An Ecologically Valid Evaluation of an Observation-Resilient Graphical Authentication Mechanism

Joseph Noel Maguire

Submitted in fulfilment of the requirements for the Degree of Doctor of Philosophy

School of Computing Science
College of Science and Engineering
University of Glasgow
November 2, 2013

Abstract

Alphanumeric authentication, by means of a secret, is not only a powerful mechanism, in theory, but prevails over all its competitors in reality. Passwords, as they are more commonly known, have the potential to act as a fairly strong gateway. In practice, though, password usage is problematic. They are (1) easily shared, (2) trivial to observe and (3) maddeningly elusive when forgotten. Moreover, modern consumer devices only exacerbate the problems of passwords as users enter them in shared spaces, in plain view, on television screens, on smartphones and on tablets. Asterisks may obfuscate alphanumeric characters on entry but popular systems, e.g. Apple iPhone and Nintendo Wii, require the use of an on-screen keyboard for character input.

A number of alternatives to passwords have been proposed but none, as yet, have been adopted widely. There seems to be a reluctance to switch from tried and tested passwords to novel alternatives, even if the most glaring flaws of passwords can be mitigated. One argument is that there has not been sufficient investigation into the feasibility of the password alternatives and thus no convincing evidence that they can indeed act as a viable alternative.

Graphical authentication mechanisms, solutions that rely on images rather than characters, are a case in point. Pictures are more memorable than the words that name them, meaning that graphical authentication mitigates one of the major problems with passwords. This dissertation sets out to investigate the feasibility of one particular observation-resilient graphical authentication mechanism called Tetrad. The authentication mechanism attempted to address two of the core problems with passwords: improved memorability and resistance to observability (with on-screen entry).

Tetrad was tested in a controlled lab study, that delivered promising results and was well received by the evaluators. It was then deployed in a realistic context and its viability tested in three separate field tests. The unfortunate conclusion was that Tetrad, while novel and viable in a lab setting, failed to deliver a usable and acceptable experience to the end users. This thorough testing of an alternative authentication mechanism is unusual in this research field and the outcome is disappointing. Nevertheless, it acts to inform inventors of other authentication mechanisms of the problems that can manifest when a seemingly viable authentication mechanism is tested in the wild.

Contents

Chapter 1

Introduction

Password problems are the stuff of legend, as is evident from the tale below:

> Two brothers lived in a small town. Cassim was wealthy and wanted for nothing, Ali Baba was poor and cut wood tirelessly to support his family. While cutting wood in the forest one day, Ali Baba spotted a flock of fierce men on horses heading his way and so hid in a nearby tree. The leader walked right up to the tree and uttered a simple, secret word. A huge rock rolled back to reveal a vast cave, covered in treasure.
>
> Stunned and shaken, Ali Baba stayed silent in the tree as the men entered the cave and the rock rolled shut. The men eventually left, at which point Ali Baba leaped from the tree and spoke the simple, secret word. The rock rolled back and Ali Baba grabbed as much treasure as he could and ran back to his family.
>
> Cassim, shocked by Ali Baba's sudden wealth, demanded answers. Ali Baba recanted the story and agreed with his brother to return in the morning. Cassim, always the greedier, set out earlier. Once at the cave, Cassim spoke the simple, secret word and, to his amazement, the rock rolled back. Cassim entered and the rock rolled shut; he began grabbing as much treasure as possible. However, when Cassim wanted to leave, he could not recall the simple, secret word. He tried in vain to remember, spouting word after word.
>
> The thieves returned to the cave to deposit more treasure. The leader spoke the simple, secret word. The rock rolled back to reveal Cassim cowering in the corner of the cave. Cassim never left.

The tale of Ali Baba and Cassim is more commonly known as *Ali Baba and the Forty Thieves* from *1001 Arabian Nights*. Some lessons emerge from this narrative:

- The first paragraph illustrates that passwords are not resistant to surveillance. Ali Baba, after all, was able to enter the cave because he was eavesdropping. The problem still persists today, as attackers can easily observe

individuals entering their passwords by shoulder-surfing. Passwords need to be resistant to eavesdropping and observation, otherwise their strength is weakened. Passwords need to be secret and kept safe.

- The second paragraph showcases secrets being shared with others. Ali Baba is able to communicate the password to his brother and, as a result, Cassim was also able to access the cave. Passwords should be known only to the individual accessing the cave, no-one else. Otherwise, it weakens the strength of the password as it is no longer secret.

- The third paragraph demonstrates that passwords are not memorable. Cassim's life literally depended on his remembering the password that Ali Baba shared with him. A good password is a lengthy string of gibberish with no meaning. The lack of meaning is what makes such passwords difficult to remember. If a user does not write it down, they will almost certainly reuse it. While a memorable password has meaning, it also makes it predictable as it has some structure or pattern. Such passwords are easily compromised by brute-force or social engineering attacks.

- The final paragraph exemplifies the fact that users and organisations only become aware of the attack when the attacker is unsuccessful. When the leader of the thieves rolls back the rock, he is confronted with Cassim. The secret has been compromised, the password is pointless. In modern day electronic systems there is often no way to detect that an attacker has successfully used another person's password.

These problems have always plagued passwords, even when they were first deployed on the Compatible Time Sharing System (CTSS) in 1962 [56]. The key difference between then and now is that CTSS was aimed at scientists and engineers. The password pitfalls could be combatted with training and education. Passwords are undeniably powerful, if used properly. Organisations can craft secure locations, make surveillance difficult, and use security policies and training to mitigate the aforementioned problems. The real problem is everyday users, managing an increasing number of passwords to achieve more and more tasks.

Users are increasingly turning to online services, relying on passwords, to complete tasks. Consequently, passwords are being pushed into the hands of users and being used in environments never envisioned. The initial envisioned password user was a scientist in a laboratory, now the list has expanded to include children in playgrounds and pensioners in living-rooms. In effect, passwords have become the mythical silver bullet for authentication; the default solution for authentication for every context and for every user.

However, the reality is that passwords are incredibly problematic, even in the hands of experts and professionals. Therefore, given the many problems associated with passwords, an alternative authentication approach is needs to be considered. It should be tailored to the needs of the average user of popular consumer products such as the Nintendo Wii and Apple iPhone. Such an authentication approach should tackle the problems of observation, memorability and sharing of

secrets. Such an authentication approach should be targeted at consumers and popular products rather than professionals and tele-type terminals.

A possible solution is graphical authentication. This approach relies on images rather than alphanumeric characters. There are several advantages to using images, a primary one being that images are more memorable than the words that name them. Unfortunately, despite considerable research into graphical authentication mechanisms, as well as other approaches, a viable alternative is yet to emerge.

There could be several reasons for a lack of a viable alternative. Dunphy et al. argues the reason for the slow uptake of graphical authentication approaches is due to designers and researchers not tackling practical issues such as deployment [76]. Nevertheless, there are many caveats or negative elements that detract from the overall story, e.g. sourcing images and creating graphical authentication secrets. The reality is that graphical authentication mechanisms, like many other alternatives, are rarely evaluated beyond the bounds of a laboratory or other controlled environments. Consequently, many aspects of authentication that are equally important, in ensuring that an authentication process is viable, are often neglected. The sourcing of images, as well as the creation of the authentication secret, are often neglected areas in the evaluation of graphical authentication mechanisms.

Furthermore, authentication context is another aspect that is often not discussed, i.e. the envisioned task and target user of the alternative authentication approach. Moreover, many alternative authentication mechanisms are often constructed as pseudo prototypes or not evaluated on actual target hardware, e.g. paper and pen is substituted for an actual software-based mechanism or an authentication mechanism designed for mobile phones is assessed on a desktop computer.

Consequently, many evaluations for an alternative authentication mechanism have relatively poor ecological validity. Nevertheless, controlled evaluations are invaluable in assessing aspects of authentication, e.g. cognitive load. The point is that controlled evaluations are not necessarily best suited for determining the overall performance or suitability of an authentication mechanism for a given context. Therefore, arguably much of the data collected in controlled evaluations is not relevant in determining the viability of an authentication mechanism *itself*.

The lack of a viable alternative to passwords is concerning as the approach is becoming increasingly stretched. The password approach was simply not designed for popular consumer based products that do not rely on physical keyboards and can used in environments with many wandering eyes.

Consequently, an alternative authentication approach is required that is designed for the aforementioned context. Here it will be argued that the viability of the alternative authentication mechanism can only be assessed through field evaluations to ensure it is fit-for-purpose.

1.1 Approach

A graphical authentication mechanism was developed that was designed to address the password problems outlined earlier, namely: *memorability*, *observation* and *sharing*. The problems were addressed (1) using memorable images, (2) entry of images in full view of onlookers without leaking the authentication secret and (3) knowledge of the images can not be easily shared with others. Nevertheless, the focus of the design is observation-resilience.

A prototype of the authentication approach was designed, implemented and assessed using a controlled evaluation to ensure the authentication mechanism was acceptable for use and ready for use in the field. Subsequently, variations of the graphical authentication mechanisms were evaluated using field investigations to determine the viability of the alternative authentication approach.

1.2 Thesis Statement

Therefore, the thesis statement is as follows:

> The viability of a recognition-based graphical authentication mechanism can only be evaluated in the wild.
> **Thesis Statement**

The steps and research questions needed to confirm or reject the above thesis statement are outlined in the next section.

