords from old [43]. Our concatenation technique follows a similar mental model and is expected to be easily comprehended by users.

Finally, the structure of the password is presented in Figure 6.2. To produce the complete password string, the password manager joins the text encoding of the graphical password (see Chapter 5, Section 5.5.3) with the randomly-generated segment.

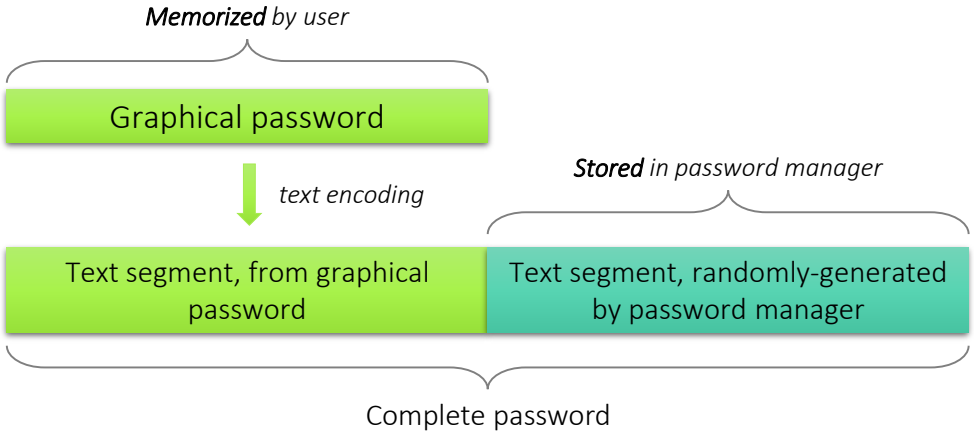


Figure 6.2: User password: segment breakdown

6.2.2 Authentication flow

The authentication process takes place as follows:

1. The user accesses the login page of an online service;
2. The password manager prompts the user with an empty grid configuration, for them to select the correct password cells;
3. The user makes a cell selection to build a graphical password;
4. The password manager validates the graphical password, and, if correct:
5. The password manager concatenates the graphical password user input with the password already stored in the manager to produce the complete password;
6. The password manager pastes the complete password into the password field of the online service.

For Item 5, note that the complete password is only built after the validation step (Item 4). The password manager must execute a validation on the graphical password input as, otherwise, an attacker could mount a replay attack and identify the password segment saved in the manager as the common substring among the failed attempts. The correctness check at Item 4 consists of a comparison between the graphical password provided by the user (Item 3) and the error-checking bits stored in the password manager (see Section 6.4.1).

6.2.3 Categorization

Considering that each password is made unique by the randomly-generated segment, it is possible to further reduce users' cognitive load by preserving the same graphical password to memorize from password to password.

As previously discussed in Chapter 2, Section 2.2, people very frequently resort to reusing their passwords across multiple accounts. By introducing **password categories**, with one graphical password to memorize *per category*, the tool allows users to **safely reuse** the same graphical password for distinct accounts. The reuse is "safe" only because it is *apparent*: internally, the tool combines the common password segment with each randomly-generated segment, producing distinct passwords at the end of the pipeline (see Figure 6.3).

Moreover, by assigning categories or labels to their passwords, users can **reduce** the number of passwords they have to memorize by at least one order of magnitude. That implies reducing the cognitive load from memorizing e.g. 50 individual passwords, to memorizing five different passwords, for five different categories of accounts.

Intuitive parallel Essentially, each password category would have a graphical *master password*, protecting all passwords of that category.

Users are able to group their passwords to completely tailor to their preferences. This enhanced customization property could be used in a number of intuitive ways:

By topic Users could categorize their online password-protected accounts by their purpose, and thus reduce all credentials of accounts on the same theme under a single graphical password to memorize.

Example If a user had online accounts for four coffee shop retailers, e.g. Starbucks, Costa, Caffe Nero and McCafe, to authenticate into *any* of the four accounts, the user would have to input one single graphical password (the one for his *Coffee shops* category) upon logging into either of the four accounts (see Figure 6.3). Based on the account information (e.g. the website requesting the password), then, the password manager would find the unique password substring assigned to that particular account and provide the final password in the password text field.

By priority The categorization feature can also serve as a measure of **effort rationing** in personal password management. Effort rationing - determining the complexity of a password based on the importance the user assigns to it - has been stated to be among users' central concerns in password use [73]. If passwords were to be categorized by importance, users could choose grid combinations requiring lower memorization effort for less important accounts, and more complex grid combinations for more important accounts. To that extent, the reader should recall the multitude of password customization options of the base tool, elaborated in Chapter 5, Section 5.3.3. Importantly, the tool allows the **flexibility and freedom of choice**, by offering a fully optional, non-mandatory added layer of security upon any password already stored in the password manager.

Auto-generated recommendations For users with a very large number of accounts,

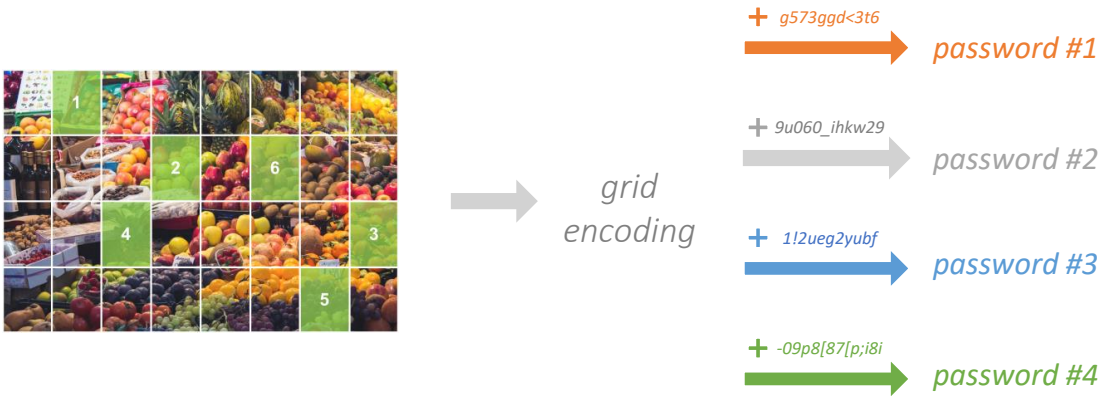


Figure 6.3: Password categorization: producing four different passwords using one grid combination and four randomly-generated strings stored in the password manager

as an extension, the tool could support auxiliary machine learning models or recommender systems to suggest automatic categorizations of users' online accounts.

6.3 Security

One of the key features of the system is **increased user control and security**, achieved by only storing a substring of a user's password in the password manager. Therefore, for a successful attack, a malevolent actor must successfully obtain the two the substrings required to produce the entire password: the password substring stored in the password manager and the graphical password combination, which produces the second password substring.

The only moment when the security of the entire system relies on the security of the graphical password tool is in the case of a password manager breach. If a password manager gets broken into, the attacker is assumed to have obtained the **unique segment** of all the passwords stored in the password manager; therefore, to reproduce an entire password string, he only needs to correctly reproduce the **graphical password** (see Figure 6.4). In that case, the same security considerations of the graphical passwords tool discussed in Chapter 5, Section 5.4 apply.

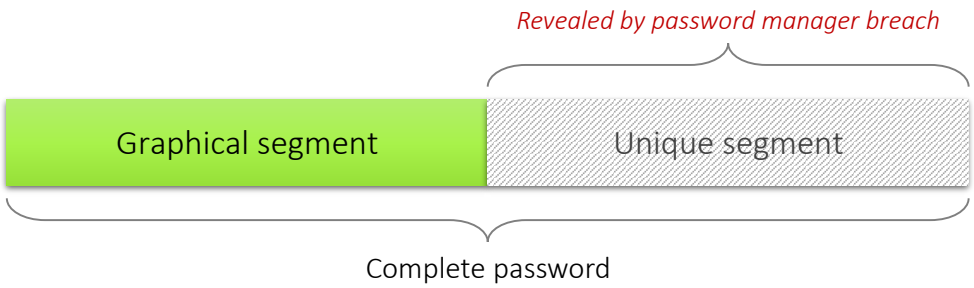


Figure 6.4: User password distribution in password manager breach

6.4 Implementation features

6.4.1 Storage requirements

In this graphical passwords tool/password manager system, the user only has to memorize the graphical passwords corresponding to the password categories they created.

A *set of credentials* is a tuple consisting of a user's *username + the website* the account is on. The password manager supports the user's reduced memorization effort by storing, for each set of credentials:

- the validation bits for the graphical password (the first password segment)
- the full randomly-generated string (the second password segment)
- the category the online service belongs to (if applicable)

For each set of credentials, the password manager must also store the additional information required by the base graphical password tool to render the appropriate empty graphical password grid upon login, previously mentioned in Chapter 5, Section 5.5.3:

- the underlying background image
- the correct grid configuration: the grid size and shape.

6.5 Target users

This joint graphical passwords tool/password manager system was designed to account for the population segment who was not comfortable giving away its passwords to a third-party password management system. As well as increased control, this tool also offers extensive customization features; another suitable population segment is one that would benefit from being able to label and categorize their passwords.

The tool description of this chapter includes a process of password refactoring; to use the tool as the design intended, potential customers must produce a classification of their passwords, then manually re-generate their password for each account (automatic password changes are not currently supported by websites). This is a relatively major initial setup effort the user must commit to. The target population for this tool must be willing to commit to this effort to benefit from the increased security the tool promises.

The possibility to *optionally* add graphical password security on top of a password stored in a manager can also appeal to users who do not necessarily wish to commit to refactoring their entire password collection, but only apply additional security to a number of particular accounts. The target population should be interested in being able to ration their memorization efforts.

Finally, since the tool essentially replaces the task of memorizing a text password with the task of memorizing a graphical one (albeit less complex), the target users should prefer to memorize visual information over text.

6.6 User study

To investigate how subject experts would respond to this proof of concept, I conducted an initial, small-scale study. The findings determined the research direction for the remainder of the report.

6.6.1 Methodology

The user study was designed as a set of semi-structured interviews with Computer Security, Human-Computer Interaction and Psychology domain experts. Semi-structured interviews, as opposed to other types of evaluation, allowed me to guide the discussion towards the key areas of my research, while also encouraging the respondents to pursue alternative discussion paths to reveal further details and expertise.

The study was approved by the Informatics Ethics Committee. The interviews were conducted by the researcher (myself). The interviews were audio recorded to facilitate note-taking. Each interview lasted approximately 45 minutes.

6.6.2 Participants

I referred to the academic staff within the University to find suitable interview candidates. I invited eight Computer Security, Cognitive Science, Human-Computer Interaction and Education experts to take part in the interviews, covering a broad area of expertise. In total, three Computer Security experts and one Cognitive Science expert agreed to take part in the study, as well as one Learning Technologist of a Midlothian high school.

6.6.3 Questions

The interviews were supported by a core set of questions across two focus areas:

1. Questions on general online security, primarily looking to identify differences in behaviour between experts and non-experts: the user categories most vulnerable to poor password security and the security considerations likely to be employed by experts, but not by lay users;

2. Questions on password managers: expert opinions over adoption, usability, security concerns, password generation features;

Afterwards, I explained and demonstrated the proof of concept for the joint graphical password/password manager tool, leading an open discussion over its advantages, disadvantages and potential improvements. The prototype was presented using several supporting illustrations.

Due to the semi-structured nature of the interview, each session slightly differed from the others; the experts were invited to elaborate on their answers, often revealing alternative discussion paths to be explored. To account for the experts' diverse areas of expertise, I slightly adjusted the focus of the question set from one interview to another.

No additional demographics information was collected from the participants; the only relevant population characteristic for the study was the participants' field of expertise, which was also the basis upon which they were invited to take part in the study.

The core interview script can be found in Appendix C.

6.6.4 Results

The interviews yielded a qualitative dataset which was analysed using an informal coding strategy, extracting the most notable remarks across the interviews. The coding was conducted over the written transcripts[1] produced from the audio recordings of the interviews. The key findings for each of the interview focus areas are detailed below. I referred to the transcripts to provide exact quotes from experts' answers.

6.6.4.1 General online security

Experts named a broad set of user categories prone to poor password security: users with an unmanageable number of accounts and casual users alike are equally likely to neglect their security measures. Experts also mentioned vulnerable categories (users with mental health problems and other disabilities, children) as potential target demographics in the development of new password administration solutions.

6.6.4.2 Password managers

The subject experts confirmed previous findings on adoption and usability of password managers. They noted that a significant part of users do not have a basic understanding of the data flow and security principles enforced by password managers. Users were said to have a "limited complexity budget" and are usually not willing to make additional cognitive efforts, relying on whatever security method the systems they use provide by default.

6.6.4.3 Prototype

Experts overall raised several positive, as well as negative points on the functionality and usability of the proof of concept.

Supporting arguments Experts commonly agreed on the fact that the tool was enforcing good security behaviours by auto-generating passwords by default:

> "If you make the secure option the default, the users will use it." (Expert 5)

They also supported the idea of effort rationing, acknowledging that less important accounts can have less secure passwords; however, they did not clearly admit that categorization was the best means to ration effort.

[1] The complete transcripts can be found in the Project Resources directory.

Opposing arguments On the other hand, the experts raised a significant number of usability concerns. One interviewee feared the complexity of the application and the training effort would discourage users from using it:

> "I worry that, actually, for a random user, they might not understand all of that and how it works." (Expert 2)

> "The questions of 'Who stores what' and 'Which part of the password is stored where' is quite complex. A lot would have to go into how you communicate it!" (Expert 2)

Experts also believed users would not immediately realise the benefits of using the tool. More severely, they worried that the system nullifies the advantages of using a password manager in the first place:

> "One of the main benefits (of password managers) is that there is only one password to remember, so, why now start introducing more?" (Expert 1)

> "It does seem like you're removing one of the main advantages of the password managers, which is to remove this cognitive load of managing all these passwords and remembering them [...] even if I could just about manage it, it sounds quite heavy..." (Expert 1)

The experts acknowledged the range of security concerns behind the decision of adding (optional) additional security to passwords, even at the cost of memorizing more than just the master password required by the password manager. While they agreed it might appeal to a certain population niche, they worried its benefits could not be generalized to a greater user population:

> "There exist people who [...] could clearly see how this system is better, it's not much more work so they prefer it. But it may also be that that number of people is actually quite small..." (Expert 2)

> "If you're only going to provide a benefit for a small number of people who are already fairly well protected..." (Expert 2)

6.7 Conclusion

The purpose of the user study was to supplement my research with the advice and opinions of security experts within the University. Their reactions towards my proposal had a negative tendency; the most notable concern was that adding the graphical passwords tool atop a password manager would diminish the password manager's efficiency instead of enhancing it.

Fortunately, this critical result was discovered early in the research process; the results motivated me to elaborate and propose a second application for the graphical password tool, this time without associating it with a password manager. The following chapter will describe the solution in detail.

Chapter 7

A second application

7.1 Introduction

Based on the results of the critical analysis conducted in Chapter 6, I decided to elaborate on a second implementation of the graphical passwords tool described in Chapter 5. While in Chapter 6 the graphical passwords tool was introduced alongside a password manager, this chapter will describe its implementation as a system that could easily be integrated into existing online authentication flows.

7.2 Description

In this second implementation, the use case of the graphical password tool changes from representing an incomplete segment of a password (as in Chapter 6) to comprising an entire password instead.

7.2.1 System integration

Host system Integrated within a host system (e.g. website or application), the tool can function as a sole authentication step or be part of a multi-step authentication process. We have seen in Chapter 5 that the tool, used on its own, can successfully protect against online attacks if a lock-out policy is put in place. Another possible application of the tool could be as part of a multi-step authentication procedure; for example, as a verification step when logging in from a new device. The host system must be adjusted to support graphical password encoding and storage, as elaborated in Chapter 5, Section 5.5.3.

Independent software The tool could also function as an independent piece of software (e.g. mobile or desktop application) or as a browser extension. In this case, the tool must return the deterministic text encoding of the graphical password (see Chapter 5, Section 5.5.3) and use that to fill the password input field of a login form (see Figure 7.1).

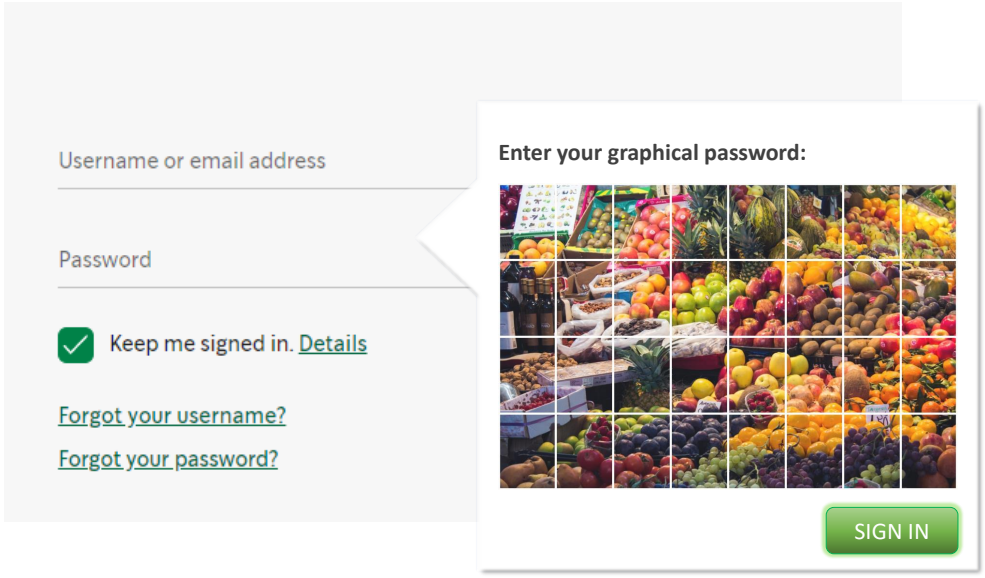


Figure 7.1: Tool as a browser extension: user login prompt (Login form from [4])

7.3 Security

In either of the implementations described above, the tool provides a layer of *adjustable* security. [32] classifies password schemes into two levels of theoretical security:

- PIN-level: password space of 12-15 bits; i.e. in the order of 10 000 (approx. 2^{12}) possible passwords,
- password-level: password space of 30-60 bits; i.e. in the order of 1 billion (approx. 2^{30}) possible passwords.

The tool can suit either security level by simply adjusting the grid configuration parameters. For example, a password grid of 15 cells out of which 4 must be chosen in order has 15-bit security (see Equation 1), while a password grid of 54 cells out of which 8 must be chosen in any order has 29-bit security (see Equation 2).

$$\text{n} = 15, \text{k} = 4; \quad {}^{n}P_k = \frac{n!}{(n-k)!} = 32\,760; \quad \log_2(32\,760) \approx 14.99 \; (Equation\ 1)$$

$${}^{n}C_k = \frac{n!}{k!(n-k)!} = \binom{54}{8} = 1\,040\,465\,790; \quad \log_2(1\,040\,465\,790) \approx 29.95 \; (Equation\ 2)$$

7.4 Target users

As opposed to the system discussed in the previous chapter, which was designed to suit a specific population, the target demographic for this system is much less constrained; only the population characteristics discussed in Chapter 3, Section 5.6 apply.

7.5 User study

To investigate how real users would respond to this proof of concept, I decided to conduct a large-scale user study. I produced a graphical password tool prototype and evaluated its performance in several experimental conditions in terms of memorability, enjoyment and ease of use.

7.5.1 Methodology

For the purposes of the study, I concluded that a web-based, online survey would be the best means of interacting with the target population to gather relevant opinions. The study was approved by the Informatics Ethics Committee.

Study type In the past years, due to the increasing accessibility of the Internet on a multitude of devices, web-based studies grew in popularity. The main advantages of web-based studies, as per [32], are that a large number of participants can be recruited, the participant pool recruited is likely to be more diverse than in another type of study and participants may behave more naturally than in a lab-based study since the study does not disrupt their natural environment.

However, web-based studies do not come without their concerns. It can be difficult to verify the correctness of the demographics information provided. Collected data may still not truly reflect participants' behaviours and opinions, although it is less likely to happen than in the "forced" environment of a lab study, where participants are particularly aware their behaviour is investigated into. Online surveys also pose the risk of participants leaving midway through, case where an unfinished response cannot be used in the analysis. Here I strove to find the balance between capturing enough relevant data from the participants without getting them to drop out because of having to answer too many questions.

Study style The study was conducted *within* subjects, with each of the participants exposed to all experiment conditions.

Implementation platform The survey was designed and ran on the Qualtrics platform [27] for which the University provides accounts with premium subscriptions. The platform supports survey design, as well as data hosting and report management for the recorded replies.

7.5.2 Participants

The recruitment pool was intentionally very loosely constrained, to capture as diverse a set of opinions as possible: any participant irrespective of age, gender or level of education was eligible to take part. Due to Informatics Ethics Committee requirements, however, participants had to respond from within the United Kingdom only.

Survey participants were recruited via snowball sampling: I advertised the survey link on the Informatics students mailing list[1] (accessing 2689 subscribers[2]) and during the

[1] The mass invite email was sent to the e-mail address students@inf.ed.ac.uk

[2] Count from https://lists.inf.ed.ac.uk/mailman/roster/students

University Project Feedback Day event, while the project supervisor posted it on his personal Twitter account. The survey was active for a period of five weeks (February/March 2020).

7.5.3 Questions

The survey is split into three different sections, as follows:

7.5.3.1 Initial questions

The first part of the survey captures background information of the participants, framing their current online security habits and perspectives:

- *demographics information*: age and level of education,
- *familiarity with technology*: hours spent online every week, typical online activities conducted every week, and
- *password management habits*: current number of password-protected online accounts, number of different passwords used, techniques used to keep track of passwords, considerations when creating passwords and security perception of current habits.

7.5.3.2 Experiments

The experimental part of the survey consists of four password creation "micro-tasks", covering four different password configurations. Each task has two parts:

1. Password **creation**: Following the instructions, the user creates a graphical password by selecting a number of cells on the grid
2. Password **confirmation**: Immediately afterwards, the user selects the same cells on the grid to confirm his password selection

The two parts have been chosen to mimic the account creation process of any regular online service (e.g. website, mobile application).

Scenario Many password-related user studies place the password creation activity in some form of scenario to simulate the desired mindset among the participants [49]. I followed the same principle, stressing the importance of not divulging any real passwords, while at the same time advising towards creating realistic, memorable passwords. The scenario is given below:

You have just created a new e-mail account for which you need to come up with a secure password. In this experiment, you will create a visual password.

You will be asked to create a password in four different ways, following four different policies - essentially, you will end up creating four different passwords. For each experiment, please treat the procedure as realistically as possible - create a password that you would be able to memorize.

Important: The passwords you will create will be saved and analysed, so please do not submit a password that you may already be using somewhere else.

Research variables The experiment conditions were designed in terms of dependent and independent variables. Recall the definitions of dependent and independent variables in research (definitions from [8]):

1. An **independent variable** is the variable that is changed or controlled in a scientific experiment to test the effects on the dependent variable.
2. A **dependent variable** is the variable being tested and measured in a scientific experiment.

As the experimenter changes the independent variable, the effect on the dependent variable is observed and recorded [8]. The experiments exercised four experimental conditions in two independent variables: password retrieval order and cell shape (see Table 7.1).