1.3 Thesis Structure

The thesis structure is outlined below:

- The direction of the graphical alternative authentication approach is outlined in the next chapter. The **Literature Survey** outlines the various forms the alternative authentication approach could take before settling on a knowledge-based approach. Subsequently, the chapter states that recognition is the most effective way to extract knowledge from the user's memory. The chapter also outlines the various graphical authentication approaches proposed and discussed in research before concluding the alternative approach will rely on images. Consequently, the chapter outlines states faces are the best image-type for use in the authentication. Lastly, the chapter ends with a review of observation-resilient approaches before stating a software-based strategy will be used for the alternative approach.

- The design, implementation and evaluation of the approach is outlined in the **Alternative Authentication Mechanism** chapter. The conclusion is that the alternative authentication mechanism showed promised in controlled evaluation and that the mechanism can move to inclusion in a task and deployment to the field.

- The potential authentication context is outlined and discussed in the **Evaluation Task** chapter. Numerous contexts and authentication tasks are discussed and evaluated for risk before a task is selected for the field evaluations of the authentication mechanism. Furthermore, several research questions are outlined for exploration in subsequent evaluations.

- The alternative authentication mechanism is evaluated as part of an application in three field investigations. The three applications are codenamed: Tom, Dick and Harry. The design, implementation and evaluation of the first application are outlined and discussed in the **Tom** chapter. The conclusion of the evaluation is that the selected image set is unsuitable and an alternative must be sourced.

- The design, implementation and evaluation of the first application are outlined and discussed in the **Dick** chapter. The application relied on a personal image collection sourced from a popular social network. The images were analysed and face images extracted. Furthermore, users were expected to select distractor and target images. Consequently, the registration process took too long to complete. Moreover, there was concern that attackers may be able to identify friends selected. Subsequently, the next iteration of the application had to source an alternative image set.

- The design, implementation and evaluation of the first application are outlined and discussed in the **Harry** chapter. The application continued to rely on a personal image collection sourced from a popular social network. However, face images were generated using tags rather than analysis. Moreover, more images were sourced for each target and distractor to cycle through images to ensure resilience against attackers and improve retention of the authentication secret. However, the approach was not well received by users.

- The penultimate chapter is the **Discussion**. The chapter essentially argues that field evaluations were the most apt tool for assessing the viability of the alternative authentication mechanism. However, the alternative authentication mechanism was simply not viable for the envisioned authentication context.

- The final chapter is the **Conclusion**. The final chapter essentially offers an overview of the entire thesis before concluding that field investigations should be used to determine the viability of an authentication mechanism for a given context. The outcome of this evaluation, the alternative authentication mechanism assessed was not viable. Alternative directions for future study are outlined in a discussion on future work.

Chapter 2

Literature Survey

Passwords are among the most powerful and prominent authentication solutions in use. Nevertheless, the problems with passwords are long established, since early use, passwords have proved difficult to use and recall [55]. Furthermore, poor passwords not only have the potential to inflict their users but countless others as well [80]. Moreover, password replacement and recovery cost is estimated at $17 per incident and generates an estimated 30% of support calls [157].

Nevertheless, while the need to advance beyond passwords is clear, the direction is not [38]. The assumption is often that as tokens are easily misplaced and knowledge easily forgotten, the only direction for authentication is the dependable and indisputable biometric [199]. However, password popularity persists, most likely due the numerous advantages and benefits the concept posses. Therefore, an *alternative* knowledge-based solution may allow authentication to progress beyond the password in some key contexts.

Consequently, the following chapter justifies the foundation of an alternative knowledge-based solution that relies on face images. The ensuing section, §2.1, offers an overview of authentication before outlining a taxonomy of various approaches, §2.2. The conclusion is that knowledge-based solutions are still advantageous for mainstream use. Consequently, various knowledge-based solutions are considered, §2.3, before determining a graphical recognition-based authentication approach is optimal, §2.4. Therefore, various images types and obfuscation solutions are considered for use in such an approach — §2.5 and §2.6, respectively. Lastly, §2.7 summarises the foundation of the alternative authentication mechanism.

2.1 Authentication Overview

There are several potential authentication definitions and uses. Section 2.1.1 defines authentication in terms of users before outlining potential uses in §2.1.2. Lastly, §2.1.3 outlines a minimal set of goals for user-facing authentication.

2.1.1 Definition

Needham and Schroeder provide one of the earliest definitions of authentication in digital networked systems, as follows:

> "verifying the identity of the communicating principles to another"
> **Needham and Schroeder [187, p. 993]**

The above definition, published in 1978, offers a clear, concise explanation of authentication. However, while succinct and applicable to authentication as whole, it less informative of the process when individuals are involved. Shockley outlines an definition, 15 years later, that incorporate users:

> "In general, we can think of the authentication data as consisting of two parts: an authenticator which is held by the actual user and an authenticand which is stored as part of the account data record. An authentication algorithm is built into the trusted computing base that, given an (authenticator, authenticand) pair determines, with an appropriately high probability of success, whether or not the given authenticator and authenticand match."
> **Shockley [251, p. 185]**

Shockley's definition is far more comprehensive and incorporates the typical steps or implementation of an authentication process. Moreover, Shockley highlights the involvement of users in the process. This aspect is admittedly absent from the succinct definition offered by Needham and Schroeder. Rao and Yalamanchili evolve the definition further, 19 years later, explaining authentication in far plainer terms, as follows:

> "Authentication determines whether a user should be allowed access to a particular system or resource"
> **Rao and Yalamanchili [215, p. 163]**

Rao and Yalamanchili, in the above definition, essentially frame authentication as a process designed to serve a user. Nevertheless, the aforementioned definitions, spanning 34 years, share a common thread: that authentication is a process of confirming a claimed identity.

The aforementioned definitions all have similar foundations. The process of authentication has remained consistent between 1978 and 2013. The definition offered by Needham and Schroeder is as relevant and accurate as the one outlined by Rao and Yalamanchili.

However, there has been an evolution in terms of the *principles* or *parties* involved in authentication. The definition offered by Needham and Schroeder delineates authentication in a purely mechanical fashion, involving seemingly reliable and infallible actors. The definitions offered by Shockley, as well as Rao and Yalamanchili, focus on users, actors that are unpredictable and unreliable.

Therefore, the user must be factored into any authentication process and must be given due consideration in the *use* and *goals* of an authentication process.

2.1.2 Use

There could be many reasons for the use of authentication in a system. Berson et al. outlines two key uses for authentication [26], as follows:

- *Access control*
  The need to regulate access to a restricted object. The regulation is granular; it could be system-wide or of very fine-grain nature. An authentication mechanism could be used to regulate access to entire operating system, e.g. Microsoft Windows, access to a specific file, e.g. password protected Microsoft Word document, or access to a system service, e.g. Google email.

- *Accountability*
  The need to attach responsibility to a user action. An authentication mechanism could be used required to authorise the removal of a document or to uninstall an application. An authentication mechanism could also be used to confirm a transaction, e.g. purchasing a song from a digital music store.

However, not all authentication solutions necessarily support such uses, e.g. group-based schemes that rely on shared authentication secrets can be used to regulate access but do not ensure accountability. Nonetheless, the aforementioned uses are likely the two most common uses of authentication and solutions that do not support both should be avoided.

Furthermore, an authentication solution can only support such uses, if the authentication mechanism itself is usable and reliable. Consequently, a minimal set of specific goals must be set to ensure the success of an authentication mechanism.

2.1.3 Goals

Kurzban outlines a minimal set of goals that all authentication mechanisms should achieve [164], as follows:

- *Avoid false positives*
  An authentication mechanism must strive to ensure that attackers or undesirable individuals are not authenticated. If the primary uses of authentication are *control over access* and *accountability for actions*, then there must be confidence in an authentication mechanism. An authentication mechanism that permits too many false positives is not functionally fit for purpose.

- *Avoid false negatives*
  An authentication mechanism must minimise the scenario where an actual user is not authenticated. If legitimate users are unable to access services or complete transactions it does not only have a serious impact on an individual but on an organisation as well, potentially impacting profits.

- *Minimal burden*
  The authentication mechanism must impose minimal burden on the user. The authentication mechanism can not be overly demanding.

- *Cost-Benefit Balance*
  The authentication mechanism should represent a cost-benefit balance. If the purpose of the authentication process is to *regulate access* or *confirm payment* then the expected user-effort should, proportionally, reflect the risk.

The user interface of an authentication mechanism is important in achieving these goals. The password approach will not necessarily be superseded by an authentication approach with a similar user interface. An authentication mechanism can take many forms and some interfaces may prove more suitable for modern applications and devices. Therefore, the user interface of the authentication mechanism must be considered.

2.2 Taxonomy of Authentication Approaches

The National Institute of Standards and Technology categorises authentication approaches [180], as follows:

- **Knowledge-based**
  The authentication secret takes the form of knowledge, known to the claimed identity.

- **Token-based**
  The authentication secret takes the form of a physical token, possessed by the claimed identity.

- **Biometric-based**
  The authentication secret takes the form of a physiological and/or behavioural trait, belonging to the claimed identity.