	In order	Any order
Triangles	Experiment 1	Experiment 3
Rectagles	Experiment 2	Experiment 4

Table 7.1: The four experimental conditions exercised in the user study

Background image I chose a stock background image from [26] (see Figure 7.2). Even though the graphical passwords tool supports (and encourages) the use of personal images, in this experiment, the background image is a **fixed variable**; otherwise, data would not be comparable between subjects. The image was chosen to be uniform across its entire surface, with a multitude of salient points.

Figure 7.2: Grid background image

Independent variable 1: Retrieval order The first independent variable - ordering of the password cells - evaluates whether different retrieval conditions perform differently

in terms of user enjoyment, memorability and ease of use. In Chapter 5, Section 5.4.2.1 I described that the tool can support two distinct memory processes: serial (in order password retrieval) and free (any-order retrieval) recall.

The retrieval order directly influences the password space, and thus the security of the passwords created; therefore, the theoretical security of the two experimental conditions had to be a **fixed variable**. This means that all the four passwords users would create in their experiments are **equally secure**. Note that the shape of the cells does not influence the security of the passwords. This has been achieved by using two grid sizes:

1. 32 shapes, out of which 6 must be chosen in order (permutations without repetition), and
2. 54 shapes, out of which 8 must be chosen in any order (combinations without repetition).

In both schemes, there are approx. 2^{30} (1 billion) possible ways to create a password. The reader must recall the tool need not withstand offline attacks, however, this configuration produces a moderately secure theoretical password space.

Independent variable 2: Cell shape The two geometric shapes (rectangles and triangles) were chosen to evaluate whether the appearance of the grid had any usability impact. The experiments exercise rectangular grids (a very common discretization form, previously used in related work [31]) and triangular grids (for their aesthetic appearance).

The password grids for the four experiments are shown in Figures 7.3, 7.4.

Experiments 1 and 2: *Create a password using 6 cells, for which you **must remember** the order they were selected in.* (See Figures 7.3a and 7.3b)

(a) Experiment 1: Rectangles (b) Experiment 2: Triangles

Figure 7.3: Experiments: in order retrieval

Experiments 3 and 4: *Create a password using 8 cells, for which you **do not have to remember** the order they were selected in.* (See Figures 7.4a and 7.4b)

Table 7.2 summarizes the research variables in the experiment.

(a) Experiment 3: Rectangles (b) Experiment 4: Triangles

Figure 7.4: Experiments: any order retrieval

Fixed variables	Independent variables	Dependent variables
Background image Theoretical password space	Retrieval order Cell space	Enjoyment Ease of use Predicted memorablity

Table 7.2: Research variables

Post-experiment questions At the end of each experiment, participants had to answer three short questions over their enjoyment of the creation process, ease of input and perceived memorability ("*Would you be able to remember this password in a week's time?*").

7.5.3.3 Final questions

After completing the four experiments, participants were asked to answer several questions on:

- comparing the *enjoyment of the creation process* between the four experiments,
- comparing the *ease of input* between the four experiments,
- comparing the *ease of memorization* between the four experiments,
- comparing the *perceived security* between the four experiments,
- preference over *ordered vs. non-ordered* cell selection,
- preference between *geometrical shapes*,
- perceived *use of the background image* during password creation,
- preference to use *own* background image instead of a stock one,
- perceptions of *ease of memorization in comparison with* text-based passwords,
- *importance of accounts* the scheme would be trusted to produce passwords for,
- preference to *use on a daily basis*.

All questions were formatted as multiple choice, on a 5-point Likert scale (*Strongly agree* to *Strongly disagree*).

The full questionnaire (except the Experiments) can be found in Appendix F.

7.5.4 Study limitations

7.5.4.1 Population

The user study population was recruited via snowball sampling as a limited number of advertisement platforms were considered; this may lead to the participant set showing similar characteristics (age, technology use, security practices).

7.5.4.2 Accessibility

It is unknown whether participants had particular accessibility requirements; this was not recorded as it was not among the purposes of the user study.

The tool prototype used in the experiment was developed under the limitations of the Qualtrics survey platform. Chapter 5, Section 5.5.1 elaborated on the input methods supported by the tool: finger/stylus taps, mouse clicks and keyboard controls. In a fully correct implementation, the prototype must provide keyboard navigation through the password grid. Unfortunately, the Qualtrics platform does not support this functionality and thus desktop respondents were only able to create passwords using mouse clicks, while mobile respondents could use fingers/stylus.

7.5.4.3 Ecological validity

The ecological validity of a study refers to the ability to design the study conditions in a manner so that its results are realistic and transferable to the real world. This property is particularly challenging to achieve for password studies involving users, which are aware they are creating a password *for a study* and not for an account they value and access repeatedly over time [57]. Ecological validity measures in web-based studies include being held in participants' natural environment and introducing realistic scenarios [32]. I designed the study to fulfill both measures, achieving a level of reliability sufficient to propose valid research claims; despite this, I cannot conclude that my results completely describe real-world user behaviours and opinions.

7.6 Results

7.6.1 Initial questionnaire results

See full breakdown of answers at Appendix G.

Participants A total of 163 participants responded to the survey. Some participants left the survey midway through; after removing all incomplete replies, I was left with 139 usable responses.

Demographics Out of all respondents, 81% were of ages 18-24, with 16% between 25-34 and 4% over 35 years old. In terms of the level of education, 62% of them completed (or are due to complete) a Bachelor's degree, while 38% completed (or are due to complete) a Master's degree or a Doctorate.

These figures suggest that the majority of survey participants are young, highly-educated adults, currently enrolled in or having completed higher education.

Familiarity with technology 52% of the participants interviewed spend 5 hours or less per day on the Internet, while the other 48% spend more than 5 hours. The survey enlisted 14 common online activities (e.g. social media, e-mail, online banking etc.) to query the variety of tasks participants do online and thus measure their level of familiarity with common online tools and platforms; 60% of participants claimed they usually conduct 9 or more out of the 14 listed activities every week.

The respondents appear to be extensive Web users; even though they are not representative of the entire population, their security behaviours are particularly important to understand when shaping future technology developments [30].

Current password management habits The subsequent questions evaluated the password management habits of the participants. To be able to draw relevant comparisons against existing research, I posed questions similar to those used in a 2017 Digital Guardian password security habits survey [58] conducted on 1000 US participants aged 18 and up.

In terms of the number of online accounts currently used, the majority (56%) of participants claimed they have "too many accounts to count", 26% of them replied with "less than 25" and 18% with "between 25 and 50". This result was completely expected; password overload has been officially flagged a problem by NCSC [9]. The Digital Guardian survey similarly recorded that 30% of respondents had too many accounts to handle.

When asked to self-report the quality of their passwords, 8% of participants reuse the same password across all their accounts, 51% alternate between a small number of distinct passwords and 41% of them have a different password for each account. Note that almost half of all participants selected the latter option, which experts recommend, while there still exists a small proportion of participants that always reuse passwords.

To keep track of their passwords, 54% of participants use a password manager, while 46% rely on personally memorizing them. Both categories use these techniques alongside writing passwords down in a safe location, or deducing them from a logical scheme they created in their own minds. In the 2017 survey [58], only 28% of respondents claimed to use password managers; the higher percentage recorded in this case may be due to the web and technology keenness of the respondents.

When it comes to their priorities when creating a password, slightly more than half of participants (56%) value security over convenience. In terms of participants' own perception of their personal online security, 69% are feeling *Extremely* or *Somewhat secure* with their current password habits, while 31% are uncertain (*Neither secure nor insecure*) or less confident (*Somewhat insecure*). Based on their answers on the

previous password management questions, their perceptions seem to be realistic and not over-confident: they indeed seem to be exercising a fairly secure behaviour. 41% of them have distinct passwords for each account, and more than half (54%) are using password managers; this shows that respondents are generally mindful of the most recent security recommendations and practices.

7.6.2 Experiment results

See full breakdown of answers at Appendix H.

Survey completion platform The metadata information captured indicates that that 58.2% of the respondents used a *desktop* device to complete the survey, while 42.8% of them used a *mobile* (touchscreen) device. This is a particularly important distinction as it distinguishes between two user input forms: mouse vs finger/stylus screen presses, which can lead to significant differences in perceived usability.

7.6.2.1 Passwords generated

To analyse the passwords created in each experiment, I produced *heatmaps* from the total frequency counts recorded at the password **creation** steps. The frequency count in this context indicates the number of times a particular cell was selected to be part of a password, across all the 139 passwords created (for each experiment). If a cell has a high frequency count it suggests that a large number of people used it as part of their passwords.

Intuitive parallel: A cell with high frequency would be equivalent of using the popular sequence "123" in a text password.

The heatmaps used in this analysis are color-coded in four color levels ranging from light yellow (low frequency count) to dark red (high frequency count). The color scale is relative to the counts in each heatmap.

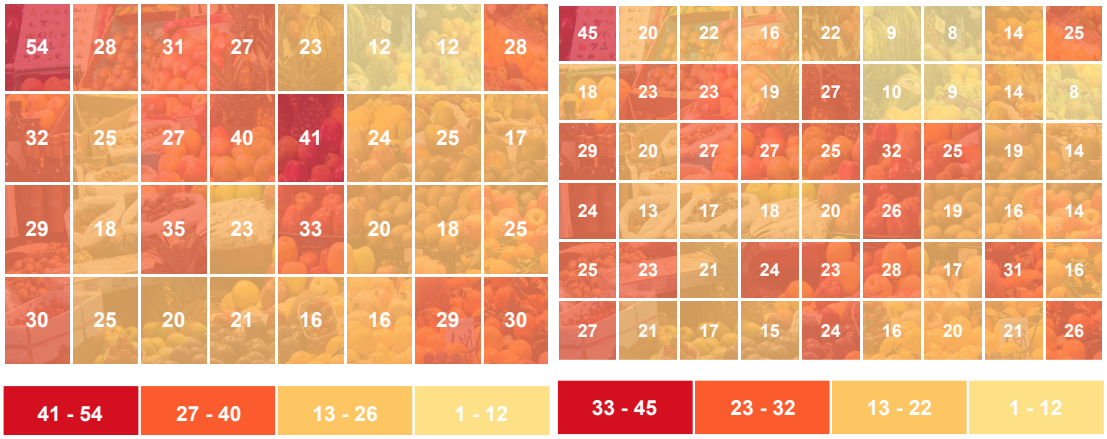
The heatmaps for experiments 1 and 3 (rectangles, in order vs. non-order) are shown in Figure 7.5[3].

Metrics	Ordered (Figure 7.5a) Total cells: 32	Non-ordered (Figure 7.5b) Total cells: 54
Minimum frequency	12	8
Maximum frequency	54	45
Average frequency	26.0	20.6
Standard deviation	8.6	6.7

Table 7.3: Metrics for heatmaps in Figure 7.5

The metrics in Table 7.3 show that the background image was relatively unsuccessful in producing a uniform probability distribution over the grid. The difference between

[3]Note that all figures can be zoomed in for more detail.



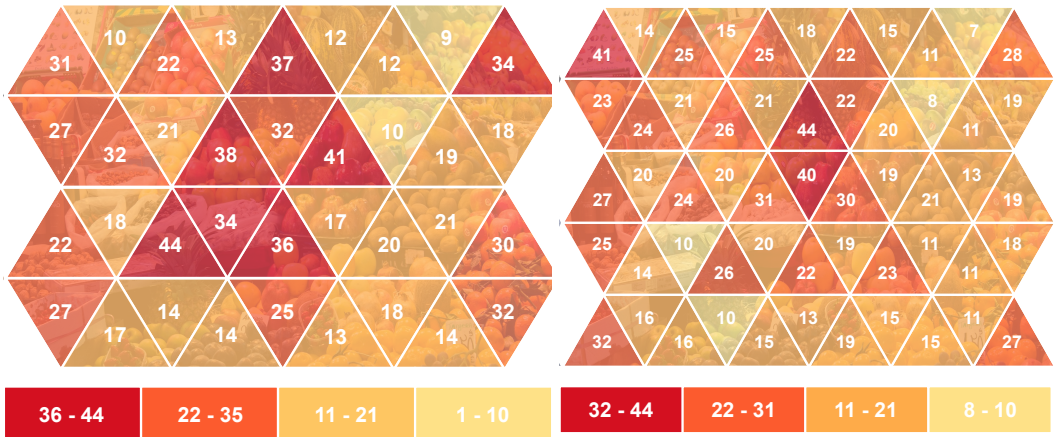
(a) Exp. 1: *Choose 6 rectangles in order* (b) Exp. 3: *Choose 8 rectangles in any order*

Figure 7.5: Heatmaps for experiments 1 and 3

minimum and maximum frequency is notable in both cases (12 vs. 54 and 8 vs. 45), indicating that the number of times cells were used vary a lot between each other.

It appears that the frequencies are distributed very similarly irrespective of the grid size. Both rectangular grids show higher interest in the top-left corner, as well as in the top-central area. The least popular cells are in the top-right area (but not the corner cell!) in both grids. Both grids show increased interest for the leftmost column, and not as much for the rightmost column.

The heatmaps for experiments 2 and 4 (triangles, in order vs non-order) are shown in Figure 7.6.



(a) Exp. 2: *Choose 6 triangles in order* (b) Exp. 4: *Choose 8 triangles in any order*

Figure 7.6: Heatmaps for experiments 2 and 4

Similarly as for the rectangles, there exists a notable difference between the minimum and maximum cell frequencies (9 vs. 44 and 7 vs. 44) (see Table 7.4) showing that cells vary a lot in their popularity.

For the triangular grid, there seems to occur less of a behavior transfer between the two

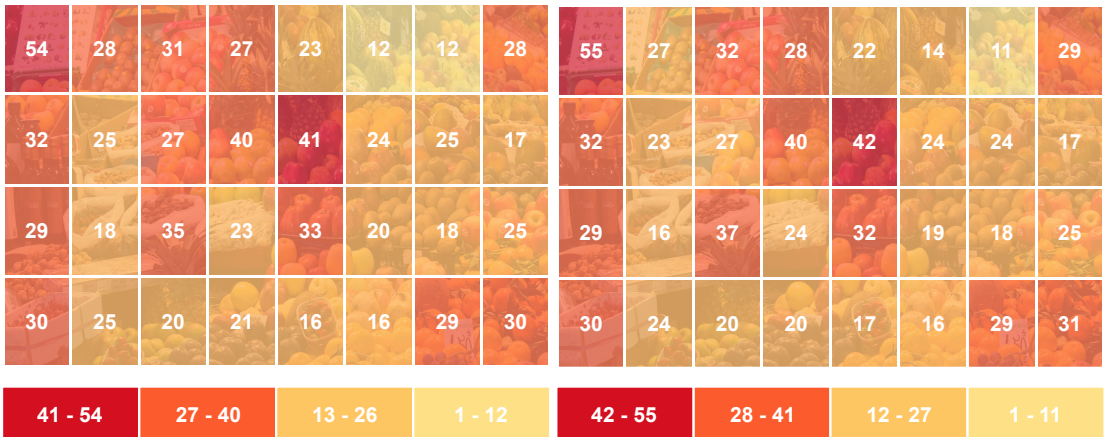
Metrics	Ordered (Figure 7.6a) Total cells: 36	Non-ordered (Figure 7.6b) Total cells: 55
Minimum frequency	9	7
Maximum frequency	44	44
Average frequency	23.1	20.2
Standard deviation	9.7	7.8

Table 7.4: Metrics for heatmaps in Figure 7.6

grid sizes. For the smaller grid, we have a strong middle focal point, as well as vertical edges and corners. On the larger grid we have many hits in corners, as well as in the top-central part of the grid.

A rather odd common fact across all the four grids is the very low-frequency area in the top right part of the image (but not including the corner!).

Patterns The spatial placement of the points suggests that participants likely made use of patterns to create their passwords. Since the Qualtrics platform does not support the particular form of input validation required, passwords input at the creation and confirmation steps were not checked for equality. Consequently, there exist slight differences in the passwords produced at the two steps. These differences further prove the manifestation of patterns: some of the areas of focus are shifted to the right (see Figure 7.7). The fact that participants went one-off at the confirmation step reveals that they focused more on patterns than on the image itself.



(a) Exp. 1: *Choose 6 rectangles in order* (Creation)

(b) Exp. 1: *Choose 6 rectangles in order* (Confirmation)

Figure 7.7: Heatmaps for Experiment 1: Creation vs. Confirmation step

Predictable patterns are, indeed, a known risk click-based passwords are exposed to [37]. If users believe patterns are the easiest strategy in creating these passwords, it indicates that the system facilitates insecure behaviour. This is a significant drawback in terms of security. Results in usability (discussed later on) then become very important to assess the overall quality of the system.

Conclusion Besides the hotspots most likely generated by patterns, the distribution is

reasonably uniform (the light orange color cells), especially for the larger grids, which showed smallest standard deviation values. This shows that although an attacker could exploit patterns, participants still generated a reasonably diverse pool of passwords.

Another note to make is that the focal attention appears to be distributed from left to right in all four experiments. This may be due to demographics, as the study was conducted in a Western country. Other studies evaluating patterns found similar behaviours [37]; results may differ if the study is conducted on a population with right-to-left readers.

7.6.2.2 Experiment: individual questions

Q1: *Creating this password was enjoyable* (options: 5-point Likert scale: Strongly agree (1) - Strongly disagree (5)) (Results in Table 7.5)

Exp	Type	Min	Max	Mean	Median	STD	Var	B. Box[4]	T. Box[5]
1	Overall	1	5	2.67	2	1.13	1.29	0.52	0.25
	Desktop	1	5	2.73	3	1.20	1.43	0.49	0.31
	Mobile	1	5	2.59	2	1.03	1.07	0.55	0.17
2	Overall	1	5	2.45	2	1.10	1.21	0.60	0.19
	Desktop	1	5	2.46	2	1.04	1.09	0.59	0.16
	Mobile	1	5	2.45	2	1.18	1.39	0.62	0.24
3	Overall	1	5	2.52	2	1.13	1.27	0.57	0.20
	Desktop	1	5	2.64	2	1.08	1.17	0.52	0.22
	Mobile	1	5	2.34	2	1.17	1.36	0.64	0.17
4	Overall	1	5	2.62	2	1.11	1.23	0.53	0.22
	Desktop	1	5	2.65	2	1.09	1.19	0.53	0.22
	Mobile	1	5	2.57	2	1.13	1.28	0.52	0.22

Table 7.5: Question 1 (Results)

Overall, E2 (triangles, in order) was deemed the most enjoyable experiment. For mobile users, E3 scored best, followed by E1, while E4 had lowest enjoyment results; it appears rectangular grids are preferred by mobile users, irrespective of the cell ordering. Desktop-based respondents were less consistent: E2 scored most highly, but the second best-rated experiment was E3, its polar opposite (triangles, in order vs. rectangles, no order).

Across all experiments, mobile respondents responded more positively, having a larger proportion of the answers in the positive bracket (the bottom box).

Q2: *Inputting this password was easy* (options: 5-point Likert scale: Strongly agree (1) - Strongly disagree (5)) (Results in Table 7.6)

[4]Strongly agree + Somewhat agree answer bands
[5]Somewhat disagree + Strongly disagree answer bands
[6]Strongly agree + Somewhat agree answer bands
[7]Somewhat disagree + Strongly disagree answer bands

Exp	Type	Min	Max	Mean	Median	STD	Var	B. Box[6]	T. Box[7]
1	Overall	1	5	2.18	2	1.25	1.56	0.68	0.23
	Desktop	1	5	2.40	2	1.29	1.67	0.60	0.28
	Mobile	1	5	1.88	1.5	1.12	1.24	0.79	0.16
2	Overall	1	5	2.07	2	1.14	1.30	0.73	0.14
	Desktop	1	5	2.31	2	1.17	1.37	0.64	0.17
	Mobile	1	5	1.74	1	1.01	1.02	0.84	0.10
3	Overall	1	5	2.18	2	1.15	1.33	0.68	0.18
	Desktop	1	5	2.28	2	1.24	1.54	0.63	0.23
	Mobile	1	5	2.03	2	1.00	1.00	0.74	0.10
4	Overall	1	5	2.40	2	1.16	1.35	0.60	0.20
	Desktop	1	5	2.42	2	1.18	1.40	0.58	0.20
	Mobile	1	5	2.38	2	1.13	1.27	0.62	0.21

Table 7.6: Question 2 (Results)

In terms of ease of input, mobile and desktop respondents agreed on their answers: E2 scored best, while E4 scored worst. For the second-best, mobile users chose E1, while desktop users preferred E3. This suggests that mobile users find serial recall (E1 and E2) easier to manage than free recall. Desktop users, on the other hand, ordered experiments the same as in Q1: E2, then E3. It appears desktop users are more supportive towards free recall than mobile users.

Again, mobile users responded much more positively than desktop users: in the case of E2, the best rated experiment, 64% of desktop users rated it positively, and 84% of the mobile users agreed it was very easy to input.

Q3: *How likely would you be to remember this password in a week's time?* (options: 5-point Likert scale: Extremely unlikely (1) - Extremely likely (5)) (Results in Table 7.7)

Exp	Type	Min[8]	Max	Mean	Median	STD	Var	B. Box[9]	T. Box[10]
1	Overall	12	16	13.99	14	1.34	1.81	0.46	0.48
	Desktop	12	16	13.89	13	1.32	1.75	0.51	0.44
	Mobile	12	16	14.14	15	1.36	1.84	0.40	0.53
2	Overall	12	16	13.97	14	1.36	1.84	0.45	0.45
	Desktop	12	16	13.93	14	1.32	1.75	0.44	0.42
	Mobile	12	16	14.03	14	1.40	1.96	0.45	0.48
3	Overall	12	16	13.83	14	1.34	1.78	0.46	0.38
	Desktop	12	16	13.89	14	1.33	1.78	0.42	0.37
	Mobile	12	16	13.74	13	1.33	1.78	0.52	0.40
4	Overall	12	16	13.63	13	1.27	1.62	0.52	0.31
	Desktop	12	16	13.79	14	1.28	1.65	0.48	0.36
	Mobile	12	16	13.40	13	1.22	1.48	0.57	0.24

Table 7.7: Question 3 (Results)

Across all experiments, participants tended to consider themselves *Less likely* to remember the passwords in the medium-long term: in Table 7.7, the top box is (with one exception: Experiment 1, Mobile) consistently under 0.50, meaning that less than 50% of the participants considered themselves *Likely* or *Extremely likely* to remember the passwords in a week's time. Bearing that in mind, participants found E1 and E2 the most memorable, regardless of the device used. For both desktop and mobile respondents, the free-recall experiments (E3 and E4) scored much poorer than the serial-recall correspondents, irrespective of cell shape.

Recall the previously-stated hypothesis:

Hypothesis Compared to ordered retrieval, unordered graphical password input can improve ease of use by decreasing the cognitive load.

The results contradict the hypothesis of free recall easing the cognitive load: it appears that even though users do not support any scheme to be more likely to remember than not, they do believe they are less likely to remember a selection of 8 cells over only 6.

Overall results show that there are differences between mobile and desktop users: while mobile users preferred E1 across all three metrics, desktop users preferred E2. The two experiments differ in the shape used, but have the same ordering - this once again confirms that serial recall is dominant regardless of the device used. We note differences when we consider the second-best ranked option: mobile users chose E2, which indicates that for mobile users, serial recall is clearly a better choice. Desktop users, however, chose E3; it appears that free recall *could* be accepted as a feature by the desktop-oriented population.

The results do not draw any clear distinction over *shape* preferences: the top two choices of both device categories (E1 and E2, E2 and E3) each contain one triangular grid and one rectangular grid.

7.6.3 Final questionnaire results

See full breakdown of answers at Appendix I.

Main set of questions

Q1: *Drag and drop to rank the four experiments by how enjoyable it was to create a password from 1 (most enjoyable) to 4 (least enjoyable)*:

Figure 7.8 shows a breakdown of users' ordering of the experiments in terms of enjoyment of use.

We notice that E1 is the most preferred experiment: it was most commonly ranked among the top two options (57%) than among the bottom two (42%).

[8]The Qualtrics platform automatically adjusted the variable values from 1 - 5 to 12 - 16, which is why Min and Max values are 12 and 16 respectively.

[9]Extremely unlikely + Somewhat unlikely answer bands

[10]Somewhat likely + Extremely likely answer bands

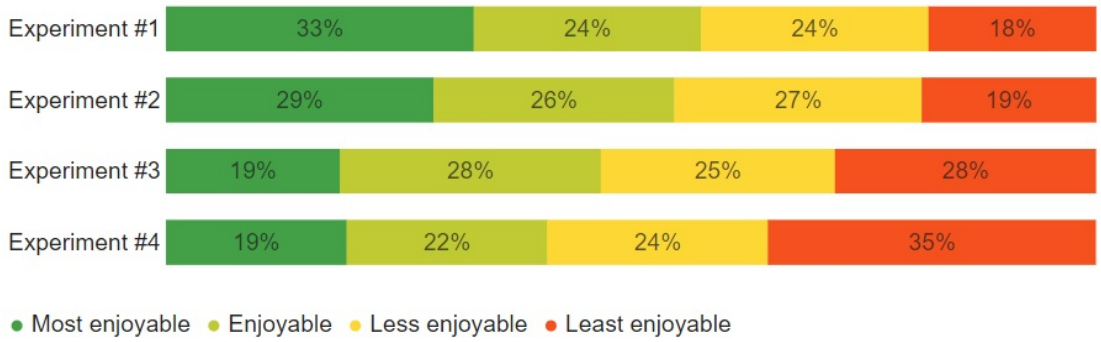


Figure 7.8: Question 1: Enjoyment of use

If we compare the experiments of the same *ordering* (E1 and E2 vs. E3 and E4), we see that they are similarly distributed: E1 and E2 are more often preferred than disliked, while E3 and E4 are more frequently put on the last positions (*Less* and *Least enjoyable*). The results continue to support those of the Experiments: users consistently ranked the serial recall options above the free recall ones.

On the other hand, if we compare between experiments of different *shapes* (E1 vs. E2 and E3 vs. E4), we find that triangular grids (E2 and E4) score more in the lower brackets than in the higher brackets, indicating that people found rectangular grids overall more enjoyable than triangular grids.

Q2: *Drag and drop to rank the four experiments by how easy it was to input a password on from 1 (easiest) to 4 (hardest)*:

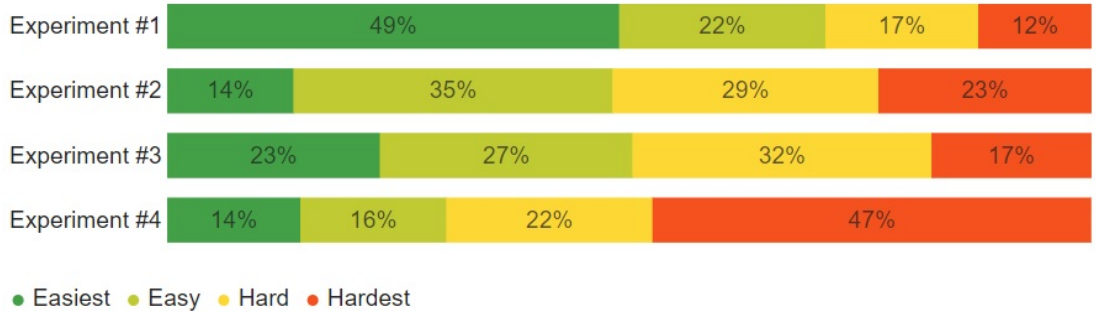


Figure 7.9: Question 2: Ease of input

The results support the same findings of Q1. E1 tops again, by far considered the easiest to input the password on.

Comparing between serial and free recall (E1 and E2 vs. E3 and E4), serial experiments continue to outperform the free recall ones. While E1 and E2 are less often marked *Hard* or *Hardest* (29% and 51%), E3 and (particularly) E4 are deemed *Hard* or *Hardest* to input in 49% and 69% of the time.

When comparing experiments of different *shapes* (E1 vs. E2 and E3 vs. E4), we observe that rectangular grids consistently score approx. 50% more in the top two boxes than in the bottom two: 71% vs 49% in E1 vs. E2, 50% vs. 30% in E3 vs. E4.

Q3: *If I created a visual password, I would prefer to...* (options: select fewer cells, but retrieve them in a specific order (1) | select more cells, but retrieve them in any order (2) | no preference (3))

Expectedly, the option to select fewer cells, but retrieve them in a specific order was preferred by both types of platforms. However, mobile respondents showed a more significant gap than desktop respondents (**mobile**: 66% (1) and 24% (2), **desktop**: 49% (1) and 40% (2)). The smaller gap for the desktop users, and the 10% of no preference (3) respondents in both cases suggest that non-ordering could also be positively received and approached by some users.

Q4: *Out of the two geometrical shapes used to create the visual passwords, I preferred using...* (options: triangles (1) | rectangles (2) | no preference (3))

Percentages were very close for both options, showing no dominant preference for any geometrical shape (**mobile**: 47% (1) and 47% (2), **desktop**: 43% (1) and 49% (2)). Although this answer implies no difference in preferences, results in Q7 show that actually there is a difference in perceived memorability between shapes.

Q5: *To what extent do you agree or disagree with the following statement: The background image influenced my choice of cells while I was creating my passwords* (options: 5-point Likert scale: Strongly agree (1) - Strongly disagree (5))

The results of this question are somewhat in opposition, with 58% of the respondents strongly agreeing (1) with the statement and 30% of them strongly disagreeing (5). The results for each of the platforms showed the same trend as the overall results, suggesting that the platform did not influence their decision. Even though this proves that the background image was important for more than half of the users, concerningly, 40% of the population only relies on patterns.

Q6: *Out of the four visual passwords you created, which do you think are most secure?* (options: E1 and E2 (1) | E3 and E4 (2))

The reader will recall that in Section 7.5.3.2 we constrained all four experiment conditions to be equally secure; this question was intentionally misleading to evaluate users' perceptions. The majority (69%) of respondents chose option (1) - this shows that serial recall is so deeply ingrained in participants' perceptions when it comes to passwords, it is considered the most secure.

Q7: *Out of the four visual passwords you created, which do you think is easiest to memorize?* (options: E1 | E2 | E3 | E4)

Serial recall on a rectangular grid (E1) was best-scoring option for each of the platforms, taking up 35% of the votes. While mobile users further proved they prefer serial over free recall (61% vs. 39%), desktop users ranked them equally memorable (48% vs. 52%).

Furthermore, for desktop users, the triangular grid (E4) scored lower than the rectangular grid (E3). Even though in Q4 desktop participants showed no significant preference of rectangles over triangles, they strongly believed the rectangular grid was easier to memorize than the triangular grid. Mobile users, however, do not seem to be interested

in the shapes; rectangular experiments scored only 5% better than triangular ones.

Final set of questions

Statement: *To what extent do you agree or disagree with the following statements about this visual password system?* (options: 5-point Likert scale: Strongly agree (1) - Strongly disagree (5))

Q1: "*I found the system simple to use*"

Overall, over 80% of respondents *Agreed* to the statement, irrespective of the platform.

Q2: "*I would use this system on a daily basis*"

The overall answer trend in this question is negative, with 25% of the answers in the *Agree* category and over 55% in the *Disagree* category. The difference between platforms, however, is quite significant. 19% of the desktop users *Agreed* to the statement, while 60% of them *Disagreed*. On the other hand, more of the mobile users (33%) *Agreed* to the statement, while fewer (47%) *Disagreed*.

The results positively indicate that although the system may not be a feasible system for everyday use for desktop users, mobile users are a potential target population.

Q3: "*These passwords are easier to remember than text-based passwords*"

Similarly to the previous question, the overall trend is towards a negative answer, with only 21% of respondents answering in the affirmative box. A notable part (25%) of the population also responded as *Neither/nor*. Comparing the trends by platform, desktop users were more notably negative, with 63% of them *Disagreeing* with the statement. For the mobile users, the answers still pointed towards disagreement, but with a less steep difference.

Q4: "*I would use such a system for my low-importance accounts*"

40% of the respondents *Agreed* with the statement, while only 20% of them *Strongly disagreed*, irrespective of the platform used. It seems that the perceived security of the system appears to be sufficient for users' low-importance accounts.

Q5: "*I would use such a system for my high-importance accounts*"

For the converse statement, 48% of the participants *Strongly disagreed*, across both platforms. Only 15% of the participants *Strongly* or *Somewhat agreed* with the statement; from this result it is overwhelmingly clear users would not trust their high importance accounts with such a system.

Q6: "*I would prefer to use my own background image instead of a stock image*"

67% of participants *Agreed* with the statement while 33% of them *Disagreed* or did not have a preference. This entails that although the option of customizing the background image is desired, stock images could still appeal to a part of the population.

7.6.4 Conclusion

The user study results lead to an interesting set of conclusions, which can serve as recommendations for future graphical authentication systems. However, the reader must keep in mind that these results are indicators and not statistically-backed evidence, as a statistical analysis was not possible at the time the results were processed.

This graphical passwords tool proposed the use of free recall in click-based graphical passwords to improve ease of use. The results ultimately contradicted the hypothesis; after having created passwords using both retrieval processes, users generally still preferred serial recall. Considering that the vast majority of existing password schemes (especially text passwords) enforce ordering, this result most likely occurred because ordered recall passwords come naturally to users. However, the results were not in entirety against free recall, as free-ordered graphical passwords have appealed to desktop users.

> **Finding** If the survey completion platform is taken into account, mobile users strongly prefer serial recall, whereas desktop users are more open to free recall.

Another independent condition exercised was the shape of the password grid cells.

> **Finding** Desktop users seem to prefer rectangular grids, while mobile users are open to alternative forms of discretization.

The study fixed a background image under the password grid. The image was chosen to have a multitude of salient points to avoid central points of attention, while the grid was placed atop to aid memorization.

The hotspot analysis of the passwords created reveals that the clear discretization has provoked probable password patterns, despite the use of the uniform image. It appears that users' predictability in creating text passwords can manifest on graphical passwords as well, through the use of patterns. The question remains on whether the lock-out policy of the system is sufficient against this drawback; this research question is recommended as *Future work* (see 8.2).

In terms of the customization features supported by the graphical passwords tool, users responded positively to the idea of using their personal images as backgrounds.

> **Finding** Participants claim to prefer to be able to personalize their graphical passwords by uploading and using their personal images as backgrounds, but would still accept using uniform stock images.

The tool also registered positive findings in terms of user adoption.

> **Finding** Desktop users were willing to rely on the tool for infrequent use, while mobile users responded positively towards more frequent use. In terms of perceived security, users would not trust the tool to protect their important accounts, but were attracted by the idea of using them to protect lower-value accounts.

Chapter 8

Conclusion

8.1 Summary

This research project explored the vast pool of password solutions proposed in the past two decades, identifying the common challenges and difficulties users face when managing passwords created using those solutions. I proposed a list of quality requirements for an alternative authentication method (Chapter 2, Section 2.3), then sought for means to fulfill the requirements. I focused my research on two popular text password alternatives: graphical passwords and password managers, using them to design and evaluate two distinct authentication solutions.

The current state of the password world seems to be a *Pareto equilibrium* [33]. Replacing text passwords with graphical passwords, password managers or other authentication means cannot guarantee a better security ecosystem for users; it just trades one set of advantages and drawbacks for another. Then, educating people to correctly make use of alternative solutions is another challenge in itself.

Bearing that in mind, the graphical password tool proposed, as well as both its implementations (Chapter 6 and 7) are only suitable for the needs and requirements of some subset of the population, while may well be unfit for others. The findings point that the tool is likely to record positive usability results if implemented on touchscreen devices.

My research concludes that a universal authentication solution is almost unfeasible to achieve; however, carefully tailored systems, dedicated to either distinct population categories, or distinct digital platforms are more likely to be successful in practice.

8.2 Future work

Future work on this project could explore different user study circumstances, as well as different aspects related to enjoyment, memorability and ease of use in click-based graphical passwords.

Long-term memorability The long-term memorability of the scheme could be evaluated in a future study. Recall that the present questionnaire only queried users on their

perceived and not their true ability to memorize the graphical passwords in medium-long term. For this purpose, however, web studies are unsuitable as per [49], which claims that memorability cannot be reliably assessed in an online study, but in a controlled lab environment.

Target population Our user study population vastly consisted of highly educated, frequent web users. For better ecological validity, similar studies should target a more diverse user population, beyond University students and academia. In the first user study conducted in Chapter 6, security experts flagged users with *mental health issues and other disabilities* as a population segment prone to poor online security; this project could not cover these demographics, however, future work should address the concerns of such users.

Security of lock-out policies The present graphical passwords tool is constrained on the existence of a lock-out policy, thus allowing its smaller password space to be sufficient against online attacks. As found in the frequency analysis of the passwords created in the study, users are likely to produce predictable password patterns. It is uncertain, however, whether the lock-out policy is sufficient to maintain user accounts secure if their graphical passwords consist of predictable patterns. Future work should investigate the effective risk salient patterns pose in the security of click-based graphical password schemes with lock-out policies in place.

Free recall This project introduced free recall in graphical passwords - retrieving elements of a password not constrained by the order at creation - and evaluated users' opinions of this memorization approach. Even though the overall results had a negative tendency, desktop users seemed enthusiastic about the added freedom. If free recall in graphical password schemes was to be further explored into, it should be integrated into a *desktop-based* password tool.

Muscle memory on touchscreen devices Last but not least, future work should consider frequent use of graphical password schemes on mobile devices. In recent years, smartphones and tablets have become increasingly popular; more authentication systems should tailor to these devices. It is long known that people can develop the ability to type their passwords using muscle memory [33]. By frequently using graphical passwords on smartphone devices, users could potentially develop similar habits.

Bibliography

[1] *Good food reads: Pinch of yum tasty food photography workshop*, Accessed April 1, 2020. https://withtwospoons.com/good-food-reads-pinch-yum-tasty-food-photography-workshop/.

[2] *The Paradox of Choice: Why Less is More in UX Design*, Accessed April 1, 2020. https://usabilla.com/blog/paradox-choice-less-ux-design/.

[3] *For the People Who Don't Trust Password Managers*, Accessed April 10, 2020. https://passwordbits.com/trust-password-managers/.

[4] *Starbucks: Sign in or create an account*, Accessed April 11, 2020. https://www.starbucks.com/account/signin.

[5] *Cryptographic hash function*, Accessed April 12, 2020. https://en.wikipedia.org/wiki/Cryptographic_hash_function.

[6] *Special Publication 800-63*, Accessed April 13, 2020. https://www.nist.gov/itl/tig/projects/special-publication-800-63.

[7] *xkcd: Password strength*, Accessed April 13, 2020. https://xkcd.com/936/.

[8] *What Is the Difference Between Independent and Dependent Variables?*, Accessed April 14, 2020. https://www.thoughtco.com/independent-and-dependent-variables-differences-606115.

[9] *Password administration for system owners*, Accessed April 15, 2020. https://www.ncsc.gov.uk/collection/passwords.

[10] *Adding Salt to Hashing: A Better Way to Store Passwords*, Accessed April 3, 2020. https://auth0.com/blog/adding-salt-to-hashing-a-better-way-to-store-passwords/.

[11] *How many possible combinations in 8 character password?*, Accessed April 3, 2020. https://math.stackexchange.com/questions/739874/how-many-possible-combinations-in-8-character-password.

[12] *Password Cracking Is Easy: Here's How to Do It*, Accessed April 3, 2020. https://towardsdatascience.com/password-cracking-is-easy-heres-how-to-do-it-875806a1e42a.

[13] *ISO 9241-11:2018 (en) Ergonomics of human-system interaction — Part 11: Usability: Definitions and concepts*, Accessed April 7, 2020. https://iso.org/obp/ui/#iso:std:iso:9241:-11:ed-2:v1:en.

[14] *Memory recall / retrieval*, Accessed April 7, 2020. https://human-memory.net/memory-recall-retrieval/.

[15] *Number of Records Exposed in 2019 Hits 15.1 Billion*, Accessed April 7, 2020. https://www.riskbasedsecurity.com/2020/02/10/number-of-records-exposed-in-2019-hits-15-1-billion/.

[16] *1Password: Password Manager for Families, Businesses...*, Accessed April 9, 2020. https://www.1password.com.

[17] *Dashlane: Password Manager App for Home, Mobile, Business*, Accessed April 9, 2020. https://www.dashlane.com.

[18] *LastPass: 1 Password Manager Vault App, Enterprise SSO...*, Accessed April 9, 2020. https://www.lastpass.com.

[19] *Saving Passwords In Google Chrome Is Better Than Nothing (And That's A Good Thing!)*, Accessed April 9, 2020. https://www.brucebnews.com/2018/10/saving-passwords-in-google-chrome-is-better-than-nothing-and-thats-a-good-thing/.

[20] *The best password managers for 2020*, Accessed February 25, 2020. https://uk.pcmag.com/password-managers/4296/the-best-password-managers-for-2020.

[21] *KeePass Password Safe*, Accessed February 25, 2020. https://keepass.info/.

[22] *Password manager*, Accessed February 25, 2020. https://en.wikipedia.org/wiki/Password_manager.

[23] *Password Safe: Simple and secure password management*, Accessed February 25, 2020. https://pwsafe.org/.

[24] *Web Confidential: The classic password manager*, Accessed February 25, 2020. http://www.web-confidential.com/.

[25] *Passfaces: Two Factor Authentication for the Enterprise*, Accessed March 11, 2020. http://www.realuser.com/.

[26] *Image by Laura Montagnani on Pixabay*, Accessed March 12, 2020. https://pixabay.com/photos/morocco-market-fruit-vegetables-3794323/.

[27] *Qualtrics: Online Survey Software*, Accessed March 12, 2020. https://www.qualtrics.com/uk/core-xm/survey-software/.

[28] Mahdi Nasrullah Al-Ameen, Kanis Fatema, Matthew Wright, and Shannon Scielzo. The impact of cues and user interaction on the memorability of system-assigned recognition-based graphical passwords. In *Proceedings of the*

Eleventh USENIX Conference on Usable Privacy and Security, SOUPS '15, page 185–196, USA, 2015. USENIX Association.

[29] Nora Alkaldi and Karen Renaud. Why do people adopt, or reject, smartphone password managers? 01 2016.

[30] Salvatore Aurigemma, Thomas Mattson, and Lori N. K. Leonard. So much promise, so little use: What is stopping home end-users from using password manager applications? In *HICSS*, 2017.

[31] Kemal Bicakci, Mustafa Yuceel, Burak Erdeniz, Hakan Gurbaslar, and Nart Bedin Atalay. Graphical passwords as browser extension: Implementation and usability study. In Elena Ferrari, Ninghui Li, Elisa Bertino, and Yuecel Karabulut, editors, *Trust Management III*, pages 15–29, Berlin, Heidelberg, 2009. Springer Berlin Heidelberg.

[32] Robert Biddle, Sonia Chiasson, and P. Oorschot. Graphical passwords: Learning from the first twelve years. *ACM Computing Surveys - CSUR*, 44, 01 2012.

[33] J. Bonneau, C. Herley, P. C. v. Oorschot, and F. Stajano. The quest to replace passwords: A framework for comparative evaluation of web authentication schemes. In *2012 IEEE Symposium on Security and Privacy*, pages 553–567, 2012.

[34] Timothy F. Brady, Talia Konkle, George A. Alvarez, and Aude Oliva. Visual long-term memory has a massive storage capacity for object details. *Proceedings of the National Academy of Sciences of the United States of America*, 105(38):14325–14329, Sep 2008.

[35] Karoline Busse, Julia Schäfer, and Matthew Smith. Replication: No one can hack my mind revisiting a study on expert and non-expert security practices and advice. In *Fifteenth Symposium on Usable Privacy and Security (SOUPS 2019)*, Santa Clara, CA, August 2019. USENIX Association.

[36] Sunil Chaudhary, Tiina Schafeitel-Tähtinen, Marko Helenius, Eleni Berki, and Harri Kiljander. Usability and security in password managers: A quest for user-centric properties and features, 11 2016.

[37] Sonia Chiasson, Alain Forget, Robert Biddle, and Paul Oorschot. User interface design affects security: Patterns in click-based graphical passwords. *Int. J. Inf. Sec.*, 8:387–398, 12 2009.

[38] Sonia Chiasson, Alain Forget, Elizabeth Stobert, P. C. van Oorschot, and Robert Biddle. Multiple password interference in text passwords and click-based graphical passwords. In *Proceedings of the 16th ACM Conference on Computer and Communications Security*, CCS '09, page 500–511, New York, NY, USA, 2009. Association for Computing Machinery.

[39] Sonia Chiasson, Elizabeth Stobert, Alain Forget, Robert Biddle, and Paul Oorschot. Persuasive cued click-points: Design, implementation, and evaluation of a knowledge-based authentication mechanism. *IEEE Trans. Dependable Sec. Comput.*, 9:222–235, 03 2012.

[40] Sonia Chiasson and P. C. van Oorschot. Quantifying the security advantage of password expiration policies. *Designs, Codes and Cryptography*, 77(2):401–408, 2015.

[41] Sonia Chiasson, P. C. van Oorschot, and Robert Biddle. A usability study and critique of two password managers. In *Proceedings of the 15th Conference on USENIX Security Symposium - Volume 15*, USENIX-SS'06, USA, 2006. USENIX Association.

[42] Sonia Chiasson, P. C. Van Oorschot, and Robert Biddle. Graphical password authentication using cued click points. In *Proceedings of the 12th European Conference on Research in Computer Security*, ESORICS'07, page 359–374, Berlin, Heidelberg, 2007. Springer-Verlag.

[43] Anupam Das, Joseph Bonneau, Matthew Caesar, Nikita Borisov, and Xiaofeng Wang. The tangled web of password reuse. 01 2014.

[44] Darren Davis, Fabian Monrose, and Michael K. Reiter. On user choice in graphical password schemes. In *Proceedings of the 13th Conference on USENIX Security Symposium - Volume 13*, SSYM'04, page 11, USA, 2004. USENIX Association.

[45] Antonella De Angeli, Lynne Coventry, Graham Johnson, and Karen Renaud. Is a picture really worth a thousand words? exploring the feasibility of graphical authentication systems. *International Journal of Human-Computer Studies*, 63:128–152, 07 2005.

[46] Ahmet Emir Dirik, Nasir Memon, and Jean-Camille Birget. Modeling user choice in the passpoints graphical password scheme. In *Proceedings of the 3rd Symposium on Usable Privacy and Security*, SOUPS '07, page 20–28, New York, NY, USA, 2007. Association for Computing Machinery.

[47] Paul Dunphy and Jeff Yan. Do background images improve "draw a secret" graphical passwords? pages 36–47, 01 2007.

[48] Michael Fagan, Yusuf Albayram, Mohammad Maifi Hasan Khan, and Ross Buck. An investigation into users' considerations towards using password managers. *Human-centric Computing and Information Sciences*, 7(1):12, 2017.