Wood labels the categories in much plainer terms, as something the user *knows*, *has* or *is* [295]. The NIST categorisation is a typical taxonomy adopted by many authentication researchers [90, 134, 212]. Nevertheless, there are authentication mechanisms that rely on cognitive processing [248] and social relationships [41] that are, arguably, unclassifiable using the above labels. Therefore, an additional category is:

- **Emerging Approaches**
  The authentication secret is not knowledge, token or biometric-based, instead the approach relies on another form of secret. The authentication secret itself, depends on the approach, e.g. a mechanism based on social relationships relies on a user's relationship with other individuals.

The aforementioned classifications can be used to categorise most authentication solutions involving individuals. Nevertheless, authentication approaches from each category or several from one category can be combined to provide additional security [194]. Therefore, an another category is:

- **Multi-factor Approach**
  There is more than one authentication secret or authentication approach involved in the authentication process, e.g. a token in the form of bank card coupled with knowledge of a personal identification number (PIN).

The aforementioned category is primarily used to improve the strength of authentication within a task. However, there are other authentication solutions that are not necessarily designed for inclusion in primary tasks but are used to recovery from authentication failure. The user may be unable to complete the primary authentication process as they have forgotten the authentication secret. Therefore, the last remaining classification in the taxonomy is authentication approaches designed to recover from failure:

- **Recovery Approaches**
  These are back-up authentication approaches, used when an individual needs to reset or recover a secret used in the aforementioned approaches.

The aforementioned categories comprise an authentication taxonomy that acts as umbrella for several authentication solutions that involve individuals. The categories are outlined and discussed in §2.2.1 to §2.2.6. Lastly, §2.2.7 outlines the optimal category for the alternative authentication solution.

2.2.1 Knowledge-based Approach

Knowledge-based authentication is the process of confirming a claimed identity through knowledge of a secret, one known only to you and the other party. The knowledge needs to remain a secret if is to be used for authentication. Therefore, individuals are advised to memorise such knowledge and not to record or share it. The knowledge itself could be a public event or record, but the use thereof must be kept a secret, e.g. date of an anniversary.

In actuality this is a terrible secret to use since a date is information, has meaning and, worse still, is connected to the individual. This makes the title of this category rather misleading, since the word *knowledge* implies information, an object, which an individual has extracted or developed personally through experiences or learning. Knowledge-based secrets are ideally prodigious, impersonal and meaningless.

Alphanumeric authentication is the best known implementation of knowledge-based authentication. There are two reasons for this: (1) the concept of passwords is one which is centuries old and is easily understood by many (2) the interaction mechanism, i.e. keyboard, is over a century old and one can easily enter passwords without additional training or expense. This made passwords the authentication mechanism of choice for early systems, such as the Compatible Time Sharing System (CTSS) [56], and operating system designers such as Ken Thompson and Dennis Ritchie.

Once again, *password* is a misnomer as, ideally, it should be a lengthy indecipherable string of characters and symbols and certainly not one with meaning or found in a dictionary. The reality is that many individuals use words which are familiar to them as passwords [231].

Morris & Thompson discuss the problems of alphanumeric authentication as early as 1979, specifically stating effort had to be spent in ensuring that individuals created passwords which were 'less predictable' [183]. Indeed, Klein was able to uncover 25% of the passwords used, in 15,000 accounts [155].

Klein used various strategies to extract passwords, the first one was to use personal information. The project leader behind CTSS, Fernado J. Corbató, stated that although passwords seemed theoretically practical, in practice, they presented many problems. Corbató argues passwords are difficult to remember, easily compromised and difficult to revoke when shared [55].

2.2.2 Token-based Approach

Token-based authentication relies on an individual presenting an object or token. The object itself acts as an authenticator and it does not need to be kept secret, just safe. Consequently, an individual is expected to carry the token at all times to avoid inconvenience.

Spender argues driving licenses are likely the most common form of token-based authenticator [256]. However, modern token-based authentication relies on increasingly sophisticated tokens or objects. These modern tokens are far more expensive than knowledge-based authentication but comparatively cheap when contrasted with biometric-based authentication. Table 2.1 outlines the principal costs of modern authentication tokens. Bojinov and Boneh state modern authentication tokens are comparatively expensive and rely on a central system or server [37].

Device	**Price**		**Power**		**Usability**
	Token	*Reader*	*Token*	*Reader*	
RSA SecurID	$50	>$10,000	Low	Low	Poor
Vasco Digipass Go	$10	$500	Low	Low	Poor
Car RKE fob	$5	$5	Low	Low	Average
HID Proximity	$2	$100	None	Average	Good
RFID (or NFC)	<$1	$50	None	Average	Good
Smartcard	$2	$10	None	Low	Poor
Magnetic stripe	<$1	$50	None	Low	Poor
QR (via camera)	<$1	$10	None	Low	Poor
Bluetooth	$10	$5	Average	Low	Average

Table 2.1: Hardware tokens and principal concerns [37, p. 15]

The tokens or devices outlined in Table 2.1 are not traditionally targeted at consumers but rather deployed by enterprises. However, variants of RSA SecurID, a popular token-based authentication approach with enterprises [234] are being deployed to consumers, e.g. online bank customers.

However, while such tokens are tamper-proof, they are still susceptible to theft. An attacker could continue to use the token until it is reported and deactivated. Moreover, while the attacker-effort necessary to steal a token may

be more laborious than compromising knowledge-based authentication, the approach does not necessary reduce attacks, merely transfers an attacker's focus and energy elsewhere [232].

Therefore, while token-based authentication may be considered by enterprise, as a more practical approach to achieve security objectives, the approach is not without flaws. Nevertheless, token-based authentication is particularly suitable for large organisations. Tokens can be produced relatively inexpensively if created on a large scale. The tokens themselves can offer dynamic entry and fine-grained access to resources, unlike traditional door-keys. Furthermore, sophisticated tokens can support the modern-day bureaucracy associated with large-scale organisations.

2.2.3 Biometric-based Approach

Prabhakar et al. defines a biometric-based authentication approach, as follows:

> "A biometric system is essentially a pattern-recognition system that recognizes a person based on a feature vector derived from a specific physiological or behavioral characteristic that the person possesses."
> **Prabhakar et al. [213, p. 33]**

Therefore, a biometric-based approach relies on authentication secrets generated from one of two characteristics [172]. The characteristics are, as follows:

- *Physiological*
  The physical characteristics of an individual, such as an individual's fingerprint, palm-print or retina.

- *Behavioural*
  The behavioural traits of an individual, such as an individual's gait, voice, signature or mouse movements.

These characteristics represent the primary advantage of biometric-based authentication, at least from a user perspective, as they are traits relatively distinct to an individual. The user is not required to carry an authentication token or memorise a meaningless string. The authentication secret is an aspect of the user, that is always with them, thus, an individual will never forget it or misplace it. There are many aspects of a user, that could be used as a biometric authentication secret. Jain et al. states any aspect of an individual could be used as a biometric authentication secret, as long as it fulfils certain requirements [127]. The requirements are, as follows:

- *Universality*
  The feature is one that is possessed by everyone.

- *Distinctiveness*
  The feature is unique, in that it is discernible from another individual's.

- *Performance*
  The feature itself should be relatively constant over a significant amount of time, i.e. not rapidly degrading or changing.

- *Collectibility*
  The feature itself can be quantified.

Biometric-based authenticators are arguably difficult and intrusive to collect when contrasted to, say, the user-generated authentication secrets used in knowledge-based approaches. However, biometric-based authentication secrets do not require the user to carry tokens or remember passwords; indeed, little is required of the user. Nevertheless, physiological biometric-based authentication secrets require an extraction step, e.g. the user's fingerprint is taken. The hardware involved in the extraction step may have a higher failure rate, such as several failed attempts in every 100, than a traditional keyboard or token reader [194]. The equipment may also suffer from any number of environmental effects, e.g. condensation, lighting or dust, all may result in errors in biometric-based authentication. Prabhakar et al. states biometric-based authentication suffers from two types of recognition error [213], as follows:

- *False match rate (FMR)*
  The biometric-based system when presented two patterns from *different individuals*, incorrectly determines the patterns belong to *a single individual*.

- *False non-match rate (FNMR)*
  The biometric-based system when presented two patterns from *a single individual*, incorrectly determines the patterns belong to *different individuals*.

Prabhakar et al. states a biometric-based system makes a balance between the aforementioned rates. The data being extracted is incredibly complex and the speed at which a sophisticated system is able to extract salient components and match these against a large database of such features could be time-consuming and pointless, especially if it generates a false match. Jain et al. argues the speed and success of results depends on the function of the biometric system [128].

Furthermore, asides from the speed and accuracy another concern is the authentication secret itself. The features used in biometric-based authentication cannot be replaced. If an individual's fingerprint has been extracted and compromised as an authentication secret, it is no longer suitable for use as an authenticator. However, a biometric secret does not need to be collected and stored in such a way that it can be repurposed. An example would be the use of finger geometry rather than a fingerprint, storing measurements rather than the print itself. This process is incredibly common and used throughout the world, including controlling access to attractions in Walt Disney World, Florida [285]. However, such questionable low-level operations is the limitation of such simplistic biometric-based authentication secrets, since geometry is not a unique, distinct or reliable secret. Thus, it holds little to no repurposing benefit.

The popularity of the geometric approach, among corporations, is partly due to the simple and inexpensive implementation cost associated with using such biometric-based authenticators. Furthermore, as advances are made and costs are reduced, biometric-based authentication is likely to increase in popularity, as it requires little from individuals and offers vast amounts of information to organisations.

2.2.4 Recovery Approaches

An authentication approach can fail for many reasons. However, there is always the possibility that authentication fails simply because the user is unable to present the necessary information or object required to authenticate. These situations are usually addressed by offering an auxiliary recovery approach or procedure. The recovery approaches for the aforementioned authentication approaches, are, as follows:

- *Knowledge-based*
  The recovery process concludes with the replacement of the authentication secret. Just argues that good recovery approaches are designed to rely on information the user already knows rather than information they are required to memorise [143]. Therefore, challenge-questions are a prominent and widespread approach for authentication secret replacement. The approach essentially presents a user is a series of questions, if answered correctly, the user is permitted to replace the authentication secret.

  Unfortunately, challenge-questions may introduce more problems than they solve. Schechter et al. discovered that 20% of users could not recall challenge-question answers within six months and that 13% of challenge-question could be guessed within five attempts [240]. Furthermore, Just and Aspinall argue challenge-questions reliant on personal information are not necessarily painless to answer or difficult to guess [145]. The reality is that many challenge-questions can be answered by consulting public records [106] or social network services [214]. However, Just and Aspinall argue security can be improved by simply asking more questions [146]. Furthermore, the strength of the approach may be improved by using other channels, such as by post [144].

  Nevertheless, researchers have proposed alternatives to the dominant challenge-question approach. Renaud and Just proposes a recovery approach based on graphical cues [221]. The approach is essentially a challenge-based approach with an added graphical hint or cue. Schechter et al. adopts a different recovery strategy and outlines an approach based on retrieving recovery-codes from various trustees [241]. The process requires an individual to specify a number of trustees or insurers, in advance of failure. Unfortunately, neither solution is able to overcome the concern of close friends and relatives successfully completing the recovery process.

- *Token-based*
  The authentication token is replaced, e.g. replacement of driver's license. However, a recovery approach does not necessarily need to result, *only*, in the replacement of a token. The system and recovery process can be designed to generate a temporary token or authentication secret. Brainard et al. proposes an approach that relies on *somebody you know*, acting as a temporary token [41]. The concept is not dissimilar to the aforementioned trustee solution, the approach expects another individual vouching for a user, i.e. another individual confirms the identity of a user.

- *Biometric-based*
  Biometric-based authentication approaches are similar to token-based approach in that immediate password recovery or replacement is not viable. If an individual loses their limbs, hands or eyes, then another biometric authentication secret would need to be generated.

The interesting aspect of the aforementioned recovery approaches is that they are not specifically distinct from authentication approaches. A recovery process is essentially an authentication process. However, recovery processes are generally unsuitable as they are typically protracted and required effort to complete [144].

Nonetheless, many of the ideas and concepts used in recovery approaches are present in authentication approaches, evident from the nascent emerging approaches that rely on somebody you know.

2.2.5 Emerging Approaches

The primary categories for authentication approaches are knowledge-based, token-based and biometric-based. Nevertheless, there are authentications approaches that may be deserving of distinct branches [41, 248]. The proposed additional branches are:

- *Something you process*
  Shah et al. proposes algorithmic-based authentication or something you process [248]. These authentication approaches rely on knowledge of a secret formula to authenticate. The authentication secret is generated with the secret formula, using variables presented by the authentication approach.

- *Somebody you know*
  Brainard et al. proposes social-based authentication or somebody you know [41]. These authentication approaches rely on social relationships to authenticate. The authentication secret could be anything but is obtained from an ally or acquaintance.

However, while the authentication approaches, contained within these categories, are certainly novel, they are arguably not distinctive enough to warrant specific branches.

2.2.6 Multi-factor Approaches

O'Gorman defines a multi-factor authentication approach, as follows:

> "Different types of authenticators can be combined to enhance security. This is called multi-factor authentication. For security purposes, each authenticator result must be satisfied; in effect a Boolean AND operation is performed for each factor's authentication results so all must be affirmative."
> **O'Gorman [194, p. 7]**

Therefore, multi-factor authentication is essentially the use of two or more authentication approaches to authenticate an individual. Table 2.2.6 outlines a comparison of the various multi-factor combinations [194].

O'Gorman argues the combination of authentication approaches, e.g. ownership of a plastic card token and knowledge of a secret PIN, is superior to relying on a single authentication approach [194]. Furthermore, Bhargav-Spantzel et al. argue that multi-factor authentication not only offers superior strength but is necessary to make some authentication approaches, such as biometrics, practical [27]. Similarly, Jin et al. argues that many of the problems associated with biometrics,

Authenticator Combination	**Security Advantage**	**Convenience Drawback**	**Example**
Knowledge and Object-based	Lost/stolen token protected by password	Must carry token and memorise password	PIN-enabled bank card
Object- and ID-based	Lost/stolen token protected by ID	Must carry token, but not ID if it is a biometric	Photo-ID
Knowledge- and ID-based	Two factors provide security in case either compromised	Have to memorise password and have ID	Password and biometric for computer access.
Knowledge-, Object-, and ID-based	A third factor to provide security in case two other factors are compromised	Have to memorise password,carry token, and have ID	Military applications requiring photo-ID checked by guard, plus password

Table 2.2: Security and convenience assessment of multi-factor combinations [194, p. 8].

i.e. difficult to revoke and inconvenient to use, can be addressed by incorporating the approach into a multi-factor solution [133]. Moreover, Aloul et al. argues expecting users to manage multiple multi-factor authentication approaches has become increasingly realistic, as mobile phones for most consumers have been superseded by slimline and sophisticated smartphones [2]. Lastly, multi-factor authentication has been widely deployed to users.

The Chip and PIN approach serves as popular and successful example of multi-factor authentication in the United Kingdom [18, 19]. Nonetheless, rather than eradicated fraud, attackers are simply adopting more sophisticated tactics and strategies [20]. Furthermore, Murdoch et al. argues that the multi-factor authentication approach is far from successful and is fundamentally broken [184]. Nevertheless, while Chip and PIN may have problems, its popularity indicates the potential for multi-factor authentication with users.

However, others argue multi-factor approaches, such as Chip and PIN, do not solve many of the problems associated with the approach they replace [5, 243]. Schneier asserts users will still write authentication secrets down, they will reuse them and they will still share them [243] . Nevertheless, Henry argues such

remarks discount the fact that authentication is only part of an organisation's infrastructure [115]. Nonetheless, Henry arguably discounts the importance users place on convenience [287]. Consequently, users will behave insecurely and adopt novel strategies to improve convenience.

Nonetheless, multi-factor could arguably be used to improve the convenience of some authentication approaches, such as biometrics. However, Rathgeb and Uhl argue multi-factor approaches encompassing biometrics make assumptions that are not necessarily valid, such as presuming other authenticators are not compromised [217]. Rathgeb and Uhl state researchers and designers of some multi-factor solutions make illogical assumptions, e.g. a password or random number can not fall into the wrong hands. However, if that was the case, there would be no need for multi-factor authentication approaches in the first instance, the impervious password or one time random string would suffice [217]. Furthermore, Rathgeb and Uhl highlight an interesting aspect of multi-factor authentication, in that such solutions are essentially strong chains of weak links. The authentication approaches contained with a multi-factor solution are essentially flawed, otherwise they would not need to be combined with others in the first instance.

Therefore, a multi-factor authentication solution is akin to making lemonade from lemons. The primary benefit of such a solution is strength [194]. The cost is complexity, users are expected to manage several authenticators, leading to a potentially less convenient solution, e.g. lengthier authentication times. Moreover, increasingly complexity only serves to undermine security, as developers and implementations struggle to manage such solutions [296]. However, such complexity costs can be avoided if the multi-factor solution is avoided in favour of a single strong authentication approach.

Unfortunately, there is no single strong authentication approach. Authentication mechanisms are invariably flawed, suboptimal or undesirable in some respects. Therefore, authentication approaches must be carefully considered to determine if such flaws represent a serious threat within a particular context before resorting to a multi-factor solution.

2.2.7 Comparison of Approaches

The reality is that when users are confronted with the choice between convenience and security, they favour convenience [287]. Consequently, users have adopted tactics when using passwords that ensure convenience but nonetheless erode security. Therefore, the aim is to progress beyond the password as the *de facto* authentication approach by developing an alternative solution that is both convenient and secure.

Consequently, the previous sections outlined the interface direction considered for the foundation of the alternative authentication mechanism. Three of these interface branches, namely: *emerging*, *recovery* and *multi-factor*, were quickly discounted for several reasons. The 'emerging' branch contained approaches that were either loosely defined or were not particularly compelling or distinct. The 'multi-factor' branch is focused on improving security rather than convenience. Lastly, the 'recovery' branch while focused on improving convenience *overall*,

arguably the branch covers authentication solutions unsuitable for primary authentication due to complexity.

Consequently, research was quickly curtailed to the traditional NIST categories [180], as any could from the foundation for the alternative authentication approach. Table 2.3 outlines the distinct aspects of all three categories [194]. The competing authentication approaches, i.e. knowledge, token and biomet-

		User Authentication		
		Knowledge-based	**Object-based**	**ID-based**
Commonly Referred to as:		Password, Secret	Token	Biometric
Support Authentication by:		Security or obscurity	Possession	Uniqueness and personalisation
Security Defense:		Closely kept	Closely held	Forge-resistant
Example	**Traditional:**	Combination lock	Metal key	Driver's license
	Digital:	Computer password	Key-less car entry	Fingerprint
Security Drawback:		Less secret with each use	Insecure if lost	Difficult to replace

Table 2.3: O'Gorman categorisation of authentication approaches [194, p. 7]

rics, all have associated costs as well as different strengths and weaknesses. Token and biometric-based authentication approaches are essentially stronger than knowledge-based approaches but are far less convenient [194]. O'Gorman argues both token and biometric-based authentication approaches should be delivered as part of a multi-factor solution [194].