[49] Sascha Fahl, Marian Harbach, Yasemin Acar, and Matthew Smith. On the ecological validity of a password study. In *Proceedings of the Ninth Symposium on Usable Privacy and Security*, SOUPS '13, New York, NY, USA, 2013. Association for Computing Machinery.

[50] Dinei Florencio, Cormac Herley, and Baris Coskun. Do strong web passwords accomplish anything? Technical Report MSR-TR-2007-64, June 2007.

[51] Google. *Create a strong password & a more secure account*, Accessed March 26, 2020. https://support.google.com/accounts/answer/32040?hl=en.

[52] Iulia Ion, Rob Reeder, and Sunny Consolvo. "...No one Can Hack My Mind": Comparing expert and non-expert security practices. In *Eleventh Symposium On*

Usable Privacy and Security (SOUPS 2015), pages 327–346, Ottawa, July 2015. USENIX Association.

[53] Ian Jermyn, Alain Mayer, Fabian Monrose, Michael K. Reiter, and Aviel D. Rubin. The design and analysis of graphical passwords. In *Proceedings of the 8th Conference on USENIX Security Symposium - Volume 8*, SSYM'99, page 1, USA, 1999. USENIX Association.

[54] Ravi Jhawar, Philip Inglesant, Nicolas Courtois, and Angela Sasse. Make mine a quadruple: Strengthening the security of graphical one-time pin authentication. pages 81–88, 09 2011.

[55] Ambarish Karole, Nitesh Saxena, and Nicolas Christin. A comparative usability evaluation of traditional password managers. In *Proceedings of the 13th International Conference on Information Security and Cryptology*, ICISC'10, page 233–251, Berlin, Heidelberg, 2010. Springer-Verlag.

[56] Krystal A. Klein, Kelly M. Addis, and Michael J. Kahana. A comparative analysis of serial and free recall. *Memory & Cognition*, 33(5):833–839, 2005.

[57] Saranga Komanduri, Richard Shay, Patrick Gage Kelley, Michelle L. Mazurek, Lujo Bauer, Nicolas Christin, Lorrie Faith Cranor, and Serge Egelman. Of passwords and people: Measuring the effect of password-composition policies. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, page 2595–2604, New York, NY, USA, 2011. Association for Computing Machinery.

[58] Nate Lord. *Uncovering Password Habits: Are Users' Password Security Habits Improving? (Infographic)*, Accessed March 12, 2020. https://digitalguardian.com/blog/uncovering-password-habits-are-users-password-security-habits-improving-infographic.

[59] Sanam Ghorbani Lyastani, Michael Schilling, Sascha Fahl, Michael Backes, and Sven Bugiel. Better managed than memorized? studying the impact of managers on password strength and reuse. In *Proceedings of the 27th USENIX Conference on Security Symposium*, SEC'18, page 203–220, USA, 2018. USENIX Association.

[60] William Melicher, Darya Kurilova, Sean M. Segreti, Pranshu Kalvani, Richard Shay, Blase Ur, Lujo Bauer, Nicolas Christin, Lorrie Faith Cranor, and Michelle L. Mazurek. Usability and security of text passwords on mobile devices. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, page 527–539, New York, NY, USA, 2016. Association for Computing Machinery.

[61] Deholo Nali and Julie Thorpe. Analyzing user choice in graphical passwords. Technical report, School of Computer Science, Carleton University, 2004. Technical Report TR-04-01.

[62] D. L. Nelson, V. S. Reed, and J. R. Walling. Journal of experimental psychology: Human learning and memory., 1975.

[63] Donald A. Norman. *The Design of Everyday Things*. Basic Books, Inc., 1988.

[64] Kenneth Olmstead and Aaron Smith. *Americans and Cybersecurity*, 2017. http://assets.pewresearch.org/wpcontent/uploads/sites/14/2017/01/26102016/Americans-and-CyberSecurity-final.pdf.

[65] Allan Paivio, T. B. Rogers, and Padric C. Smythe. Why are pictures easier to recall than words? *Psychonomic Science*, 11(4):137–138, 1968.

[66] Tobias Seitz. Supporting users in password authentication with persuasive design. July 2018.

[67] Richard Shay, Patrick Gage Kelley, Saranga Komanduri, Michelle L. Mazurek, Blase Ur, Timothy Vidas, Lujo Bauer, Nicolas Christin, and Lorrie Faith Cranor. Correct horse battery staple: Exploring the usability of system-assigned passphrases. In *Proceedings of the Eighth Symposium on Usable Privacy and Security*, SOUPS '12, New York, NY, USA, 2012. Association for Computing Machinery.

[68] Richard Shay, Saranga Komanduri, Adam Durity, Philip Huh, Michelle Mazurek, Sean Segreti, Blase Ur, Lujo Bauer, Nicolas Christin, and Lorrie Cranor. Can long passwords be secure and usable? 04 2014.

[69] Richard Shay, Saranga Komanduri, Patrick Kelley, Pedro Leon, Michelle Mazurek, Lujo Bauer, Nicolas Christin, and Lorrie Cranor. Encountering stronger password requirements: User attitudes and behaviors. 09 2010.

[70] Lionel Standing. Learning 10000 pictures. *Quarterly Journal of Experimental Psychology*, 25(2):207–222, 1973. PMID: 4515818.

[71] Elizabeth Stobert. Graphical passwords and practical password management. 2015.

[72] Elizabeth Stobert and Robert Biddle. Memory retrieval and graphical passwords. In *Proceedings of the Ninth Symposium on Usable Privacy and Security*, SOUPS '13, New York, NY, USA, 2013. Association for Computing Machinery.

[73] Elizabeth Stobert and Robert Biddle. The password life cycle: User behaviour in managing passwords. In *SOUPS*, 2014.

[74] Xiaoyuan Suo, Ying Zhu, and G. Owen. Graphical passwords: A survey. pages 463–472, 01 2005.

[75] Julie Thorpe and P. C. van Oorschot. Human-seeded attacks and exploiting hotspots in graphical passwords. In *Proceedings of 16th USENIX Security Symposium on USENIX Security Symposium*, SS'07, USA, 2007. USENIX Association.

[76] Endel Tulving and Michael Watkins. Continuity between recall and recognition. *The American Journal of Psychology*, 86:739, 12 1973.

[77] Rick Wash, Emilee Rader, Ruthie Berman, and Zac Wellmer. Understanding password choices: How frequently entered passwords are re-used across web-

sites. In *Proceedings of the Twelfth USENIX Conference on Usable Privacy and Security*, SOUPS '16, page 175–188, USA, 2016. USENIX Association.

[78] Nancy C. Waugh. Free versus serial recall. *Journal of Experimental Psychology*, 62(5):496–502, 1961.

[79] Susan Wiedenbeck, Jim Waters, Jean-Camille Birget, Alex Brodskiy, and Nasir Memon. Passpoints: Design and longitudinal evaluation of a graphical password system. *International Journal of Human-Computer Studies*, 63:102–127, 07 2005.

[80] Jeff Yan, Alan Blackwell, Ross Anderson, and Alasdair Grant. Password memorability and security: Empirical results. *Security Privacy, IEEE*, 2:25 – 31, 10 2004.

[81] Shikun Aerin Zhang, Sarah Pearman, Lujo Bauer, and Nicolas Christin. Why people (don't) use password managers effectively. In *Fifteenth Symposium on Usable Privacy and Security (SOUPS 2019)*, Santa Clara, CA, August 2019. USENIX Association.

Appendix A

User study (Chapter 6): Participant Information Sheet

Participant Information Sheet

Project title:	Stimulating mindfulness while generating passwords
Principal investigator:	Robin L. Hill
Researcher collecting data:	Bianca Burtoiu
Funder (if applicable):	N/A

This study was certified according to the Informatics Research Ethics Process, RT number RT 70309. Please take time to read the following information carefully. You should keep this page for your records.

Who are the researchers?

4th year Computer Science student at The University of Edinburgh, Bianca Burtoiu and project supervisor Robin L. Hill.

What is the purpose of the study?

The study aims to gather information on password security practices and password managers. The information will be used in the researcher's Honours project.

Why have I been asked to take part?

You have a high level of experience and expertise in either Computer Security, Human-Computer Interaction or Psychology.

Do I have to take part?

No – participation in this study is entirely up to you. You can withdraw from the study at any time, without giving a reason. Your rights will not be affected. If you wish to withdraw, contact the PI. We will stop using your data in any publications or presentations submitted after you have withdrawn consent. However, we will keep copies of your original consent, and of your withdrawal request.

What will happen if I decide to take part?

The researcher will conduct a semi-structured interview, where you will be invited to discuss various proposed topics related to password security practices and password managers. If permission is given, audio recording will be taken during the interview.

You will be interviewed a single time, and a session is estimated to last up to 45 minutes. Location can vary for your convenience. The session will take place between the 29th of October 2019 and 2nd of April 2020 (report submission deadline).

Compensation.

You will not be compensated for your participation in this study.

Are there any risks associated with taking part?

There are no significant risks associated with participation in this study.

Are there any benefits associated with taking part?

There are no benefits associated with participation in this study.

What will happen to the results of this study?
The results of this study may be summarised in the researcher's Honours project (reports and presentations). Quotes or key findings will be anonymized: We will remove any information that could, in our assessment, allow anyone to identify you. With your consent, information can also be used for future research. Your data may be archived for a minimum of 2 years.

Data protection and confidentiality.
Your data will be processed in accordance with Data Protection Law. All information collected about you will be kept strictly confidential. Your data will be referred to by a unique participant number rather than by name. Your data will only be viewed by the researcher Bianca Burtoiu and project supervisor Robin L. Hill.

All electronic data (written transcripts, audio recordings) will be initially stored on an encrypted Apple smartphone (without access to iCloud), then transferred on the School of Informatics' AFS secure file servers. All paper records will be stored in a locked filing cabinet in the PI's office. Your consent information will be kept separately from your responses in order to minimise risk.

What are my data protection rights?
The University of Edinburgh is a Data Controller for the information you provide. You have the right to access information held about you. Your right of access can be exercised in accordance Data Protection Law. You also have other rights including rights

of correction, erasure and objection. For more details, including the right to lodge a complaint with the Information Commissioner's Office, please visit www.ico.org.uk. Questions, comments and requests about your personal data can also be sent to the University Data Protection Officer at dpo@ed.ac.uk.

Who can I contact?
If you have any further questions about the study, please contact the lead researcher, Robin L. Hill (r.l.hill@ed.ac.uk, +44 (0) 131 650 4426).
If you wish to make a complaint about the study, please contact inf-ethics@inf.ed.ac.uk. When you contact us, please provide the study title and detail the nature of your complaint.

Updated information.
If the research project changes in any way, an updated Participant Information Sheet will be made available on https://web.inf.ed.ac.uk/infweb/research/study-updates.

Alternative formats.
To request this document in an alternative format, such as large print or on coloured paper, please contact Bianca Burtoiu (s1634680@sms.ed.ac.uk).

General information.
For general information about how we use your data, go to: edin.ac/privacy-research

Appendix B

User study (Chapter 6): Consent Form

Participant number: _____

Participant Consent Form

Project title:	Stimulating mindfulness while generating passwords
Principal investigator (PI):	Robin L. Hill
Researcher:	Bianca Burtoiu (s1634680@sms.ed.ac.uk)
PI contact details:	r.l.hill@ed.ac.uk, +44 (0) 131 650 4426

Please tick yes or no for each of these statements.

		Yes	No
1.	I confirm that I have read and understood the Participant Information Sheet for the above study, that I have had the opportunity to ask questions, and that any questions I had were answered to my satisfaction.		
2.	I understand that my participation is voluntary, and that I can withdraw at any time without giving a reason. Withdrawing will not affect any of my rights.		
3.	I agree to being audio recorded.		
4.	I consent to my anonymised data being used in academic publications and presentations.		
5.	I understand that my anonymised data can be stored for a minimum of two years.		
6.	I allow my data to be used in future ethically approved research.		
7.	I agree to take part in this study.		

Name of person giving consent	Date dd/mm/yy	Signature

Name of person taking consent	Date dd/mm/yy	Signature

THE UNIVERSITY of EDINBURGH informatics

Appendix C

User study (Chapter 6): Interview Script

General

1. Which categories of people do you think are most prone to poor password security? Why? (See how experts categorize people)

2. How should lay people be approached when it comes to educating them about password security? As an expert, which factors are you usually considering, which you think people without your knowledge would not figure out? (Are there any fundamental security aspects simpler to convey to non-experts?)

Password managers - A password manager assists in generating and retrieving complex passwords, potentially storing such passwords in an encrypted database or calculating them on demand.

1. Are you using a password manager? Why / why not? If yes, what made you use them / If no, why are you not using them? Follow up: (Why do you use a password manager vs another?)

2. Why do you think people do not use password managers? Even though they are widely advertised, their adoption is still low. Follow-up: What do you think would encourage people to use password managers?

3. One of the reasons people do not want to use password managers is the risk of ”keeping all their eggs in one basket”. How could we make password managers appeal to users' sense of autonomy (being in control of their passwords)?

4. On drawbacks of password managers, what do you think would be their most concerning points of weakness? Follow-up (if any found): Do you have any potential solutions to the issues?

5. Lay people seem to have trouble understanding (and thus trusting) the how strong encryption and security are achieved in password managers. How could we make the password manager architecture and underlying encryption functions easier to understand for non-experts?

6. Discussion: advantages and disadvantages between desktop-based (encrypted and stored locally), cloud-based (stored on a third-party server) and web browser-based password managers.

Which do you think a non-user would be most inclined to adopt? What about by a non-expert user?

7. What do you think about a password manager having password generation facilities? Research says they are not widely used and, consequently, some people are still storing weak passwords in their password managers.

Prototype - Description

The tool will work alongside a password manager, with the purpose to support people's existing coping strategies with passwords. It consists of an optional grid combination as a password.

The grid could be either used:

1. To ration effort: adding additional security to a password of high importance

2. By categorizing and labeling passwords, users could "reuse" the same grid to gain access into several different accounts, thus imitating password reuse in a safe manner.

Images used as password grid backgrounds are known to aid as memorization cues.

1. How do you think the trade-off should be made between technical notion of strength and usability? What is enough to work but won't make people think it is too much of a hassle?

2. How should the prototype convey the idea that that 7 squares on a grid are more secure than 3? How can you convey the level of security passwords have into a more easily understandable form?

3. What do you think about how the concept of effort rationing was applied in this prototype? What are your thoughts on the notion of password categorization?

4. To what segment of the population do you consider this tool would be most suitable for?

Appendix D

User study (Chapter 7): Participant Information Sheet

Participant Information Sheet

Project title:	Stimulating mindfulness while generating passwords
Principal investigator:	Robin L. Hill
Researcher collecting data:	Bianca Burtoiu
Funder (if applicable):	N/A

This study was certified according to the Informatics Research Ethics Process, RT number 2019/60527. Please take time to read the following information carefully.

Who are the researchers?

Bianca Burtoiu, 4th year Computer Science student at The University of Edinburgh and project supervisor Robin L. Hill.

What is the purpose of the study?

The study aims to gather information on people's perceptions of several different graphical password creation schemes. The information will be used towards the development of a graphical password creation tool, as part of the researcher's Honours project.

Why have I been asked to take part?

The research target group is the general adult public, individuals over 18 years old with varying levels of education.

Do I have to take part?

No – participation in this study is entirely up to you. You can withdraw from the study at any time, without giving a reason. Your rights will not be affected. If you wish to withdraw, close the browser. We will stop using your data in any publications or presentations submitted after you have withdrawn consent. If you choose to withdraw, we will not store any partially completed data.

What will happen if I decide to take part?

You will be asked to complete an online survey. The survey consists of three parts:

1. A short initial questionnaire concerning demographics and your relationship with technology and password management.
2. A role-playing experiment: You have just created a new e-mail account for which you need to come up with a secure password. In this experiment, you will create a visual password. You will be asked to create a password in four different ways, by following four different policies - essentially, you will end up creating four different passwords. For each experiment, please treat the procedure as realistically as possible - create a password that you would be able to memorize. The passwords you will create will be saved and analysed as part of the research.
3. A short follow-up questionnaire concerning your opinions on the password creation scheme you had just used.

Your screen resolution and device metadata information will be captured to determine the platform (e.g. mobile, desktop) you have completed the survey on. See more information at https://www.qualtrics.com/support/edit-survey/editing-questions/question-types-guide/advanced/meta-info-question/.

You can only participate in the study once. Completion of the survey is estimated to take approximately 15 minutes, but you can take as much time as you need. The survey will close on 2nd of April 2020.

Compensation.

You will not be compensated for your participation in this study.

Are there any risks associated with taking part?

There are no significant risks associated with participation in this study.

Are there any benefits associated with taking part?

There are no benefits associated with participation in this study.

What will happen to the results of this study?

The results of this study will be summarised in the researcher's Honours project (reports and presentations). Quotes or key findings will be anonymized: We will

remove any information that could, in our assessment, allow anyone to identify you. With your consent, information can also be used for future research. Your data may be archived for a minimum of 2 years.

Data protection and confidentiality.

Your data will be processed in accordance with Data Protection Law. All information collected about you will be kept strictly confidential. Your data will be referred to by a unique participant number rather than by name. Your data will only be viewed by the researcher Bianca Burtoiu and project supervisor Robin L. Hill.

All the data collected will be stored on the Qualtrics survey platform, run on the School of Informatics' secure file servers. No paper records will be produced.

What are my data protection rights?

The University of Edinburgh is a Data Controller for the information you provide. You have the right to access information held about you. Your right of access can be exercised in accordance Data Protection Law. You also have other rights including rights of correction, erasure and objection. For more details, including the right to lodge a complaint with the Information Commissioner's Office, please visit www.ico.org.uk. Questions, comments and requests about your personal data can also be sent to the University Data Protection Officer at dpo@ed.ac.uk.

Who can I contact?

If you have any further questions about the study, please contact the lead researcher, Robin L. Hill (r.l.hill@ed.ac.uk, +44 (0) 131 650 4426).
If you wish to make a complaint about the study, please contact inf-ethics@inf.ed.ac.uk. When you contact us, please provide the study title and detail the nature of your complaint.

Updated information.

If the research project changes in any way, an updated Participant Information Sheet will be made available on https://web.inf.ed.ac.uk/infweb/research/study-updates.

Alternative formats.

To request accessibility adjustments for this document, please contact Bianca Burtoiu (s1634680@sms.ed.ac.uk).

General information.

For general information about how we use your data, go to: edin.ac/privacy-research

Appendix E

User study (Chapter 7): Consent Form

Participant number:_____

Participant Consent Form

Project title:	Stimulating mindfulness while generating passwords
Principal investigator (PI):	Robin L. Hill
Researcher:	Bianca Burtoiu
PI contact details:	r.l.hill@ed.ac.uk

Please tick yes or no for each of these statements.

		Yes	No
1.	I confirm that I have read and understood the Participant Information Sheet above, that I have had the opportunity to contact the researchers, and that any questions I had were answered to my satisfaction.		
2.	I understand that my participation is voluntary, and that I can withdraw at any time without giving a reason, by closing the browser. Withdrawing will not affect any of my rights. If I choose to withdraw, no partially gathered data will be stored.		
3.	I consent to my anonymized data being used in academic publications and presentations.		
4.	I understand that my anonymized data can be stored for a minimum of two years.		
5.	I allow my data to be used in future ethically approved research.		
6.	I agree to take part in this study.		

Name of person giving consent	Date dd/mm/yy	Signature

Name of person taking consent	Date dd/mm/yy	Signature

THE UNIVERSITY of EDINBURGH informatics

Appendix F

User study (Chapter 7): Questionnaire

Demographics

1. What is your age?

- 18 - 24
- 25 - 34
- 35 - 44
- 45 - 54
- 55 - 64
- 65 or older
- I prefer not to say

2. What is the highest degree or level of education you have completed? If you are currently enrolled in school, please indicate the highest degree you will receive.

- Less than a high school diploma
- High school degree or equivalent
- Bachelor's degree
- Master's degree
- Doctorate
- Other (please specify)
- I prefer not to say

Technology

1. On average, how much time do you spend on Internet-related activities (e.g. e-mail, browsing, social media etc.) every day?

- Less than 1 hour

- 1 to 2 hours
- 3 to 5 hours
- More than 5 hours
- I do not use the Internet every day

2. Which of these Internet-related activities do you typically do every week? Select all that apply:

- Social media
- Navigation
- E-mail
- Music / Video streaming
- Messaging
- Planning (Calendar, Notes)
- Voice / Video calling
- Travel websites / apps
- Online banking
- Job search
- Online shopping
- Gaming
- Web browsing
- Education (e.g. online courses)

Password management

1. How many password-protected online accounts do you have?

- Less than 10
- Between 11 and 25
- Between 25 and 50
- Too many to count

2. Which statement concerning password behaviour best represents your current situation? Select one of the following:

- I use the same password across all my online accounts.
- I alternate between several different passwords for my online accounts.
- Each of my accounts has a different, unique password.
- Other (please specify)

3. How do you keep track of your passwords? Select all that apply:

- I memorize my passwords.
- I came up with a scheme that allows me to deduce the password of an account.
- I wrote my passwords onto a note stored in a safe place that I consult if needed.
- I am using a password manager that stores my usernames and passwords for me.
- Other (please specify)

4. What is your most important consideration when creating a password for an online account? Select one of the following:

- Security: making it complex and difficult to guess
- Convenience: making it easy to remember and/or type

5. How secure are you feeling with your current password management habits? (5-point Likert scale: Extremely secure - Extremely insecure)

Intro to experiment

Please read the following information before proceeding to the experiments.

The scenario: You have just created a new e-mail account for which you need to come up with a secure password. In this experiment, you will create a visual password.

You will be asked to create a password in four different ways, following four different policies - essentially, you will end up creating four different passwords. For each experiment, please treat the procedure as realistically as possible - create a password that you would be able to memorize.

Follow the on-screen instructions for each of the four experiments. Please note you will not be able to go back on your answers.

Important: The passwords you will create will be saved and analysed, so please do not submit a password that you may already be using somewhere else.

You can begin the first experiment by pressing the ”Next” button.

Experiment

1. To what extent do you agree or disagree with the following statements? (5-point Likert scale: Strongly agree - Strongly disagree)

1. Creating this password was enjoyable
2. Inputting this password was easy

2. How likely would you be to remember this password in a week’s time? (5-point Likert scale: Extremely likely - Extremely unlikely)

Grid comparison questions

1. Drag and drop to rank the four experiments by how enjoyable it was to create a password on from 1 (most enjoyable) to 4 (least enjoyable):

- Experiment 1: Choose 6 rectangles in order
- Experiment 2: Choose 6 triangles in order
- Experiment 3: Choose 8 rectangles in any order
- Experiment 4: Choose 8 triangles in any order

2. Drag and drop to rank the four experiments by how easy it was to input a password on from 1 (easiest) to 4 (hardest):

- Experiment 1: Choose 6 rectangles in order
- Experiment 2: Choose 6 triangles in order
- Experiment 3: Choose 8 rectangles in any order
- Experiment 4: Choose 8 triangles in any order

3. If I were to create a visual password, I would prefer to...

- Select fewer cells, but retrieve them in a specific order
- Select more cells, but retrieve them in any order
- I don't have a preference

4. Out of the two geometrical shapes used to create the visual passwords, I preferred using...

- Triangles
- Rectangles
- I don't have a preference

5. To what extent do you agree or disagree with the following statement: The background image influenced my choice of cells while I was creating my passwords. (5-point Likert scale: Strongly agree - Strongly disagree)

6. Out of the four visual passwords you created, which do you think are most secure?

- Passwords 1 and 2: Choose 6 rectangles/triangles in order
- Passwords 3 and 4: Choose 8 rectangles/triangles in any order

7. Out of the four visual passwords you created, which do you think is easiest to memorize?

- Password 1: Choose 6 rectangles in order
- Password 2: Choose 6 triangles in order
- Password 3: Choose 8 rectangles in any order
- Password 4: Choose 8 triangles in any order

Final questions

1. To what extent do you agree or disagree with the following statements? (5-point Likert scale: Strongly agree - Strongly disagree)

    1. I found the system simple to use
    2. I would use this system on a daily basis
    3. These passwords are easier to remember than text-based passwords
    4. I would use such a system for my low-importance accounts
    5. I would use such a system for my high-importance accounts
    6. I would prefer to use my own background image instead of a stock one

Appendix G

User study (Chapter 7): Results - Initial questionnaire

Survey answers: Participant Consent Form 1

C1 - I confirm that I have read and understood the Participant Information Sheet above, that I have had the opportunity to contact the researchers, and that any questions I had were answered to my satisfaction.

C2 - I understand that my participation is voluntary, and that I can withdraw at any time without giving a reason, by closing the browser. Withdrawing will not affect any of my rights. If I choose to withdraw, no partially gathered data will be stored.

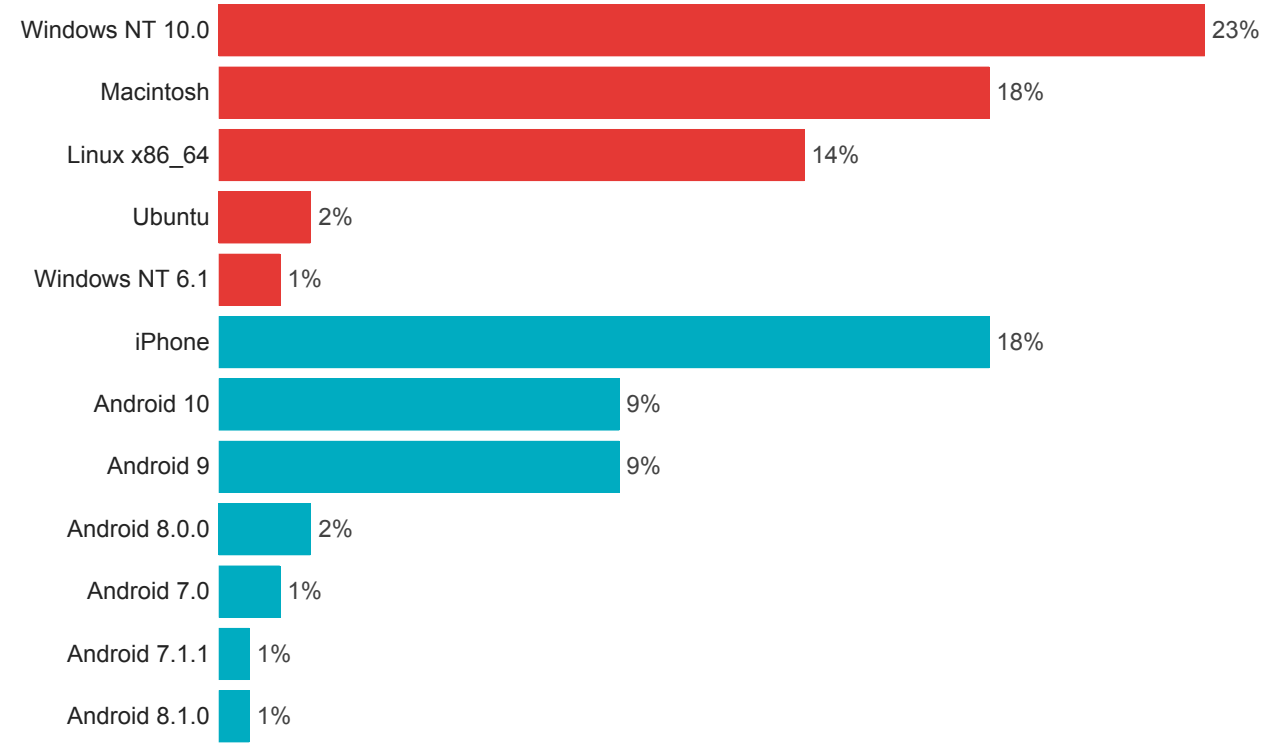
C3 - I consent to my anonymized data being used in academic publications and presentations.

C4 - I understand that my anonymized data can be stored for a minimum of two years.

C5 - I allow my data to be used in future ethically approved research.

C6 - I agree to take part in this study.

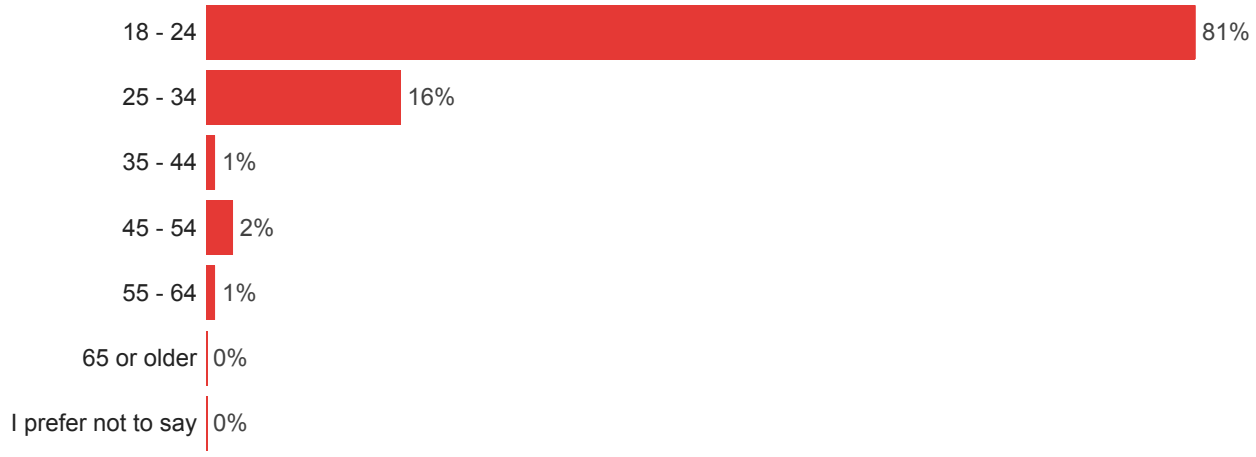
Browser Metadata Information - Operating System



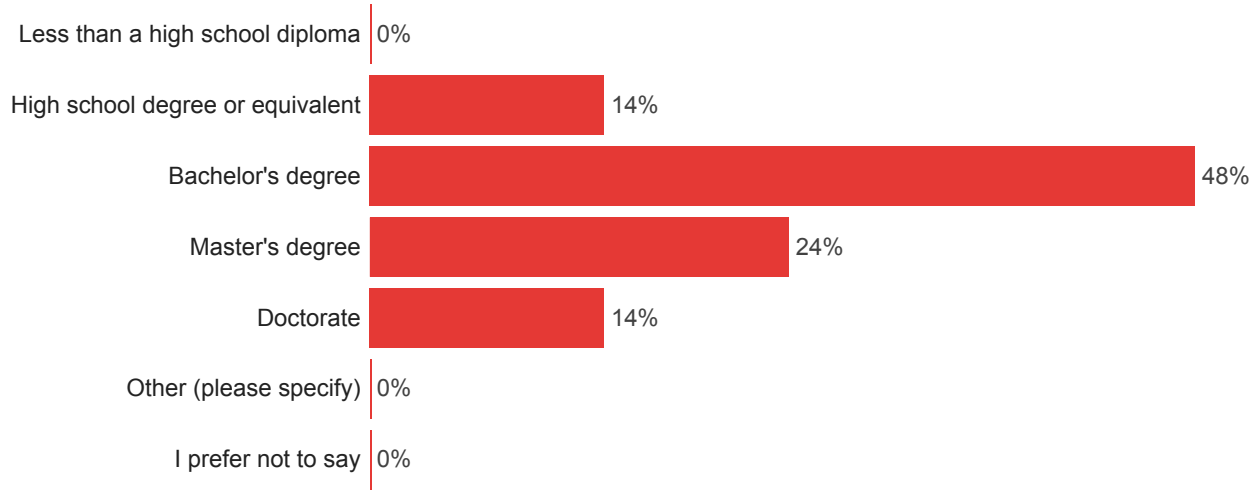
Red: Desktop devices — Total count: 81 — Total percentage: 58%
Blue: Mobile devices — Total count: 58 — Total percentage: 42%

Survey answers: Demographics

D1 - What is your age?



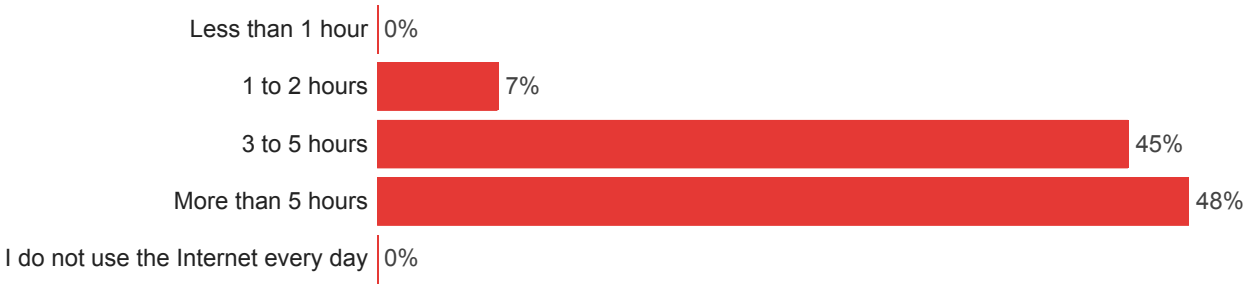
D2 - What is the highest degree or level of education you have completed? If you are currently enrolled in school, please indicate the highest degree you will receive.



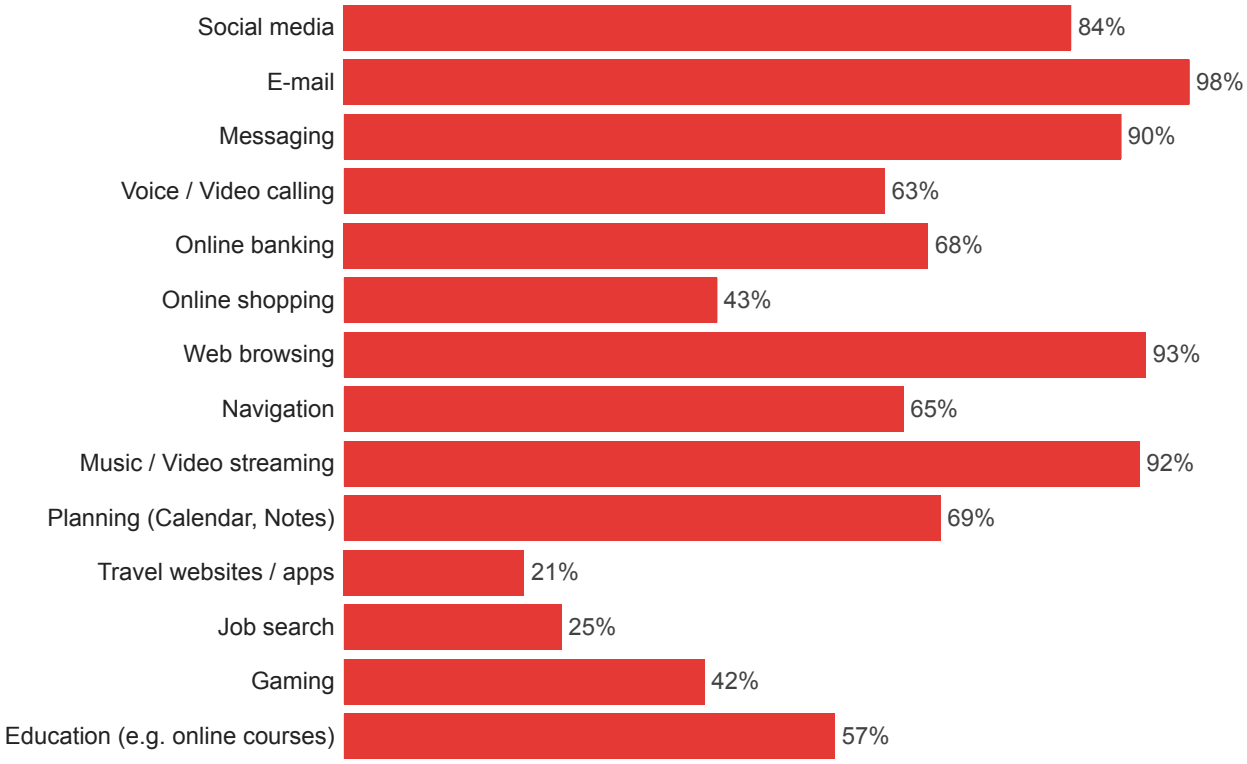
Survey answers: Familiarity with technology

T1 - On average, how much time do you spend on

Internet-related activities (e.g. e-mail, browsing, social media etc.) every day?

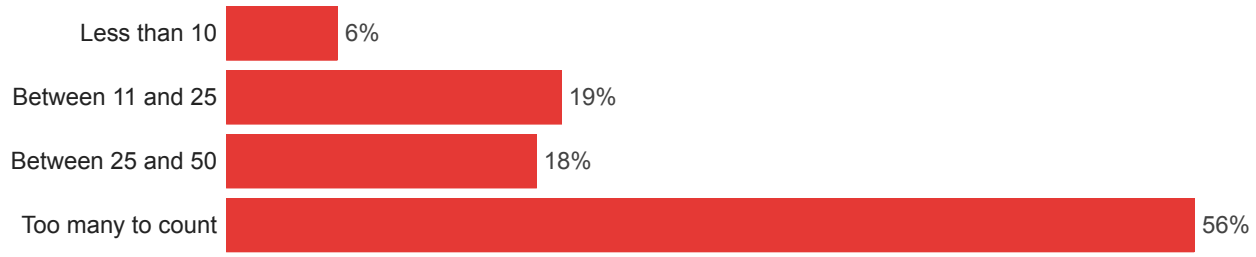


T2 - Which of these Internet-related activities do you typically do every week? Select all that apply:

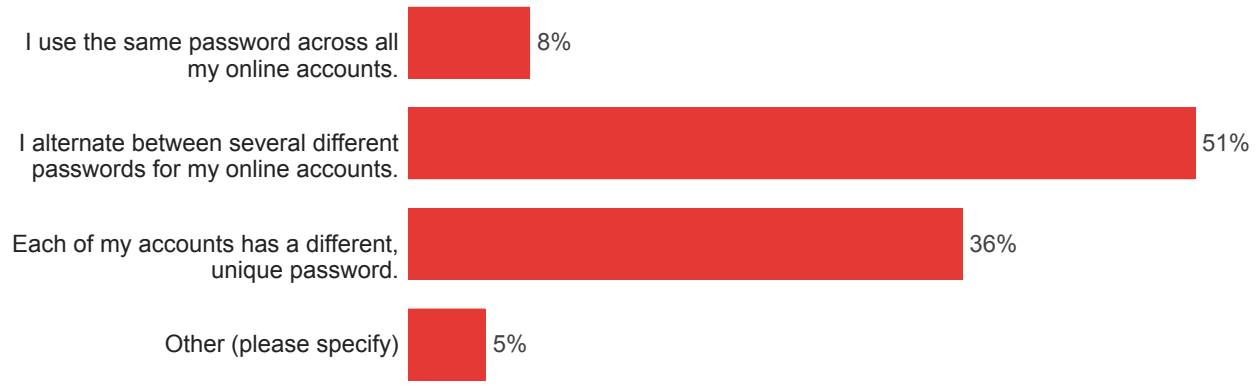


Survey answers: Current password management habits

P1 - How many password-protected online accounts do you have?



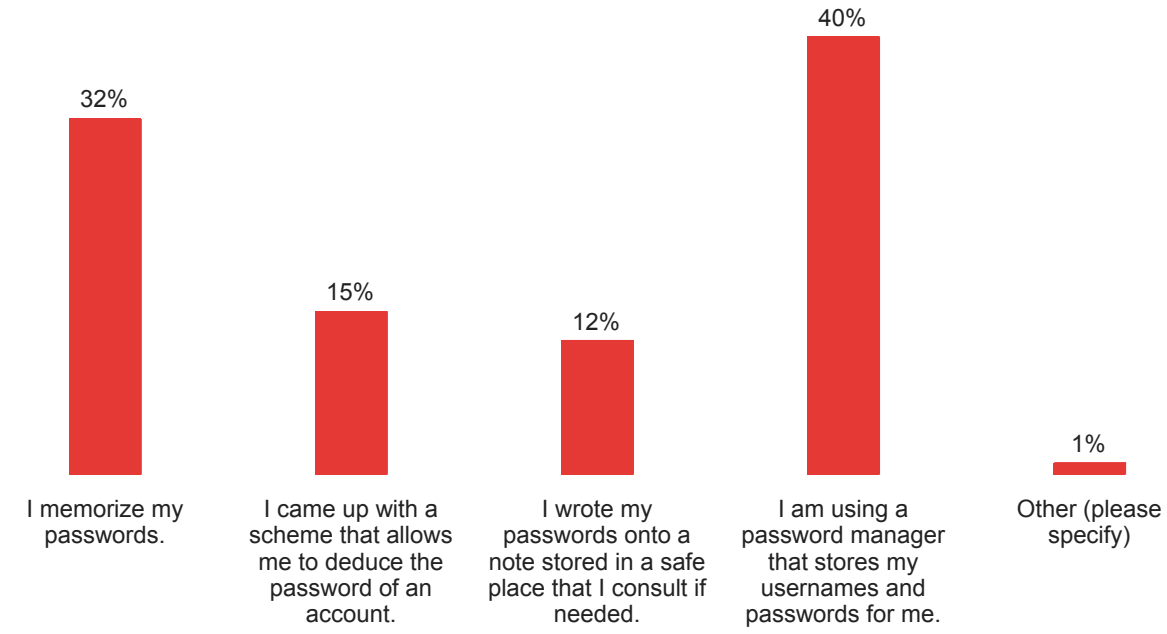
P2 - Which statement concerning password behaviour best represents your current situation? Select one of the following:



Other (please specify) - Text
A mixture of option 2 (alternating between several) and 3 (every one different). The passwords I have had before I started using a password manager were always taken from a pool of several passwords, but not unique, and I haven't yet changed these passwords. The ones after I started using the manager are all different.
Between 2 - 3, depending on the account importance and website reputation
I use about five variations of one password for all my accounts
Major accounts have the same long password that I remember. Most other accounts have different, unique passwords.
Majority of my accounts have a different unique password. Some accounts alternate between various versions of the 3/4 different passwords
Some have unique passwords others alternate (sort of a transition phase from the latter to the former)
it depends on the account, some website requires capital letters or symbols.

Survey answers: Current password management habits

P3 - How do you keep track of your passwords? Select all that apply:

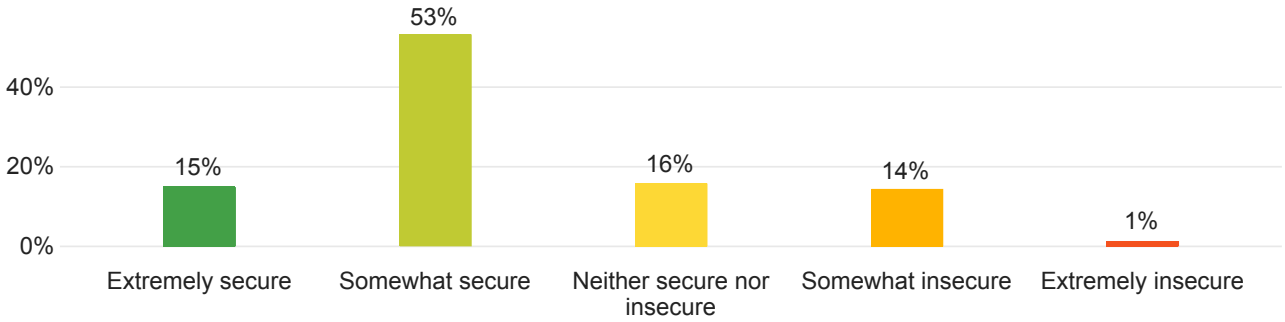


Other (please specify) - Text

My note is on my computer and it doesn't explicitly say the password, but allows me to remember it
using a protected document to store personalised hints for my passwords

P4 - What is your most important consideration when creating a password for an online account? Select one of the following:

P5 - How secure are you feeling with your current password management habits?