The motivation is that a token without an additional factor is easily compromised if stolen. In case of biometrics, coupling the approach with a token containing identifying information can make the authentication process more efficient. However, as part of networked systems and services both tokens and biometrics are far more problematic [194]. The recover and replacement approaches are far more complex than knowledge-based approaches, as users need to visit specialised facilities or wait on delivery of sophisticated tokens. The approaches are simply far less convenient and cost considerably more due to the complexity of additional hardware.

Furthermore, most successful consumer applications are coming from small software houses and sole-traders [10]. Apple recently announced that 40 Billion such applications had been downloaded by users, 20 Billion being delivered in the past year alone [235]. Consequently, sole-traders and small software houses will heavily scrutinise administration costs of authentication solutions. Nevertheless, smartphone manufacturers could simply provide hardware and software solutions [2] and subsume the administrator costs. However, authentication is part of many systems and services, not just smartphones, consequently the device would need to communicate with a wide-range of others, introducing further complexity and potentially weakening the overall solution.

Furthermore, another concern is that a small group of large and powerful smartphone manufacturers would have amass vast amounts of information about individuals. The smartphone manufacturer would have an overview of all the systems and services an individual uses. Moreover, in terms of biometrics, an organisation would have access to a wealth of data that may have applications beyond authentication. The data an individual provides for authentication could be applied in health or even advertising ventures.

The user may simply not be aware of the potential applications for the data they are providing for authentication. Furthermore, another aspect of token and biometric authentication is that the transaction can be performed passively. An arguable advantage of knowledge-based authentication is that the user has to be engaged to authenticate. The authentication approach has to interact with the user and keep them in the loop. A potential concern for biometric and token-based authentication approaches is that users can be removed from the authentication loop.

Therefore, there are several concerns surrounding biometric and token-based authentication. Furthermore, these concerns only serve to highlight the advantages of knowledge-based solutions. The reality is that knowledge-based authentication secrets are are vastly superior in terms of convenience to token and biometric approaches [194]. A knowledge-based authentication secret is (1) easily replaced, (2) has little to no applicability beyond the process itself and (3) users need to be engaged for it to be extracted. Therefore, given these advantages knowledge-based authentication was selected as the foundation for the alternative authentication solution.

Therefore, the aim was to develop a knowledge-based alternative to the password. Nonetheless, the key concern of knowledge-based authentication is that users do not memorise the authentication secret. In term passwords, users create simple strings that are memorable but are not strong enough to withstand assaults from attackers. The concern can be characterised as the 'password problem', i.e. users are expected to generate lengthy meaningless strings and memorise them [292]. The problem is that users find it very difficult to memorise meaningless information, never mind memorising a meaningless string for every system and service they use. Wiedenbeck et al. states:

> "A better way to overcome the password problem is to develop password systems that reduce fundamental memory problems"
> **Wiedenbeck et al. [292, p. 105]**

Consequently, memorability became the focus of the design of the alternative knowledge-based authentication approach. The aim was to improve the memorability of the authentication secret used within the alternative knowledge-based approach to negate the need for coping strategies.

2.3 Taxonomy of Knowledge-based Authentication Approaches

The reality is that many users adopt undesirable coping strategies to bear the burden of memorising knowledge-based authentication secrets. Users resort to creating simple authentication secrets, reusing them or recording them. These actions are unsurprising, as users are motivated out of concern, namely that they will be unable to complete tasks as they will not be able to complete authentication.

Therefore, in designing and developing knowledge-based authentication approaches, researchers have focused on improving memorability of authentication secrets. Unfortunately, the human brain is still very much a black-box in some respects - making it perfect for storing secrets but difficult for scientist to understand and investigate. Nevertheless, many researchers have contributed greatly to the study of memory.

Scoville and Milner investigated patients who had received bilateral brain surgery and reported that a memory defect was always present when the hippocampus and hippocampal gyrus had been bilaterally corrupted [246]. Thus, these regions of the brain are deemed central to memory.

Therefore, forgetting could be the result of natural decay within this region. However, Underwood argued forgetting was not necessarily natural decay but caused by previous memories [277]. Craik and Lockhart suggested development of memory traces should be considered in terms of *depth of processing*, i.e. extent of sensory and semantic interpretation [59]. Memory traces deemed *weak* are those which progress little beyond sensory interpretation, thus are susceptible to decay. Furthermore, Craik and Lockhart suggest information that sits well with existing knowledge will be rapidly processed to greater depths. However, Wixted argues that there are limited resources in memory creation and development, when those resources are taxed it leads to the detriment of previous memories [294].

Therefore, while there is still much discussion surrounding the creation and destruction of memories, there does appear agreement on the key stages of memory. Memory can be decomposed into three key stages or processes, as follows:

1. *Encoding*
   Information is extracted from various senses and encoded into a format suitable for storage.

2. *Storage*
   An encoding is stored in memory, for retrieval.

3. *Retrieval*
   An encoding is extracted from storage in response to some sensory information or activity.

The aforementioned steps are important in the design of alternative solutions. The aim of alternative authentication solutions to the password have often focused on one, if not all three stages to improve the memorability of knowledge-based

authentication secrets. The proposed solution may aim at improving the initial encoding or may focus on improving the interface to retrieve the authentication secret when required.

The authentication solutions to overcome the shortcomings and concerns of passwords can be roughly classified using one the following categories:

- *Alphanumeric authentication*
  The proposed solutions are iterations or improvements on the original password, still relaying on alphanumeric characters.

- *Alternatives to alphanumeric authentication*
  The proposed solution is arguably still an iteration on the original password concept but relying on sound and/or graphics rather than alphanumeric characters.

The knowledge-based authentication solutions proposed by researchers to overcome the problems of passwords are outlined discussed, in terms of the aforementioned classification.

2.3.1 Alphanumeric Authentication

The many advantages of passwords has lead to much interest in addressing the primary disadvantage, namely memorability. Therefore, rather than disregarding passwords completely, researchers have instead attempted to iterate on the approach to improve the strength and memorability of authentication secrets. Zviran & Haga outline various alternative approaches that could be used to generate stronger passwords [305]. The approaches are, as follows:

- *System-generated passwords*
  Allowing a system to create a secret, rather than a user, ensures the secret is closer to gibberish than predictable information. Random strings of gibberish are strong against brute-force attacks, but are difficult for users to recall [298].

- *Passphrases*
  Porter proposes the concept of a passphrase [212]. The user creates a sequence of words, that together, act as an authentication secret. Porter argues the phrase is more memorable to the user but harder for an attacker to guess, as it contains many more characters than a simple traditional password.

- *Associative passwords*
  Smith proposes the use of word association to authenticate users [253]. The user creates a collection of *cues* and *responses*. A pair is randomly selected and a user must provided the respective response for the cue presented.

- *Cognitive passwords*
  Zviran refines the process of associative passwords by probing a user's personal experience [304]. This quiz-based approach extracts several pieces of

knowledge from the user. The individual is presented a series of *fact-based* and *opinion-based* questions. A fact-based question would be 'What was the first school you attended?', while an opinion-based question would be 'What is your favourite film?'.

However, while the proposed approaches aim primarily to strengthen the memorability of alphanumeric authentication secrets, others have focused on increasing the strength of them, namely:

- *Persuasive Text Passwords*
  Forget et al. propose persuasive text passwords (PTPs), an approach that attempts to couple the strength of user and system-generated passwords [86]. The approach expects users to initially create a password, then the system randomly inserts characters into the secret authentication string. The user then using the result string to authenticate.

- *Password Algorithms*
  Crabb adopted the use of password algorithms at the NASA-Ames Research Center [58]. Passwords sheets were widely distributed to employees at the research centre. However, out of concerns that individuals other than employees could observe the passwords an algorithm was separately communicated to employees. Consequently, the freely distributed passwords acted more like seeds for a secret algorithm that individuals were required to internalise.

- *One-time Passwords*
  Lamport originally proposed the idea of one-time passwords (OTPs) to tackle the problems of insecure communication and password breaches [165]. The concept is simple, an individuals uses a different password every time they authenticate with a system. Therefore, a used password is worthless as it can only be used once.

The aforementioned iterations and improvements represent an attempt to address the problems of passwords with novel authentication solutions. However, an alternative is to support the user in using the password approach, *properly*. The password is incredibly powerful, versatile as well as inexpensive to implement and explain. Unfortunately, the concept is riddled with problems as users continue to wield it improperly. Therefore, the following approaches could be used to support users in creating passwords:

- *Policy and Education*
  A combination of strict policies and education could be used to improve passwords [306]. Users could be provided with training and tools to create stronger, more superior passwords.

- *Proactive Password Compliance*
  Proactive password checkers could be used to assess the suitability of user-generated passwords [34]. The user would still generate a password but the system would provide feedback informing the user of the potential strength

of the authentication secret. Furthermore, the system could enforce compliance, rejecting passwords that have perceived, insufficient strength.