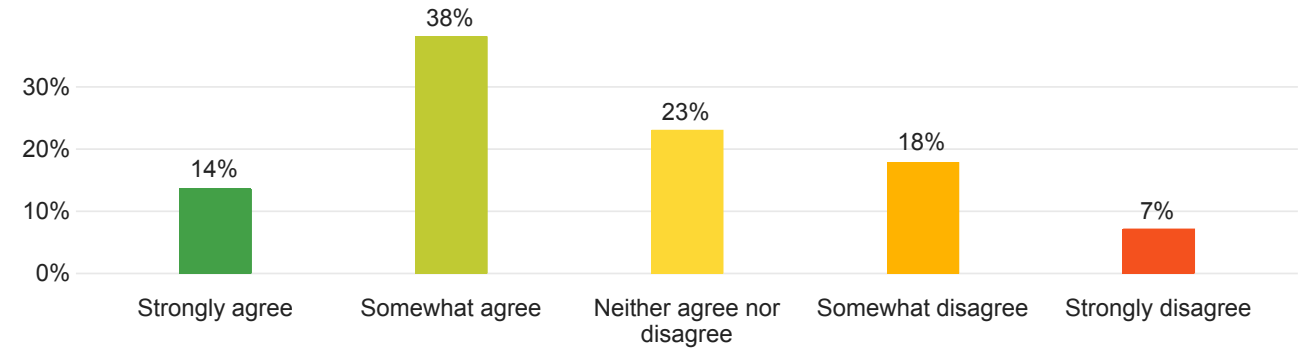


Appendix H

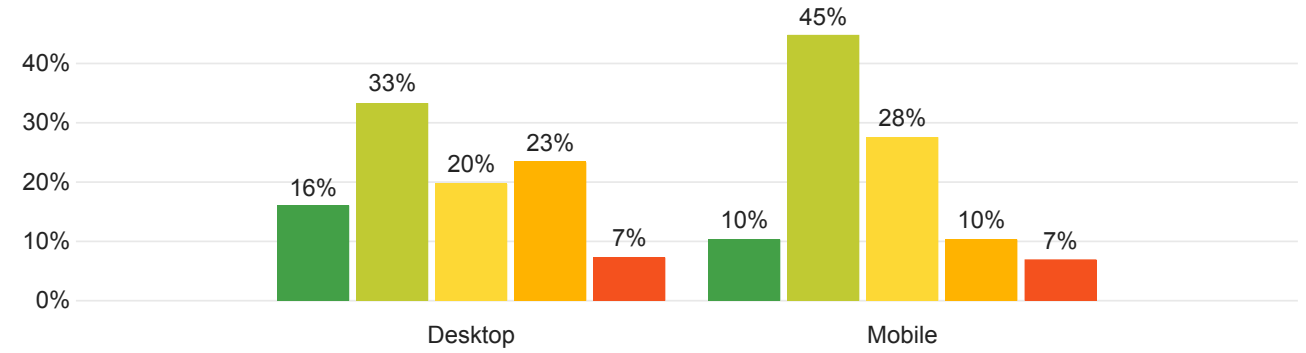
User study (Chapter 7): Results - Experiments

Survey answers: Individual experiments, Q-by-Q 1

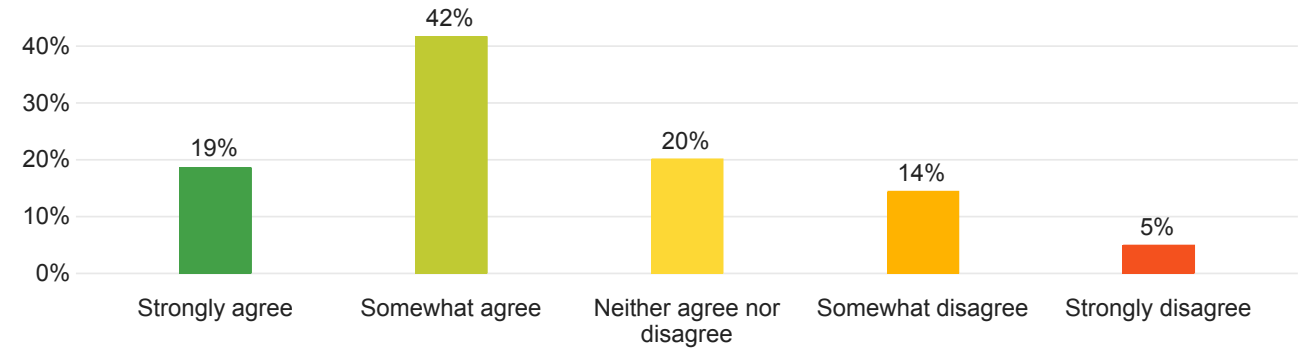
Experiment #1 - Creating this password was enjoyable (Overall)



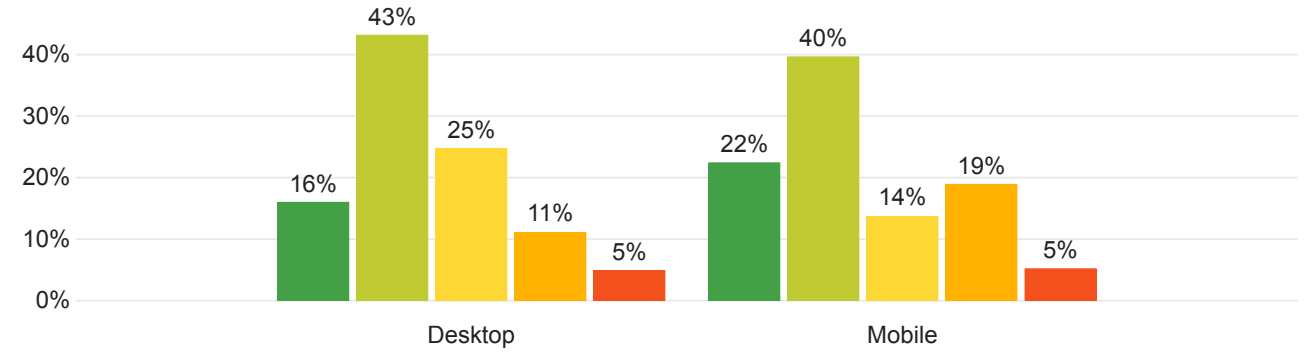
Experiment #1 - Creating this password was enjoyable (Mobile vs Desktop)



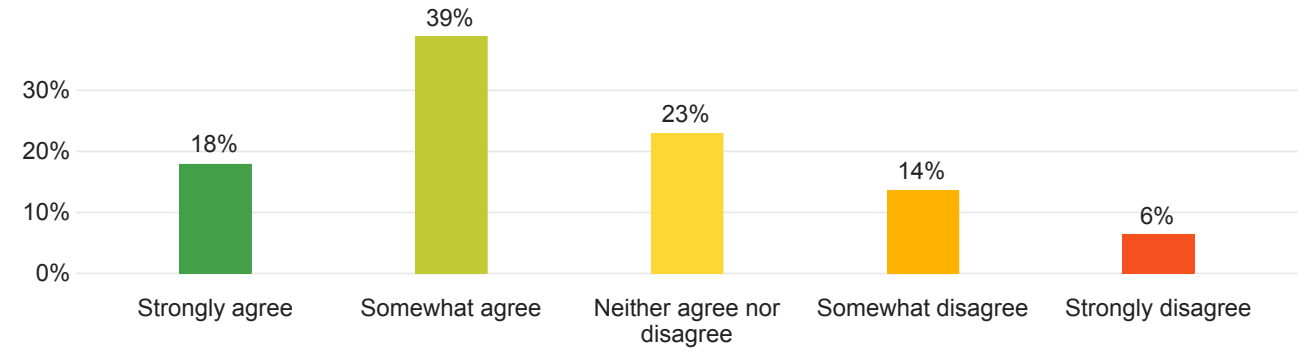
Experiment #2 - Creating this password was enjoyable (Overall)



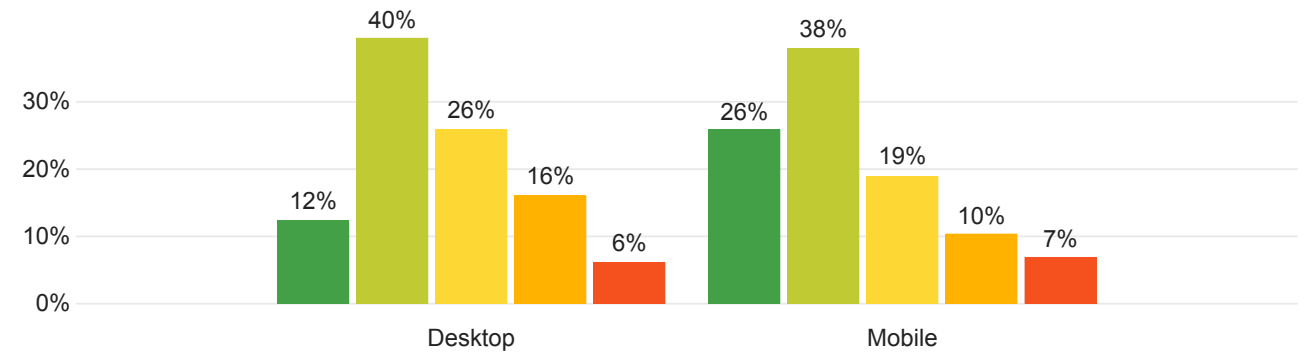
Experiment #2 - Creating this password was enjoyable (Mobile vs Desktop)



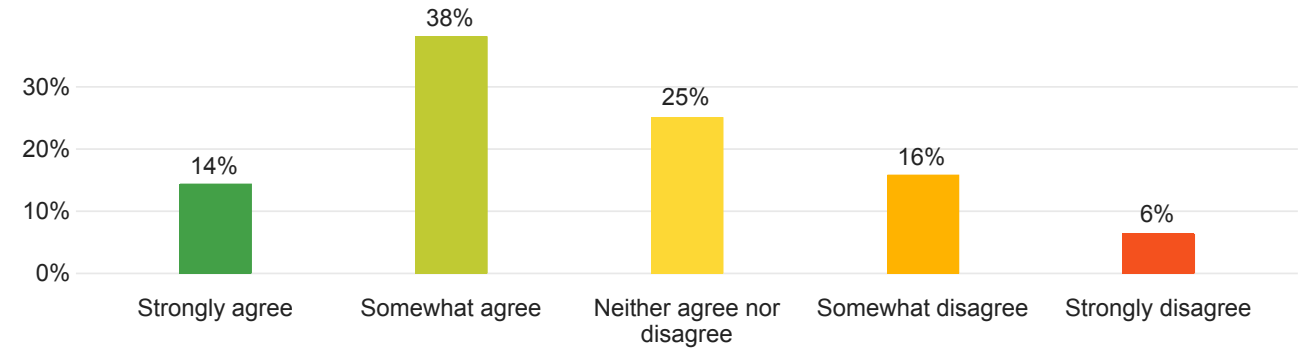
Experiment #3 - Creating this password was enjoyable (Overall)



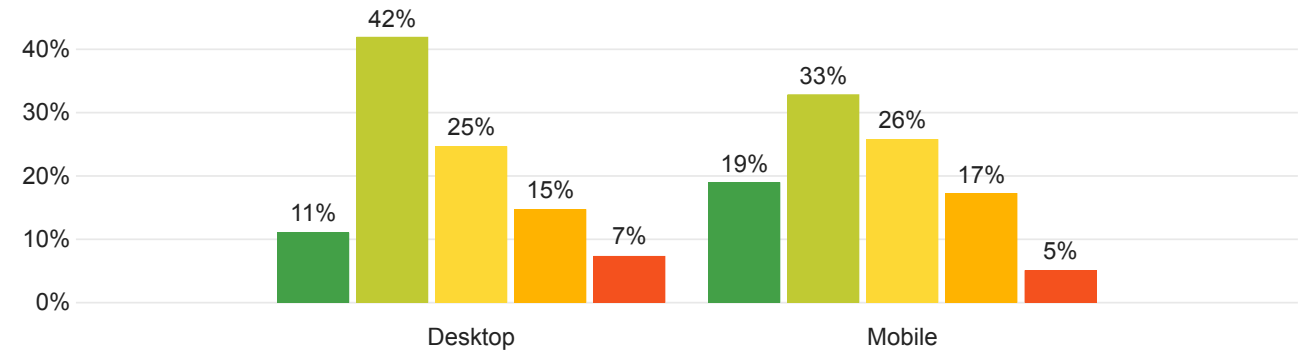
Experiment #3 - Creating this password was enjoyable (Mobile vs Desktop)



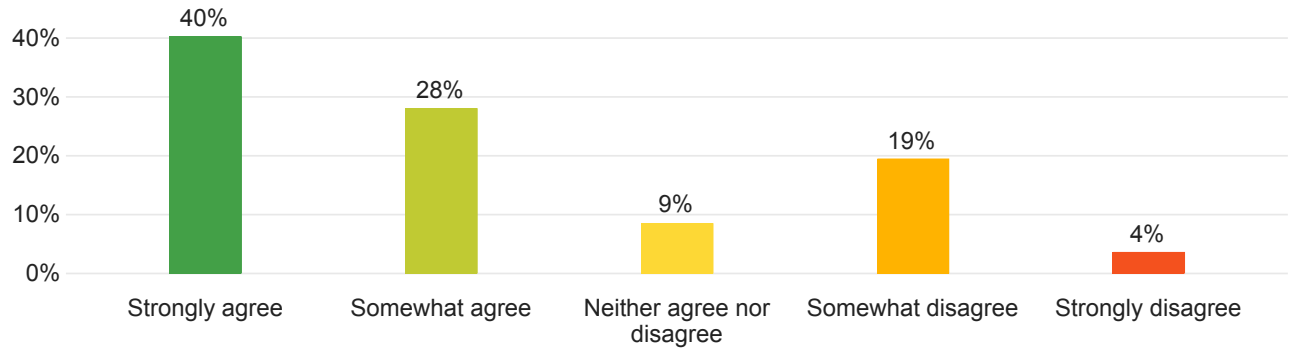
Experiment #4 - Creating this password was enjoyable (Overall)



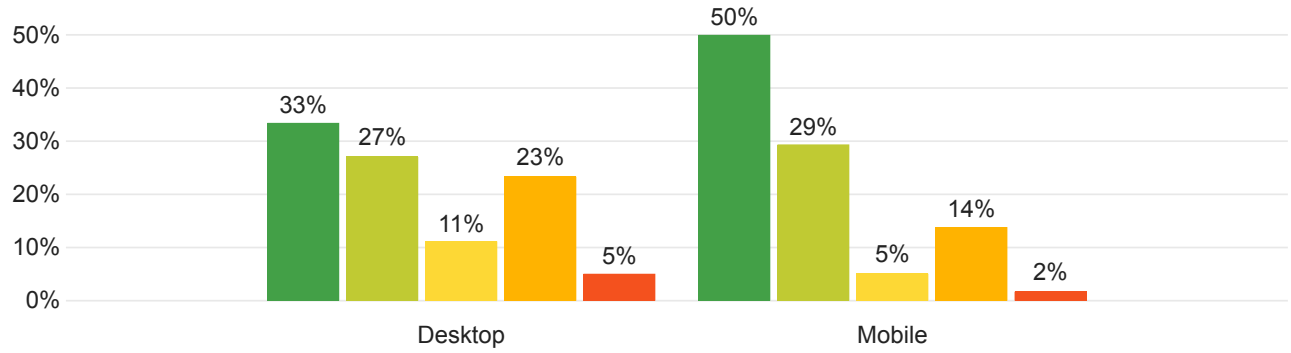
Experiment #4 - Creating this password was enjoyable (Mobile vs Desktop)



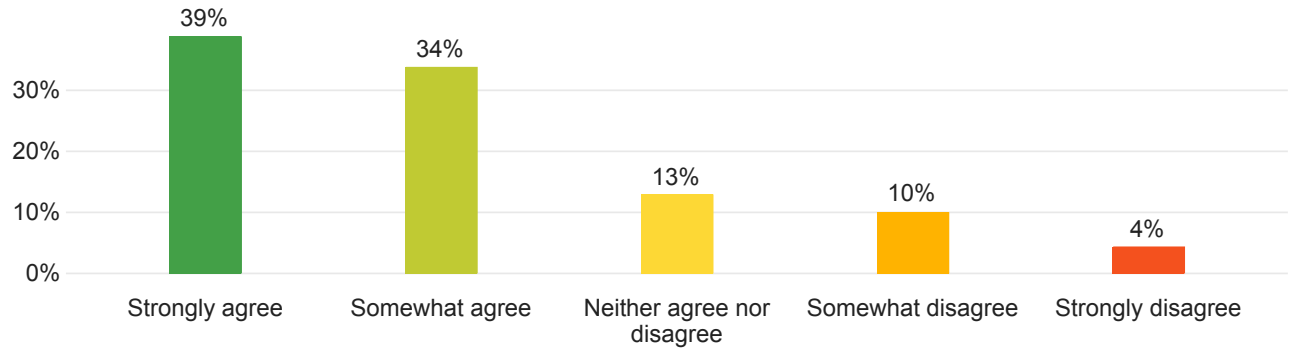
Experiment #1 - Inputting this password was easy (Overall)



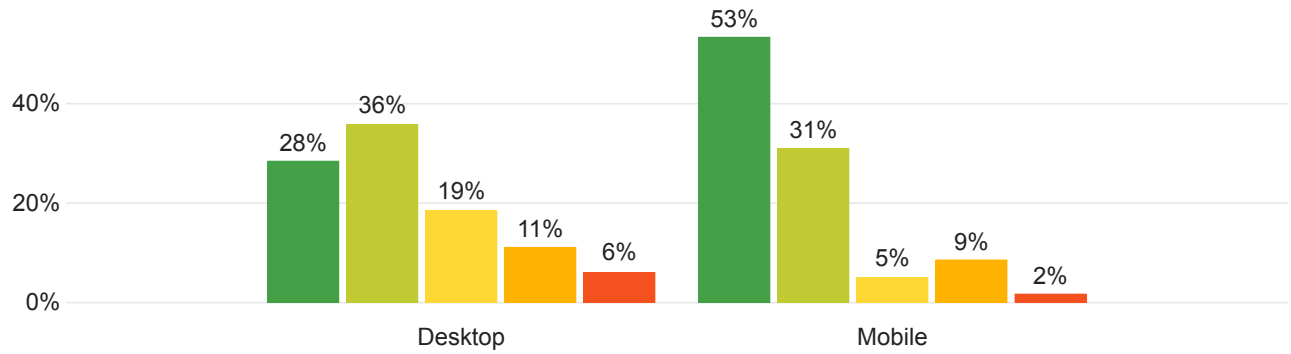
Experiment #1 - Inputting this password was easy (Mobile vs Desktop)



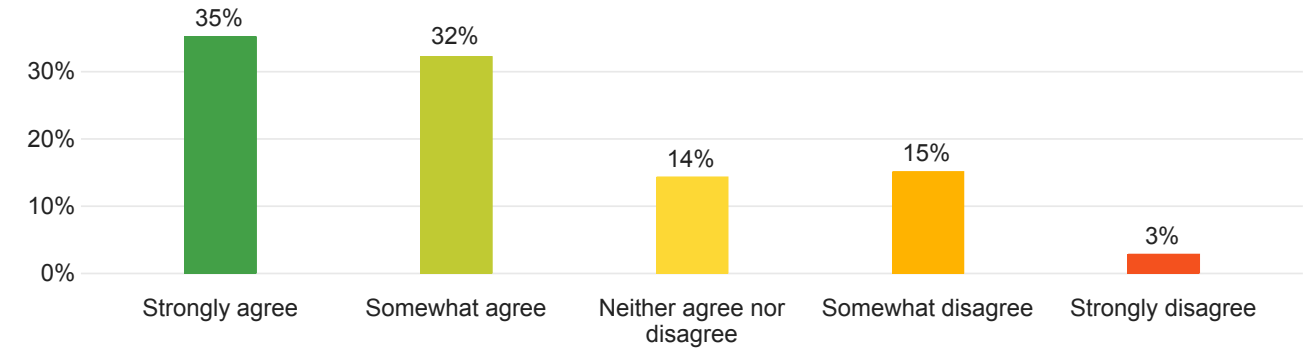
Experiment #2 - Inputting this password was easy (Overall)



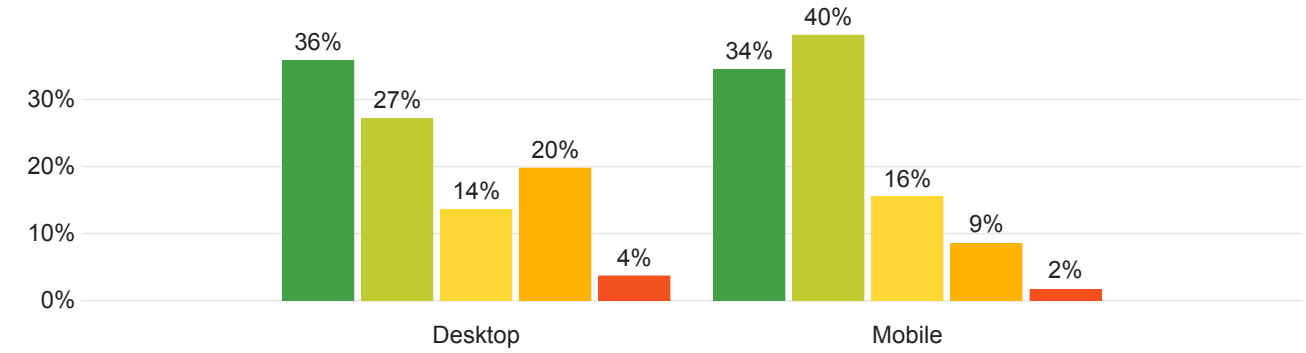
Experiment #2 - Inputting this password was easy (Mobile vs Desktop)



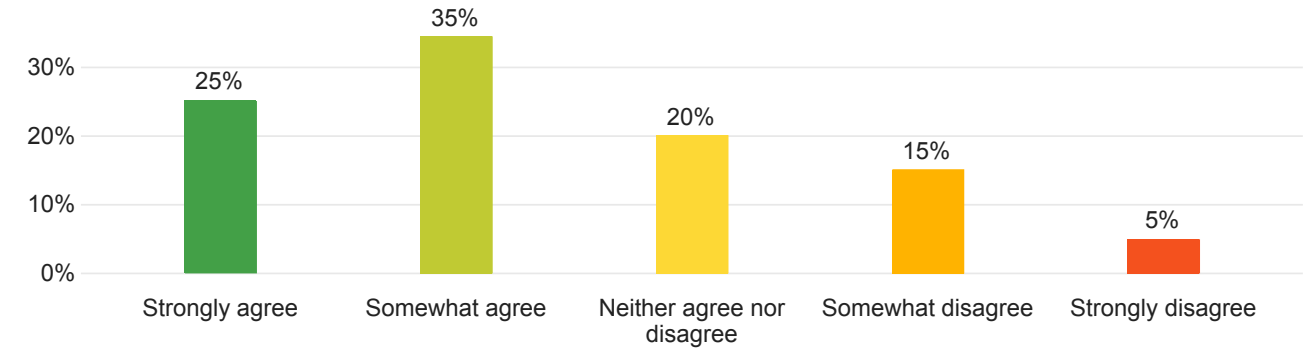
Experiment #3 - Inputting this password was easy (Overall)



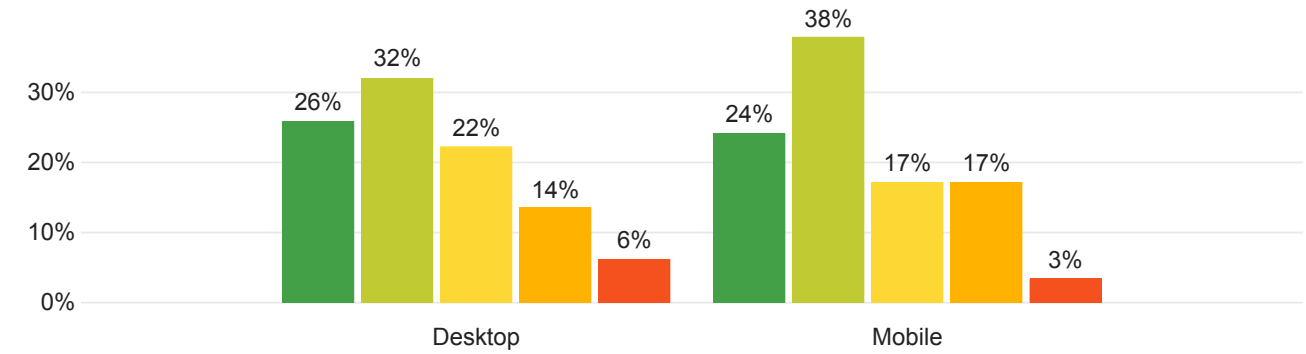
Experiment #3 - Inputting this password was easy (Mobile vs Desktop)



Experiment #4 - Inputting this password was easy (Overall)

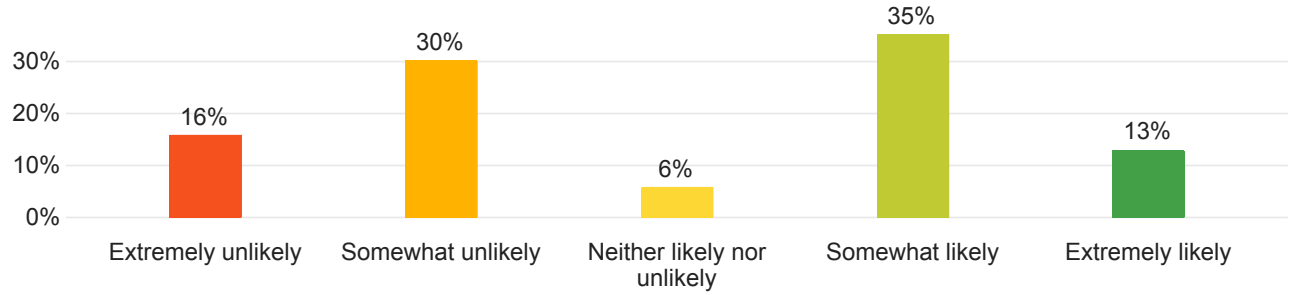


Experiment #4 - Inputting this password was easy (Mobile vs Desktop)

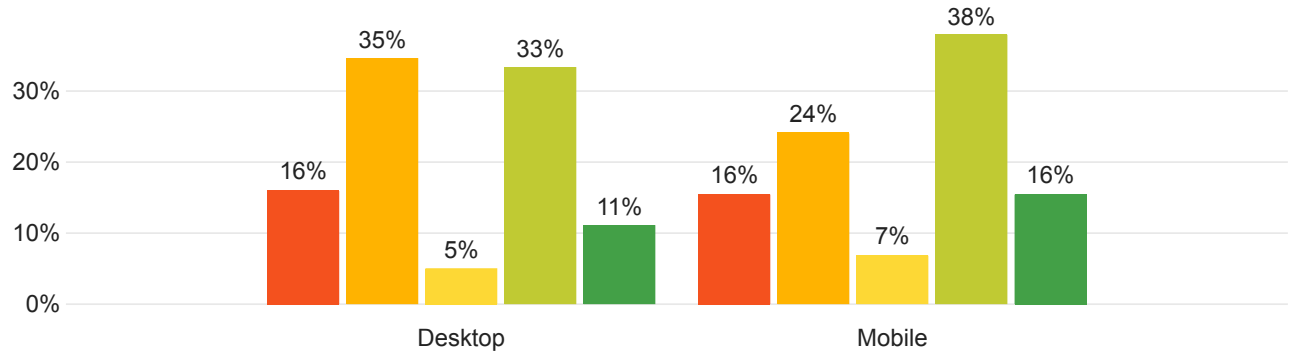


Survey answers: Individual experiments, Q-by-Q 5

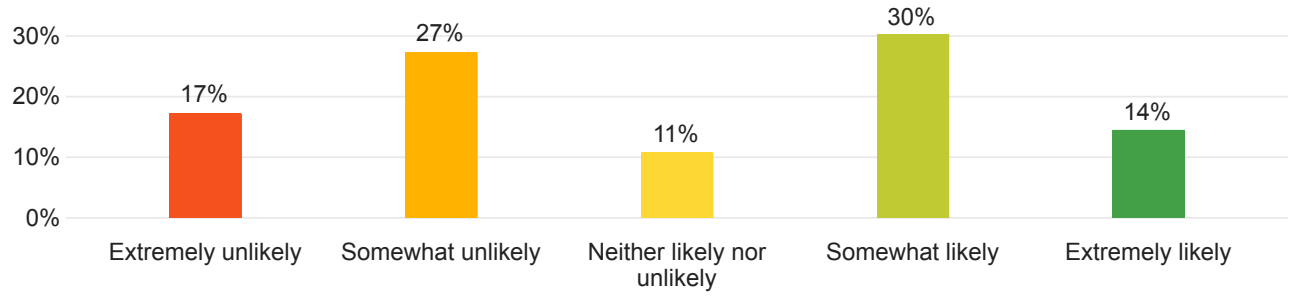
Experiment #1 - How likely would you be to remember this password in a week's time? (Overall)



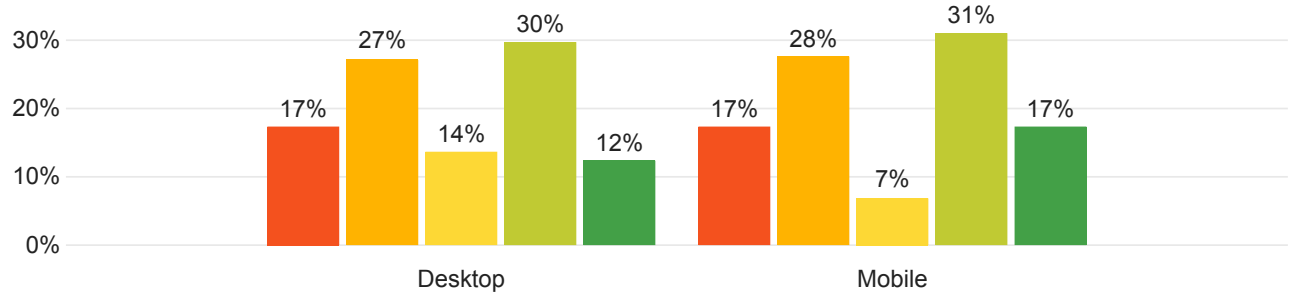
Experiment #1 - How likely would you be to remember this password in a week's time? (Mobile vs Desktop)



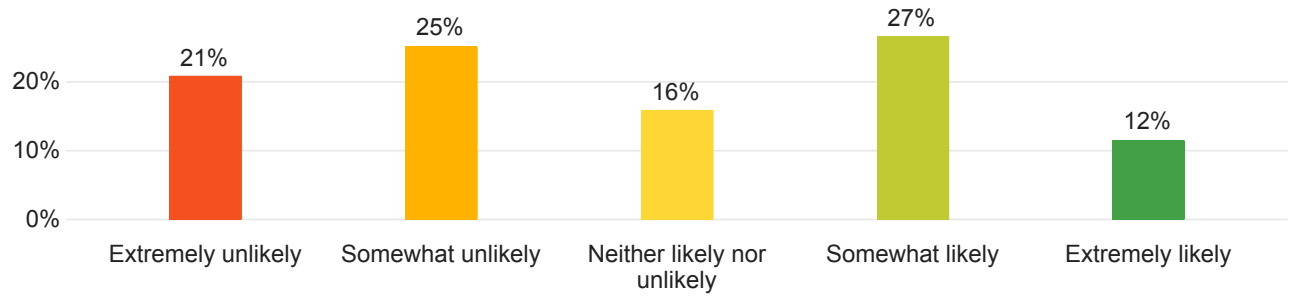
Experiment #2 - How likely would you be to remember this password in a week's time? (Overall)



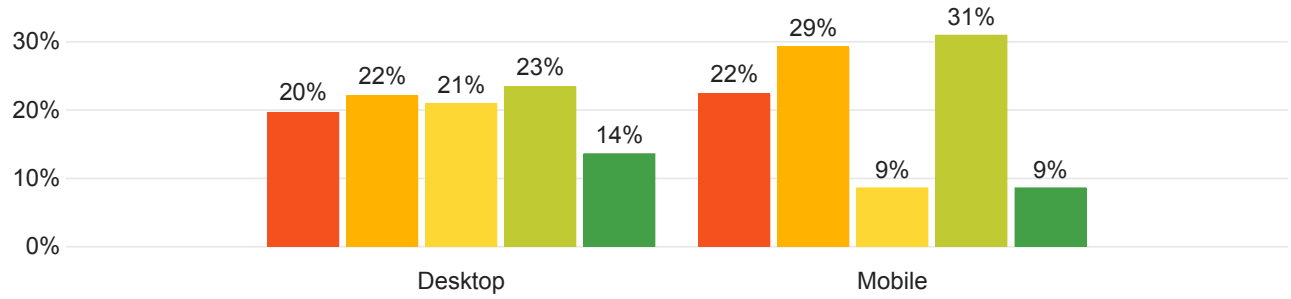
Experiment #2 - How likely would you be to remember this password in a week's time? (Mobile vs Desktop)



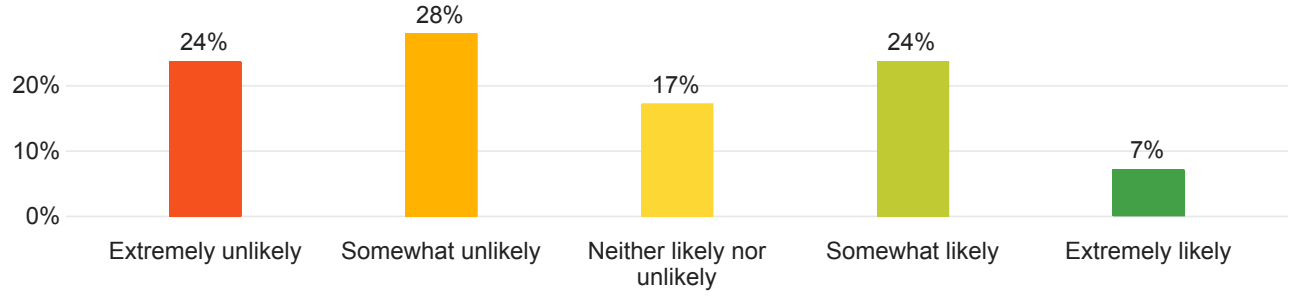
Experiment #3 - How likely would you be to remember this password in a week's time? (Overall)



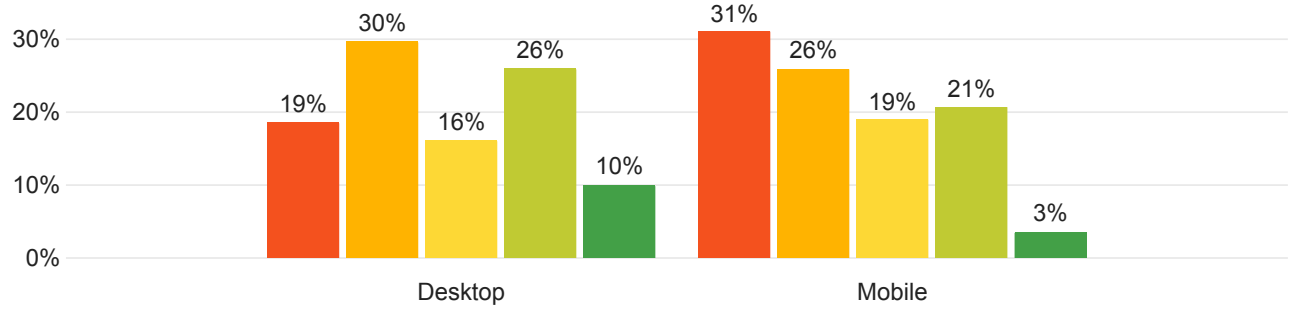
Experiment #3 - How likely would you be to remember this password in a week's time? (Mobile vs Desktop)



Experiment #4 - How likely would you be to remember this password in a week's time? (Overall)



Experiment #4 - How likely would you be to remember this password in a week's time? (Mobile vs Desktop)

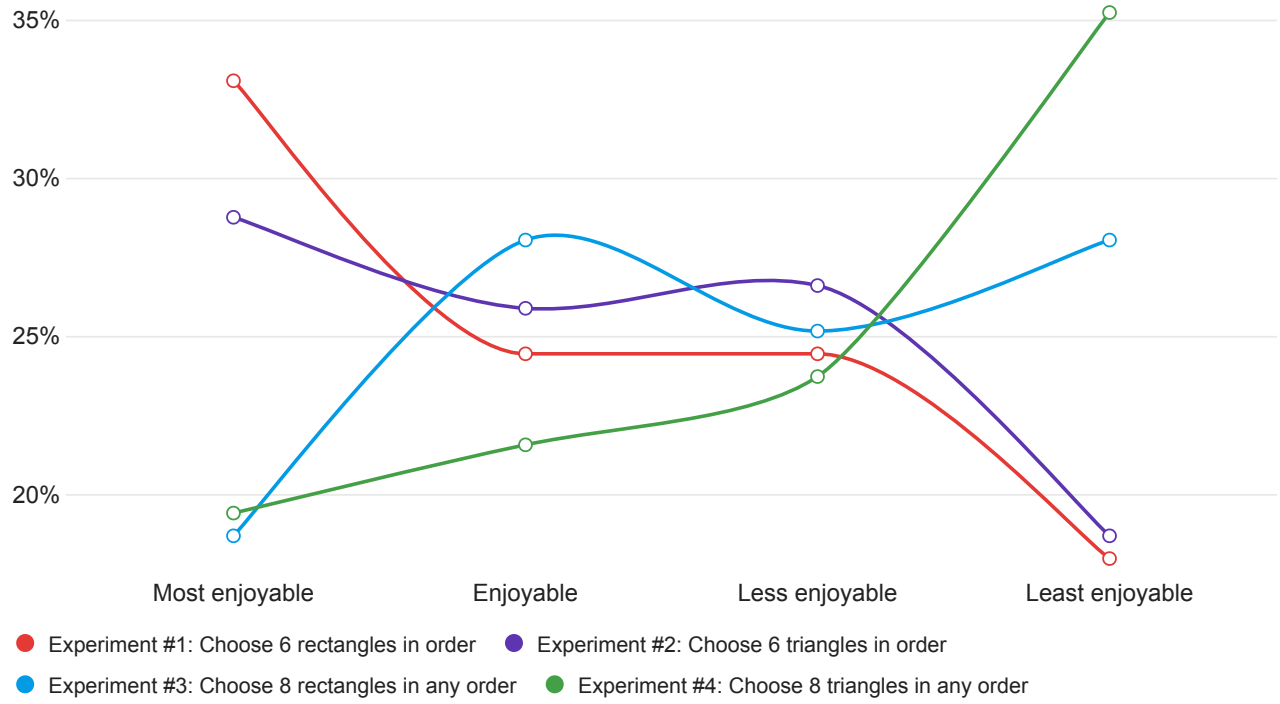


Appendix I

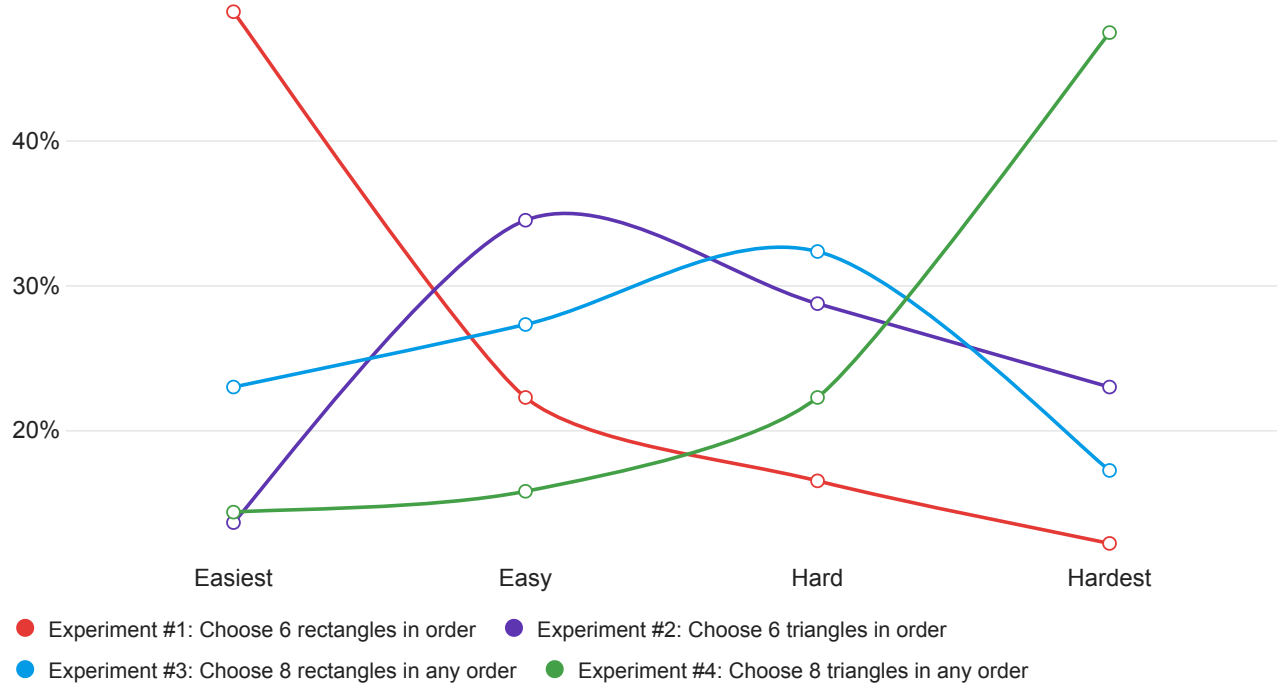
User study (Chapter 7): Results - Final questionnaire

Survey answers: Visual password system, main questions

Q1 - Drag and drop to rank the four experiments by how enjoyable it was to create a password from 1 (most enjoyable) to 4 (least enjoyable):

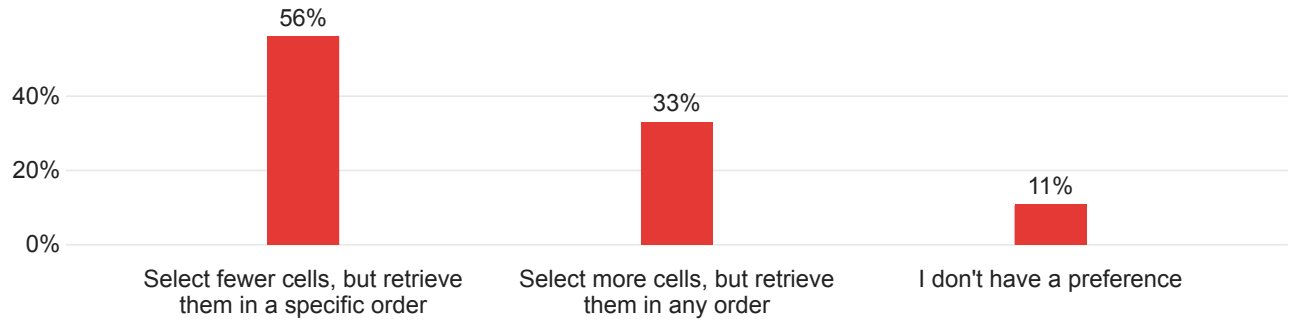


Q2 - Drag and drop to rank the four experiments by how easy it was to input a password on from 1 (easiest) to 4 (hardest):

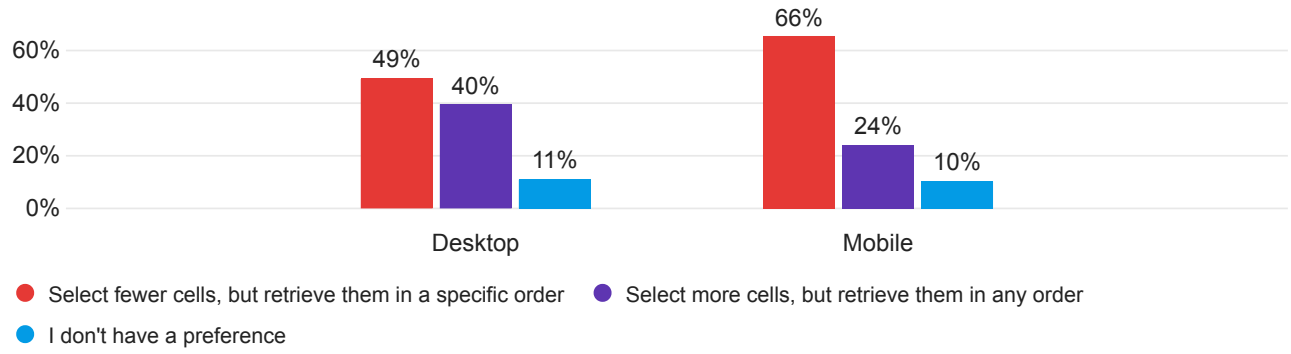


Survey answers: Visual password system, main questions

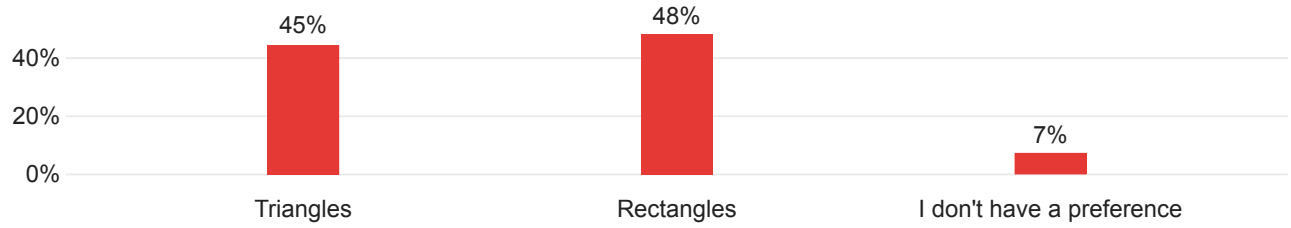
Q3 - If I were to create a visual password, I would prefer to... (Overall)



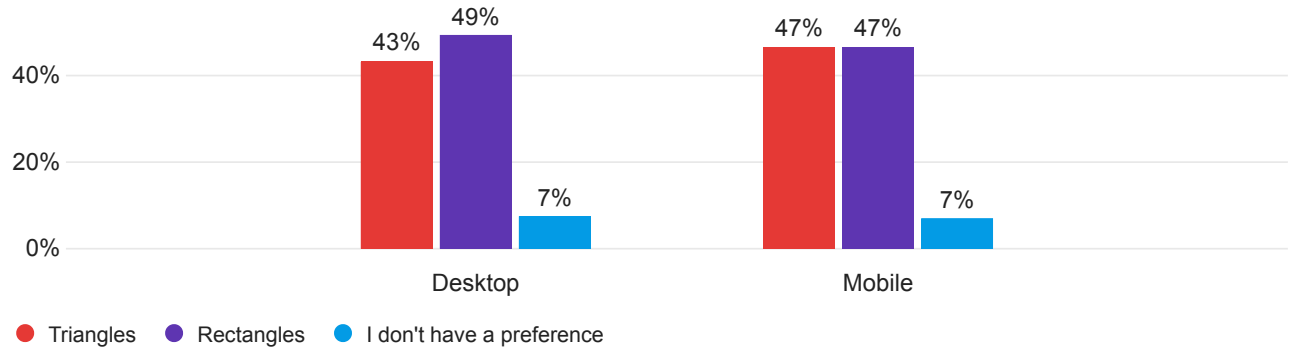
Q3 - If I were to create a visual password, I would prefer to... (Mobile vs Desktop)



Q4 - Out of the two geometrical shapes used to create the visual passwords, I preferred using... (Overall)

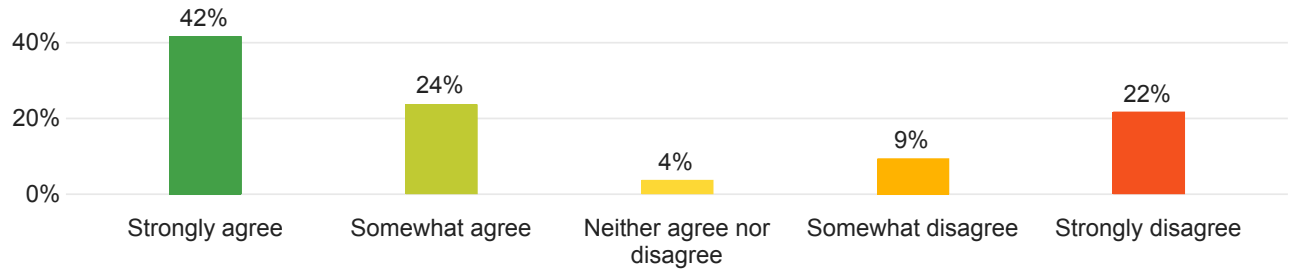


Q4 - Out of the two geometrical shapes used to create the visual passwords, I preferred using... (Mobile vs Desktop)