- *Password Managers*
  A software solution that simply adopts the burden of users memorising passwords. The software would simply manage the passwords issued and created by users. The user would simply be expected to memorise a single, strong password to access the manager. Alternatively, the password manager could rely on another form of authentication to access it, such as token or biometric-based authentication.

The aforementioned approaches for improving passwords as well as strategies for improving user choice are discussed over the ensuing sections.

Passphrases

Shay *et al.* explored the performance of system-assigned passphrases and discovered that they perform similar to system-assigned passwords [250]. Passphrases were not easier to recall than passwords and users still recorded them. Furthermore, Shay *et al.* states passphrases take a longer time to enter and suffer from 'typos'. Therefore, Shay *et al.* propose use of auto-correction to avoid typing mistakes in passphrase entry. However, Shay *et al.* caution about the relevancy of their results as the authentication mechanism was not explored as part of a system or workflow. They argue the performance of passphrases in the real-world could be dramatically different.

The focus of passphrases is to promote the use of lengthier, alphanumeric strings. If an individual mistypes a character, the entire phrase is rejected. Therefore, the concept of passphrases seems little more than marketing. There is no technical difference between a long password and a passphrase. However, Spector & Ginzberg propose Pass-sentence [255], an authentication approach based on semantics and natural language analysis to authenticate an individual. The approach allows for the interchange of words, as long as those words have the same *semantic primitive*, e.g. 'buy', 'sell' and 'trade' as all distill down to the concept of ownership transfer. Spector & Ginzberg state user could neglect words from the sentence, that system could subsequently probe or base access controls on. Spector & Ginzberg claim superior memorability with a developed prototype, although these claims have yet to be assessed in real-world use.

Associative & Cognitive Passwords

Similarly, associative and cognitive passwords have displayed, strong levels of recall, with both approaches being resilient to guessing, even from allies in the laboratory [109]. However, Bunnell *et al.* argues cognitive passwords are superior to word association [44]. Bunnell *et al.* evaluated the performance of both on 86 psychology students and their significant others by issuing questionnaires to each. Bunnell *et al.* reveals fact-based cognitive elements have superior recall to opinion-based elements but fact-based elements are easily compromised by close associates. Bunnell *et al.* states specific opinion-based elements had strong

recall and are not easily guessed, arguing these strong elements be extracted and assessed to improve the overall approach. However, Bunnell *et al.* emphasise that performance in the field may be dramatically different from their findings. Furthermore, they caution about the relevancy of results that do not closely model a system in the real-world.

Furthermore, while associative and cognitive passwords are promising and interesting they are still fairly complex and elaborate for regular use. Consequently, small iterations on the original password approach may be more desirable as many strengths of the approach can be maintained while addressing weaknesses.

Persuasive Text Passwords

Forget et al. proposes Persuasive Text Passwords (PTPs) as an iterative password approach that maintains much of the strength of the original approach but tackles the weaknesses of it [86]. The primary problem being that desirable secrets are essentially system-generated gibberish and users tend not to recall or create such passwords. Forget et al. argues users ignore security policies and instead should be influenced or persuaded to create better passwords. Forget et al. resolve this issue by mixing user and system-generated approaches in their proposed PTP approach. The proposed approach expects users to initially generate a password. The system then injects random characters throughout the password. The user can then 'shuffle' characters until they are happy with the resulting password. Forget et al. evaluated the approach and found that when the system contributed more characters, users reacted by creating simpler seed passwords. Therefore, they argue PTP should be limited to insertion of 3 characters, anymore and users resort to coping mechanisms. However, Forget et al. emphasises that these results are not indicative of real-world performance and PTP would need to be deployed as a field-study to determine actual practicality.

Password Algorithms

Password algorithms are one of the few examples of an alternative alphanumeric authentication approaches that have been deployed to actual users. Crabb adopted the use of a password algorithm at the NASA-Ames Research Center [58]. Crabb states that support workers had to be issued "password sheets" regularly, when passwords change for various systems. These changes could result in 40 different passwords. These password changes were printed onto sheets and distributed among staff. However, in-order to tackle the problem of an attacker obtaining the sheets, a password algorithm was issued alongside a password sheet. The password algorithm would require individuals to alter the printed password to obtain the actual password. Crabb offers an example of such an algorithm, "capitalise the first vowel and add a dash". Therefore, if the printer password was "freem3", the actual password would be "frEem3-", Crabb states. The approach adopted by Crabb can be considered, as a primitive password manager. If the user does not have the algorithm or 'master password', the passwords are inaccessible.

One-time Passwords

However, while password algorithms may allow for potentially strong alphanumeric authentication secrets, stronger passwords do not alleviate many of the problems associated with authentication, namely phishing, key-logging or shoulder-surfing [85]. Furthermore, strong passwords do not stop an individual reusing the password [126]. A strong incentive for an attacker, as they will target weaker systems to gain passwords for stronger systems. Although, these attacks could be tackled by using another alternative approach, one-time passwords.

Lamport. envisioned a system where users generate a set of one-time passwords (OTPs) [165]. The user creates the set by entering their password locally. The password acts as a seed for a function that generates a OTP. The OTP is then fed back into the function to generate another OTP. The process continues, until an arbitrary number of OTPs are generated, for example, a 1000. The server stores the last OTP and discards the rest. The user discards the last OTP and authenticates with the penultimate one. The server receives the penultimate OTP from the user, feds it into the function and compares the output with the OTP stored. If they match, the user is authenticated and the previously stored OTP is discarded by the server. The next time the user authenticates they would enter the 998th OTP. The S/KEY authentication approach is based on Lamport's vision [110]. Long & Blumenthal argue a drawback is the fact an individual would need to carry around several lists of OTPs, one for each service they use, and instead propose a Manageable One-Time Password (M-OTP) [173]. Long & Blumenthal argue an individual would log-on to a M-OTP module, e.g. a browser extension, using a typically password and the module would issue OTPs to the respective services.

There have been variations of OTP deployed widely to both enterprise and consumer markets. RSA SecurID outlined in §2.2.2 is one such example of an OTP variation deployed widely to users. RSA SecurID has proved popular with various users but is not without problems, also outlined in §2.2.2. Consequently, attackers merely evolve and escalate attacks to undermine the authentication solution. Therefore, a different direction may be to teach and train users to produce better passwords, rather than offer alternative authentication solutions.

Policy and Education

An alternative direction is to improve password strength through education and policy. The password approach is powerful but is flawed due to the passwords created by individuals. Consequently, if strict policies are adopted and users are offered education, individuals could potentially create powerful passwords and keep them safe and secret. Vu, Bhargav & Proctor argue such password policies do not guard against weak passwords [283]. Similarly, Dell'Amico, Michiardi & Roudier in their analysis of password datasets, argue such policies do not prevent weak passwords and individuals spend more time creating their username, than their password [66]. Hoonakker found users regularly deviate from security policies, after surveying 836 employees at a large organisation [119]. Hoonakker found users either created simple passwords, regularly reused them or recorded them.

Furthermore, Hoonakker states the reality is probably much worse, as users are unlikely to fully confess to breaking rules outlined by the organisation. Hooknakker argues research should focus on user perceptions, such as convenience, as a balanced approach will be more appealing to users.

Shay *et al.* investigated the response of users to a change in password policy at Carnegie Mellon University by surveying them [249]. The institution moved to a new password policy, requiring users to create stronger alphanumeric authentication secrets. Shay et al. discovered that although users were annoyed by the change, they still complied with it, as they perceived a benefit. However, Shay *et al.* states users did experience difficulty in adopting the policy. Although able to create compliant passwords, 19% of users could not recall it. Furthermore, nearly half of participants claimed they simply altered the existing password to comply with the change in policy.

Barton & Barton argues such policies are scripted from the perspective of a system, rather than a person [15]. Therefore, it is not surprising when users deviate from them. Barton & Barton propose password creation should harness cognitive processes, relying on semantic and episodic memory. They argue passwords drawn from a user's personal experience and transformed using simple rules, e.g. transliteration, would be more memorable. However, Barton & Barton concede that such passwords would not be as strong as system-generated secrets.

Therefore, strict policies and education appear to have advantages but are not without faults and difficulties. However, tools and techniques can still be used, outside of strict policies to improve the strength of passwords.

Proactive Password Compliance

Proactive password compliance affords users feedback on the password they have created, informing individuals of the perceived strength of it. Users can then tailor and adapt the password inline with the feedback given. The approach is aimed at persuading users to create stronger passwords. The process typically conveys strength through a meter. The meter often comprises of a colour, segmented bar, oscillating between red and green, representing *low-strength* and *high-strength*, respectively. Users can then edit their password in response to the rating, until they arrive at a strong password.

Ur et al. argues the use of password meters do influence users to create lengthier passwords [278]. However, Ur et al. caution that for all the user-effort expelled there is little benefit, as passwords are only incrementally stronger. Nevertheless, Ur et al. does state meters that are more strict, but not overly, result in stronger passwords. However, such strict meters are not in wide-spread use.

The assessment of strength and acceptance of *low-strength* passwords is determined by the organisation. An organisation can either use proactive password compliance *independently of* or *in conjunction with* password policies. Yan et al. argues proactive password checking could improve compliance with security policies [299]. However, the process needs to be configured properly to ensure they reject weak passwords [300].