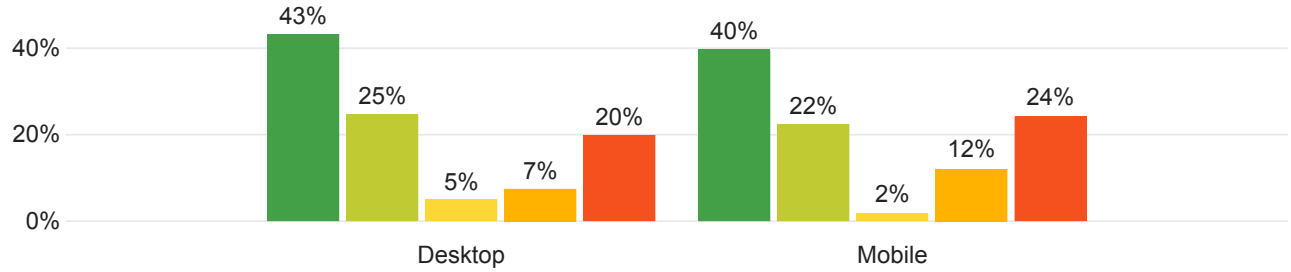


Survey answers: Visual password system, main questions

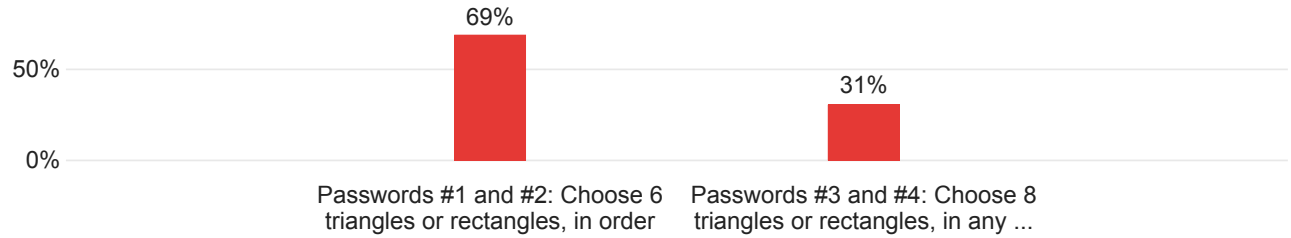
Q5 - To what extent do you agree or disagree with the following statement: "The background image influenced my choice of cells while I was creating my passwords." (Overall)



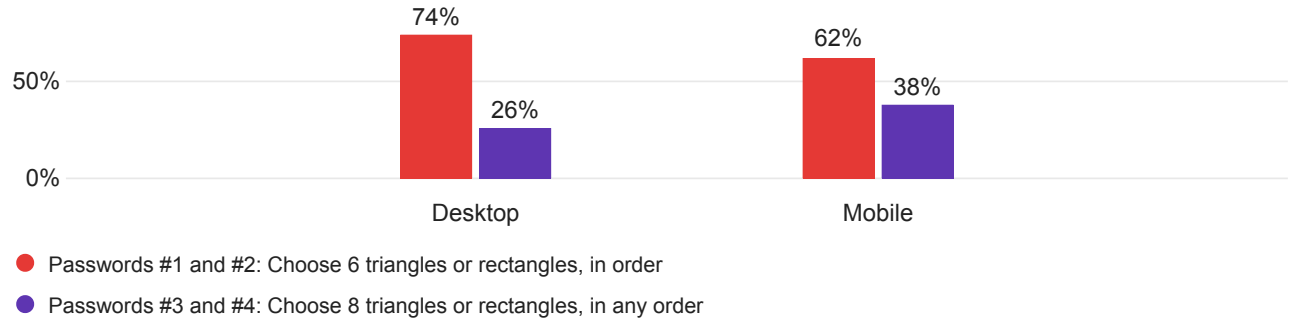
Q5 - To what extent do you agree or disagree with the following statement: "The background image influenced my choice of cells while I was creating my passwords." (Mobile vs Desktop)



Q6 - Out of the four visual passwords you created, which do you think are most secure? (Overall)

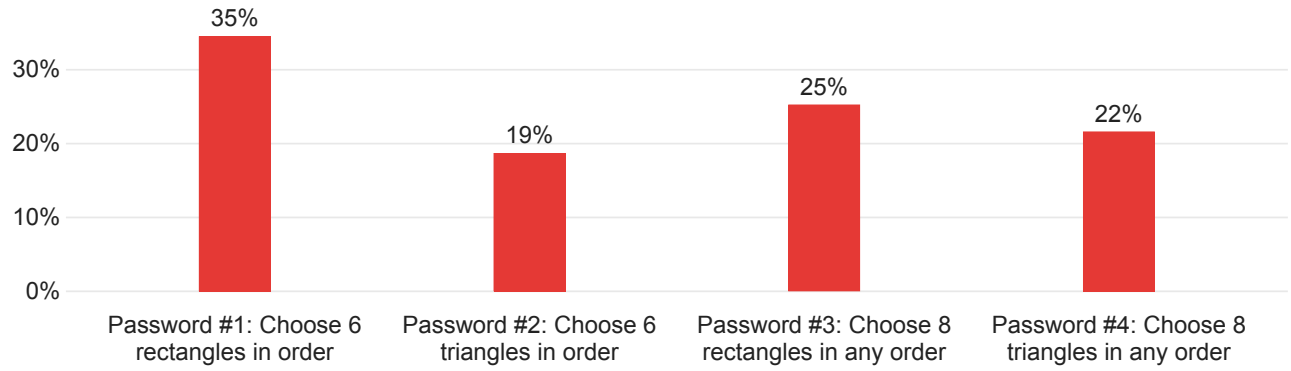


Q6 - Out of the four visual passwords you created, which do you think are most secure? (Mobile vs Desktop)

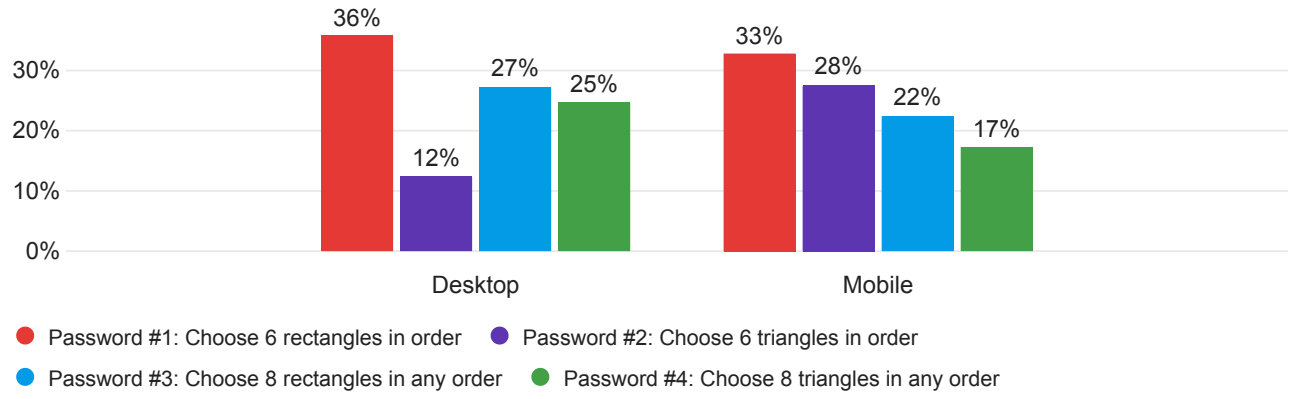


Survey answers: Visual password system, main questions

Q7 - Out of the four visual passwords you created, which do you think is easiest to memorize? (Overall)



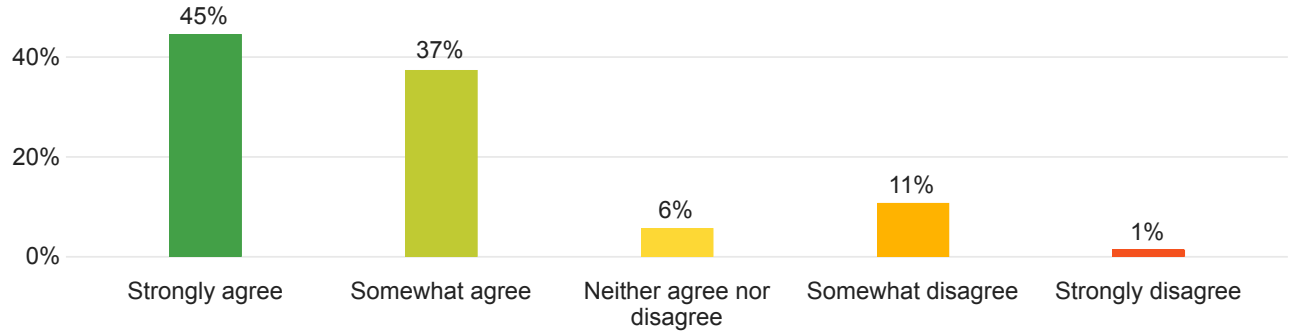
Q7 - Out of the four visual passwords you created, which do you think is easiest to memorize? (Mobile vs Desktop)



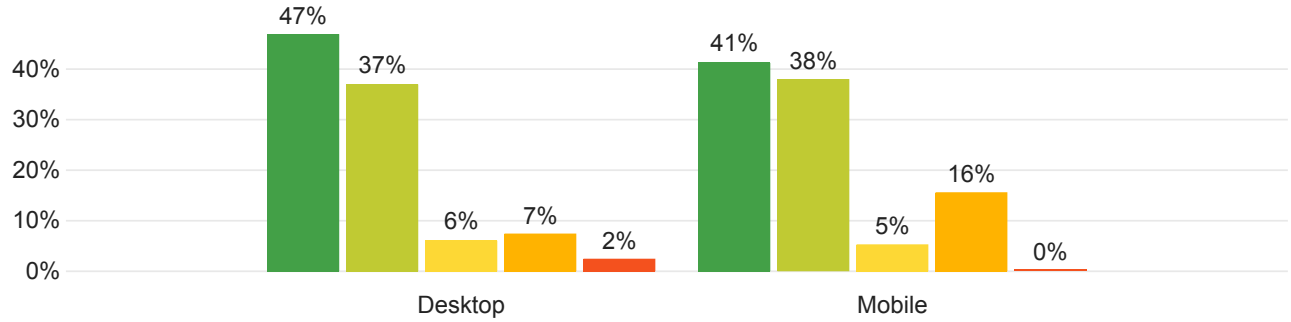
Survey answers: Visual password system, final questions

To what extent do you agree or disagree with the following statements about this visual password system?

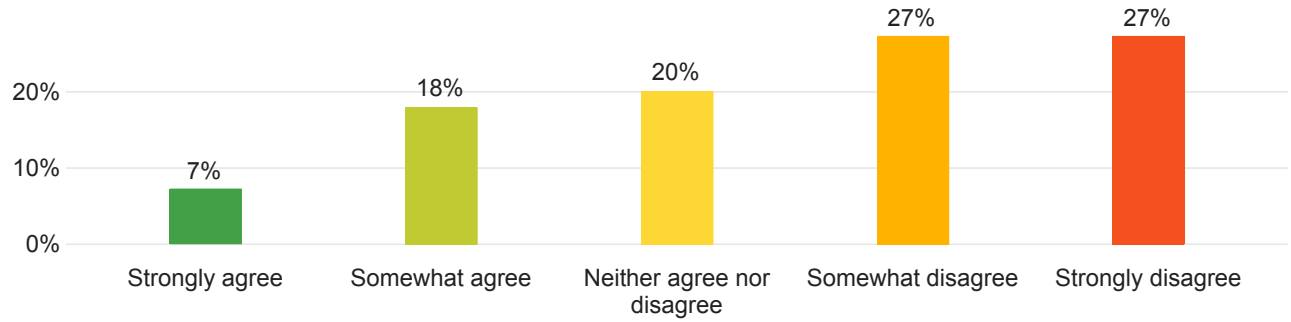
"I found the system simple to use" (Overall)



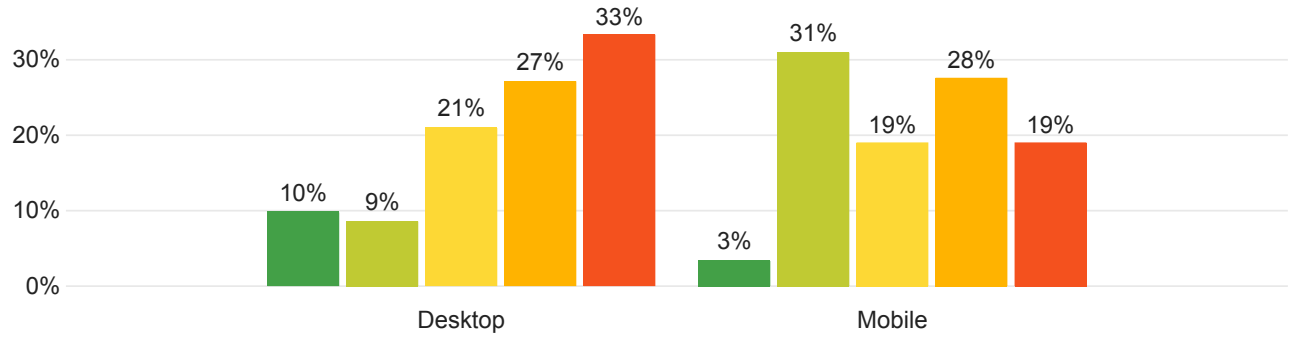
"I found the system simple to use" (Mobile vs Desktop)



"I would use this system on a daily basis" (Overall)

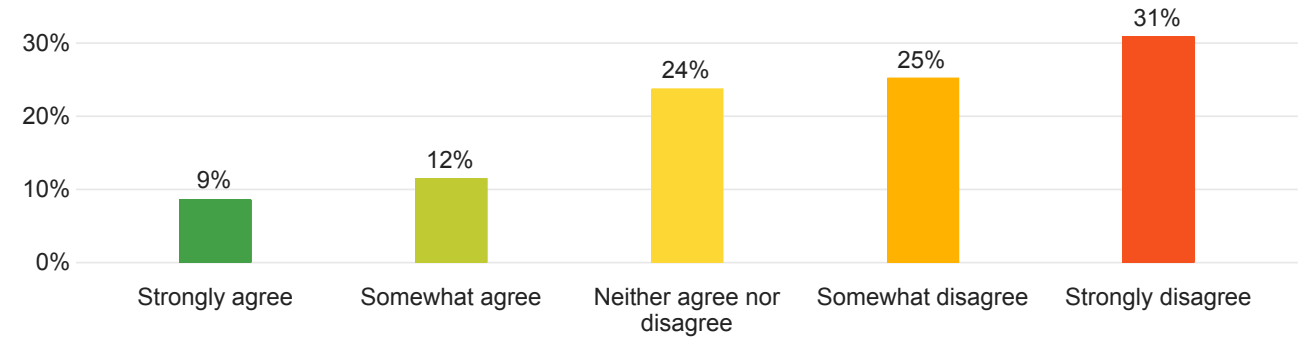


"I would use this system on a daily basis" (Mobile vs Desktop)

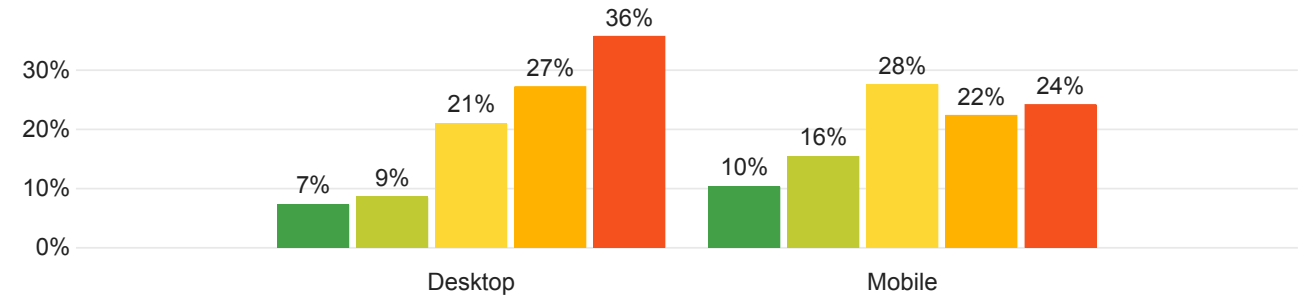


Survey answers: Visual password system, final questions

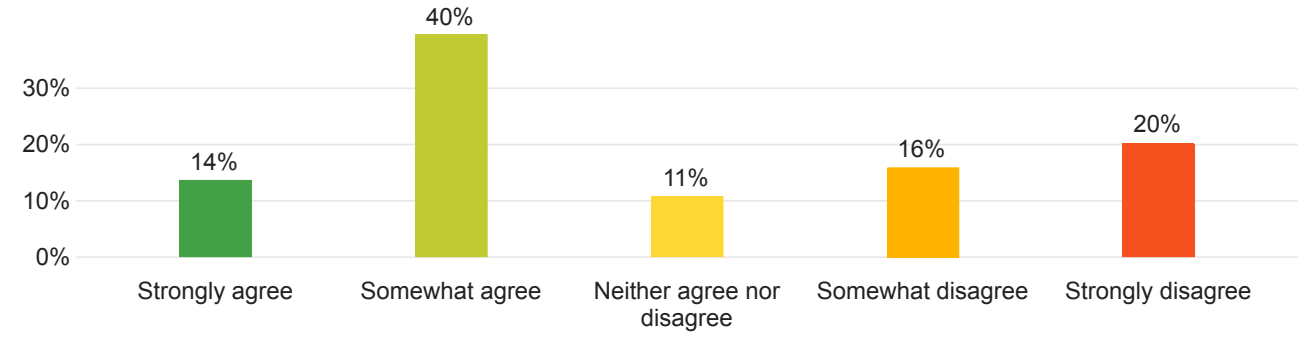
"These passwords are easier to remember than text-based passwords" (Overall)



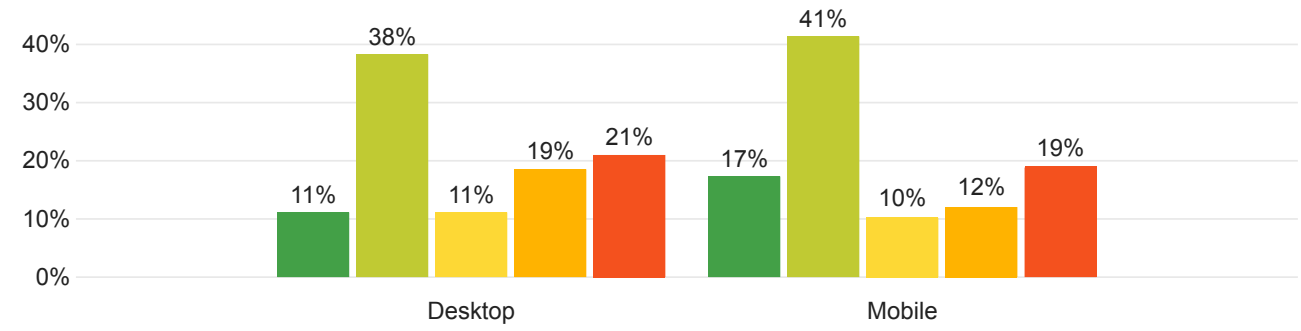
"These passwords are easier to remember than text-based passwords" (Mobile vs Desktop)



"I would use such a system for my low-importance accounts" (Overall)

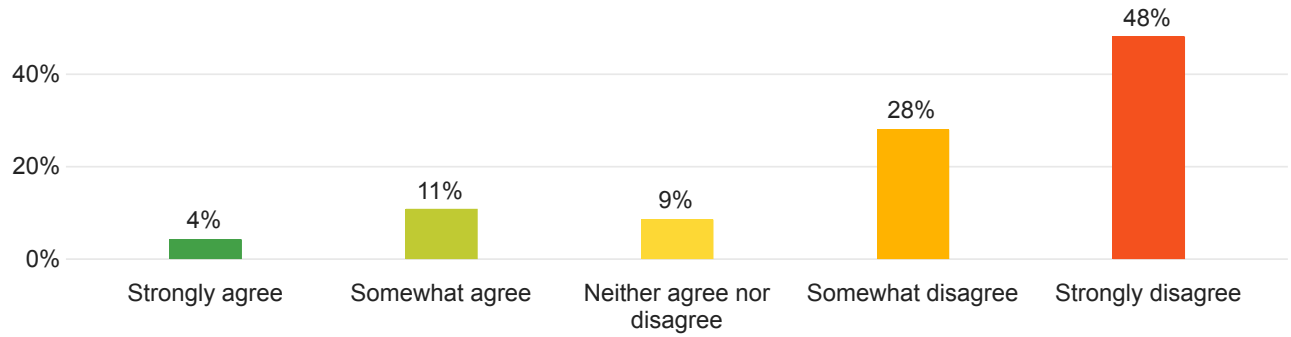


"I would use such a system for my low-importance accounts" (Mobile vs Desktop)

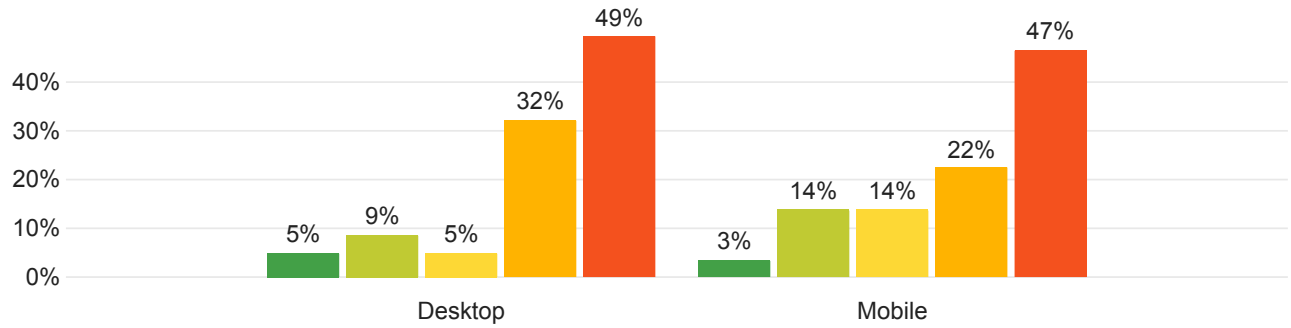


Survey answers: Visual password system, final questions

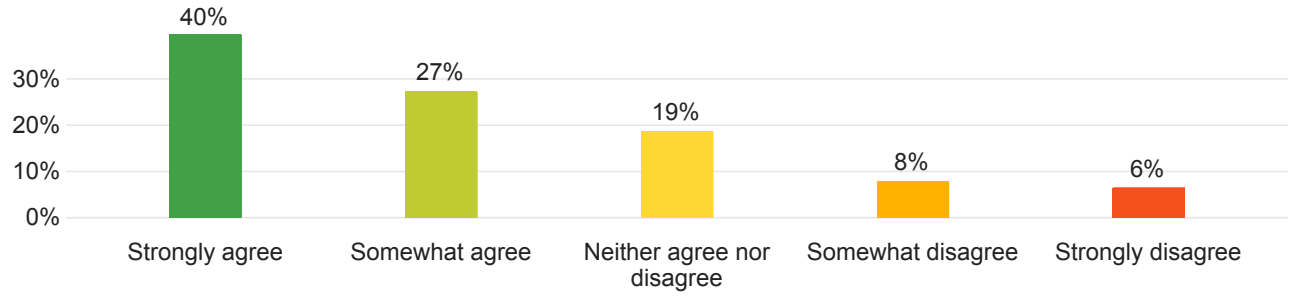
"I would use such a system for my high-importance accounts" (Overall)



"I would use such a system for my high-importance accounts" (Mobile vs Desktop)



"I would prefer to use my own background image instead of a stock one" (Overall)



"I would prefer to use my own background image instead of a stock one" (Mobile vs Desktop)