Nevertheless, even if strict policies are adopted and password meters only accept incredibly powerful password, there is no guarantee an individual can

remember such passwords. The problem is not necessarily the generation of strong password, system-generated gibberish is a powerful password. The problem is user being unable to manage such complex and meaningless password. Consequently, another direction is to adopt software that manages complex passwords for an individual, relieving them of the burden.

Password Managers

Password managers maintains and stores all the various passwords an individual creates, negating the need to memorise them. Moreover, an individual can use complex and lengthy system-generated passwords as they are not required to memorise themselves. The user merely has to create one strong, solid password to regulate access to the password manager. However, Chiasson et al. argues users are hesitant to relinquish control to password managers [50]. Chiasson et al. investigated the usability of two passwords managers with users and found that users had inaccurate or incomplete mental models of how they worked. Users thought a new password was being generated every-time they accessed their email, for example. Furthermore, Chiasson et al. emphasised that users were aggravated by the fact they did not know the actual password for many services. This represents a limitation of password managers, as users would need to have access to them everywhere, in-order to access all the services and systems they use.

Nevertheless, Jobusch and Oldehoeft outline the need for stronger passwords by highlighting some specific examples where weak passwords resulted in several systems being compromised [135]. Jobusch and Oldehoeft highlight how the Morris Worm relied on weak passwords to compromise 6,000 machines connected to the Internet [77].

2.3.2 Alternatives to Alphanumeric Authentication

The usability problems associated with passwords coupled with advances in modern technology has led to an increasing number of researchers considering knowledge-based alternatives.

The majority of alternative authentication research has focused on three of the five senses, namely: sound, vision and touch. The remaining senses of smell and taste, although not explored in terms of knowledge-based authentication have been proposed for use in biometric authentication [14, 216]. The primary motivation for exploring alternative authentication solutions that rely on senses is memorability. However, memorability is only a single aspect of many to an authentication solution.

Renaud and De Angeli propose a number of criteria to assess authentication mechanisms for use on the web, from a user perspective [220]. The web is arguably the biggest single consumer application of the enterprise-era, if not the genesis of the consumer-era. The Renaud and De Angeli criteria are percipient as they are concerned with the user and not with application or platform specific concerns. Therefore, criteria designed by Renaud and De Angeli are at least applicable to assess authentication mechanisms tailored for consumers. The criteria are, as follows:

1. *Accessible*
   The mechanism should not require any additional software, hardware or expertise. Furthermore, support must be provided for those suffering with disabilities, affecting sight and hearing.

2. *Convenient*
   The mechanism must not be time-consuming, either at the enrolment or authentication phases.

   (a) Enrolment Time
   Lengthy enrolment times could deter users from experimenting with an application or service.

   (b) Authentication Time
   A time-consuming authentication process could deter on-going use of an application or service.

3. *Memorable*
   The authentication secret and process is easy-to-remember.

4. *Secure*
   The authentication secret and process is suitable for use in authentication.

Renaud and De Angeli state the *secure* principle is multi-dimensional, defining these dimensions as *unpredictable*, *abundant* and *undisclosed*. De Angeli, Coventry, Johnson and Renaud outline the following dimensions to assess the security of authentication mechanisms [63]. They are, as follows:

(a) *Guess-ability*
The ability for an attacker to simply guess the authentication secret.

(b) *Observability*
The ability for an attacker to observe entry of the authentication secret.

(c) *Record-ability*
The ability for an attacker to utilise a user-generated recording, either *of* or *associated with* an authentication secret.

Therefore, the combination of the criteria outlined by Renaud and De Angeli and the security dimensions stated by De Angeli et al. offer an overall set of criteria. The criteria can be used to assess the suitability of an authentication mechanism for consumers. Consequently, many of the aforementioned aspects as well as memorability should be considered when determining the value of alternative sense-based solutions. The following sense-based solutions were considered:

- **Vision**
  Alternative authentication solutions based on graphics.

- **Sound**
  Alternative authentication solutions based on audio.

- **Touch**
  Alternative authentication solution based on touch.

The advantages and disadvantages of each sense as the basis for an authentication solution are discussed over the following sections.

Vision

Craik and Tulving's research provides interesting insight into the development of memory traces. Nonetheless, such limitation are arguably limited to the retention of words and depth of processing does not necessarily apply other stimulus, such as images, sound and touch. Bower and Karlin reported that depth of processing does indeed have relevance in the retention of faces [39]. Bower and Karlin requested individuals to make judgements of faces based on sex, likableness and honesty, i.e. progressing similarly from physical to semantic analysis. Bower and Karlin reported improved recognition for faces processed to greater depths, i.e. judged on honesty.

Therefore, depth of processing is applicable beyond the retention of words. However, Intraub and Nicklos argue the *depths* are reversed when processing images or scenes, in that analysis of traditionally shallow physical characteristics result in better recall [124]. However, Bower et al. argue comprehension of images improves retention, specifically when those images are meaningless to the individual [40].

The difference between images and words is not particularly surprising. Paivio proposes a dual-coding approach, in that we processes verbal and visual information through separate channels. The information from each is encoded differently, verbal information produces *symbolic codes* while visual information generates *analogue codes* [198]. Nelson et al. report images have superior recognition to the words that name them, i.e. *picture superiority effect*, due to the coding employed [188]. Furthermore, Anderson and Bower report that encoding of verbal information is strengthened when coupled with a suitable visual stimulus [4]. However, Kinjo and Snodgrass argue that although there is a picture superiority effect in explicit tasks, there is not necessarily such effect in implicit tasks [154]. Nevertheless, the potential of images sparked the imagination of many individuals.

Blonder first proposed the notion of a graphical password [35]. The proposal was that an individual would create a secret *sequence* of *pixels*, limited by the boundaries of a specific image. The user was expected to replicate the sequence of pixel points to authenticate.

Similar concepts had been explored previously. IBM Technical Disclosure Bulletin proposed the use of menu-items that only become active when an individual enters a sequence of specific pixels [120]. Nevertheless, the concept did not rely on a specific image but others have proposed the use of images in authentication. King proposed the use of images to improve the memorability of system-generated passwords [153]. King suggests users form associations between images and chunks of the complex, meaningless passwords.

However, the approach proposed by King is more an attempt to improve the password approach rather than offer an alternative graphical approach. Blonder

proposed approach expects users to create a sequence of pixels using an image and then recreate that sequence, over the same image to authenticate. Nevertheless, while interesting, Blonder never throughly explored his proposal but the concept inspired others to create graphical authentication mechanisms.

Nonetheless, Blonder proposed approach has inspired many researchers. Moreover, Blonder has clearer inspired designers and developers as graphical passwords have also by deployed by large organisations such as Google and Microsoft in their respective operating systems offerings. However, while inspiring and well received, graphical authentication is not without its challenges.

There are, for example, real concerns surrounding the observability of graphical authentication solutions. Wiedenbeck et al. argues that graphical authentication solutions are incredibly vulnerable to shoulder-surfing [293]. The reality is that many authentication alternatives rely on interaction that is vulnerable to observation and recording. The advantage in alphanumeric authentication in many ways is that the keyboard can be wielded or shielded from the wandering eyes of others.

Therefore, while there are potential advantages to using graphical authentication solutions, they are not without concerns of flaws. Consequently, it is not surprising that researchers have been exploring sound and touch-based authentication.

Sound

There is a great deal of potential for sound in authentication. LeDoux suggests that sound or rather music is less susceptible to interference as processing occurs at lower levels of the brain [167]. Such resilience is desirable due to the potential impact of interference on memory and in-turn knowledge-based authentication secrets [277]. Furthermore, much like vision, Peretz et al. suggests the brain may contain specialised channels for processing audio [202].

Moreover, Bigand and Poulin-Charronnat argue users are primed for music and additional training is not necessarily required [33]. Bigand and Poulin-Charronnat assessed different in terms of processing between trained and untrained individuals in regards music and argue that the human brain is already specialised and trained to process music through extensive exposure in everyday use [33]. Scherer and Zentner also argue that listeners typically agree on the particularly emotion conveyed in a particular emotion [242].

The power of emotions are important, as positive emotions are potentially important in forming memories [129]. Jäncke states that music stimulates emotions, that in turn act as memory enhancers [130]. However, the potential of music has both negative and positive applications. Alpert et al. demonstrated that when music within a product commercial stimulated emotion in an individual, purchase was far more likely [3]. While Hirsch discussed the weaponisation of music, outlining its use in torture and in deterring loitering [117].

Nevertheless, the strengths of sound have inspired authentication researchers to design and develop alternative solutions. Liddell et al. propose Audio Visual Associative Protocol (AVAP) an authentication solution that uses a combination of sound as well as images to authenticate users [169]. The approach relies

mnemonic associations, users are required to associate or pair a series of audio clips and images. The users is required to create five pairs of images and sounds. The user is then present an audio clip during authentication and select the associated image from a collection, containing distractor images. The user is expected to recreate all five associations to authenticate. Liddell et al. evaluated the authenticate approach and concluded that it was cumbersome and complex while providing little to no advantage to users [169]. Nevertheless, Liddell et al. stress that users enjoyed the authentication approach found it interesting and inspiring. Consequently, Liddell et al. argue efforts should be made by researchers to ensure authentication approaches are not only functional but delightful and enjoyable as well.

Figure 2.1: Musipass authentication screen [98, p. 127]

However, while AVAP may not have presented any advantage, it was arguably not a pure audio password. Gibson et al. argues pure audio passwords may be superior, specifically for certain users and context and propose Musipass [98]. The authentication approach expects users to create an audio password comprising of four audio clips. The user is then presented four collections of nine audio clips and asked to select their audio clip from each to authenticate. Gibson et al. report that when evaluated with users, Musipass does offer improved performance in terms of memorability over alphanumeric authentication. However, Gibson et al. stress the time taken to complete authentication was undesirable.

Furthermore, Gibson et al. outline the challenges of guess-ability, observability and record-ability. The reality is that there are several tools that allow an individual to share their music preferences, that may make it trivial to guess an authentication secret. Moreover, eavesdropping and recording an audio authentication secret is a concern as sound travels and could be heard or recorded by hidden attackers. Nevertheless, Gibson et al. argues observability is a bigger challenge for graphical password than audio, as individuals could use headphones to shield music. Lastly, Gibson et al. argues record-ability is an issue for almost all authentication approaches not just the novel sense-based authentication secrets.

Therefore, much like graphical passwords, while audio passwords certainly

have potential they also have numerous challenges. Consequently, it is unsurprising that some researchers have explored other senses for authentication, namely touch.

Touch

Touch is another sense that researchers have been inspired to explore for use in authentication. The memorability of tactile information is argued to be similar to audio [176]. Mahrer and Miles argues the capacity for tactile information is between four and six items [175]. However, there is disagreement about the specific capacity and decay of tactile information [176].

Furthermore, Millar argues tactile information is best recalled when coupled with other sensory information [181]. Therefore, touch or tactile information may not be a strong candidate for alternative authentication as there is potentially no real memory advantage. However, tactile information is potentially stronger than audio and visual alternatives in terms of observability and record-ability. An alternative authentication solution based on touch has the potential to afford the user the ability to enter an authentication secret, free from observation. Moreover, such solutions could be more accessible to users with specific requirements.

The potential of tactile information has sparked the imagination of a few researchers. Deyle and Roth proposes a tactile PIN entry mechanism that affords users the ability to enter an authentication secret free from observation [67]. The authentication secret compromises of a sequence of specific fingers, e.g. right index, left middle, left middle, left index finger. The user enters each finger through a series of rounds and using a custom tactile PIN entry device. Deyle and Roth argues the mechanism has the potential to improve authentication for all users and at the very least has the potential to improve authentication for blind users at ATMs.

Similarly, Bianchi et al. proposes the Secure Haptic Keypad (SHK) as an authentication approach that affords users the ability to authenticate in public spaces [29]. The authentication secrets comprises of a series of vibration patterns. The user enters the correct sequences of patterns by recognising the correct pattern among a set of distractor patterns. Bianchi et al. evaluates the SHK and concludes that while the approach shows potentials there are concerns attackers could observe and record audio vibrations emitted by the device. Moreover, Bianchi et al. emphasises more research is required to determine the actual memorability and learning processes of tactile information to ensure it is appropriate for use in authentication.

The reality is that there is far more research surrounding the senses of sight and sound, than touch. Nevertheless, researchers and designers are not deterred from proposing authentication approaches based on tactile information. Kuber and Yu proposes the Tactile Authentication System (TAS), an authentication approach that relies on recognition of PIN-patterns [161]. The authentication secret is a sequence of four PIN-patterns. The patterns are presented to an individual through a tactile mouse that communicates the patterns by raising pins. The user places one hand over the tactile mouse and is able to sense the various pattern arrangements. The user is presented a series of 3×3 grids, each

containing a single target pattern and 8 distractor patterns. The user needs to select the target pattern in each grid to authenticate. Kuber and Yu evaluate TAS and argue that the process performs similar to alphanumeric and graphical authentication mechanisms and that tactile information could form the basis of an alternative authentication solution.

Kuber and Sharma performed additional investigation of TAS and adapted it for use with blind users [160]. Kuber and Sharma evaluated the process with five blind users and argue the process is more secure but the user experience was less than desirable. Kuber and Sharma state that while the individuals could authenticate, the time taken to do so was undesirable and would need to be improved to make the authentication approach acceptable.

Therefore, tactile information clearly has a place for use in authentication. However, while tactile information could be used to reduce the observation of authentication entry and improve accessible, it may not be the basis of an alternative authentication solution.

2.3.3 Importance of Context

The aforementioned criteria outlined by Renaud and De Angeli is useful in the assessment of authentication mechanism. However, the resulting combination is arguably flawed as all the criteria are *relative*.

An authentication mechanism reliant on sentences, for example, is not necessarily accessible if a user is illiterate. The authentication mechanism is not particularly convenient either, if a user struggles to process sentences, e.g. dyslexics. The sentence is also not memorable, as memorising a string of characters would likely present a time-consuming, difficult task. Lastly, using a sentence as an authentication secret is not particularly secure, as illiterate users will adopt any number of coping strategies, such as relying on other literate individuals.

Hence, an authentication mechanism reliant on sentences would be deemed sub-optimal. However, in fairness, the aforementioned assessment criteria were outlined when the web was tailored to highly literate, wealthy westerners [302]. The criteria were crafted making the explicit assumption that any authentication mechanism was being used as part of a web application or service.

The context of an authentication mechanism in the aforementioned criteria was clear. The platform was an expensive computer with mouse and monitor peripherals. The system would need a wired connection, placing it in a small office, study or living-room. The people using the mechanism are those who can afford such infrastructure and navigate the world-wide-web. Finally, the assumption was that users would encounter an authentication mechanism when they accessed protected content or services.

While the aforementioned assumed context still persists, authentication now happens in many more contexts. The modern mobile computer or smartphone has pushed powerful computation and access to the Internet, into many more hands. Therefore, illiterate fishermen, on boats, in India will be able to use many applications and services, similar to the wealthy white westerners have done from their living rooms [151]. Moreover, in developed countries, mobile computers are

becoming increasingly common and commoditised, e.g. some print adverts in magazines are embedding smartphones to push live content [274].

Therefore, the suitability or success of an authentication mechanism, for a given task, is dependent on a given context. The same task is not necessarily performed by the same people, in the same place, on the same platform. Furthermore, the mechanism may serve one purpose in one setting but another, in another setting.

Consequently, the aforementioned criteria need to be extended by incorporating a *context specification*: A statement that provides a clear outline of the envisioned user, platform, place and purpose of an authentication mechanism. The context specification will be provide a solid foundation for a more nuanced assessment criteria. The context specification is defined as follows:

Context Specification A definition of the envisioned user-base, device and purpose of the authentication mechanism. • *People* The envisioned target users of an authentication mechanism. • *Platform* The envisioned device and/or software of an authentication mechanism. • *Place* The envisioned environments where an authentication mechanism will be used. • *Purpose* The reason for deploying an authentication mechanism.

Table 2.4: Context Specification for Authentication Mechanism Design

The previously defined assessment criteria can and should be assessed relative to the context specification. Therefore, if the targeted users of the authentication mechanism are illiterate, the accessible criteria should consider whether the mechanism is accessible to these specified users.

Therefore, in determining the design direction of the prototype authentication mechanism, the context of the process must be defined and considered. Consequently, when comparing the aforementioned authentication approaches, context will be considered along side the mechanism.

2.3.4 Comparison of Approaches

The primary design motivation of almost all knowledge-based or password alternatives is to improve the memorability of authentication secrets. The aim is to deter users from adopting undesirable coping strategies for knowledge-based authentication secrets. Consequently, researchers have proposed several knowledge-based alternatives to the password. The typical design direction taken by re-

searchers is either (1) iterate over the original text-based password or (2) construct sense-based solutions.

The first design direction is desirable as any iteration would retain the advantages of passwords but address the disadvantages. The proposed solutions, such as cognitive and associative passwords, do indeed improve the memorability of authentication secrets while maintaining many of the benefits. However, there are concerns surrounding the increased complexity of interaction that essentially leads to increased authentication times. Furthermore, attackers may undermined many of the proposed approaches simply by adopting different tactics, e.g. social engineering.

Consequently, policy and education of users may be the best direction to improve passwords [47]. The use of policy and education coupled with proactive password checkers or password algorithms could improve the strength of passwords. However, there is arguably limited mileage in education and policy as there is no guarantee that users will actually follow policy [119, 123]. Furthermore, even if guidelines are followed they have the potential to be more harmful than helpful [116]. Moreover, education and policy contributes nothing to the upcoming problems facing alphanumeric authentication, namely the diversity in consumer devices and lack of physical keyboard.

Popular consumer devices, such as the Apple iPhone, Nintendo Wii and Microsoft Kinect, forgo physical keyboards and mouse interaction in favour of touch and natural motion. Furthermore, the audience for such devices is arguably far wider, encompassing many more individuals with differing abilities. Moreover, audience members are not sitting in a single spot but are interacting with these devices in a variety of environments. Consequently, context is becoming an increasingly important aspect of authentication mechanism design. Table 2.4 outlines aspects that should considered about the context in the design of any authentication mechanism. Furthermore, specific contexts may favour sense-based solutions such as audio and graphical passwords rather than alphanumerics.

Therefore, the second design direction taken by researchers has been to develop sense-based alternatives to the password. Table 2.5 summarises the strengths and weakness of vision, sound and touch-based passwords. The superior sensory encoding of vision and sound-based stimulus suggests they are better suited to knowledge-based authentication than alphanumerics, in terms of memorability. The same may not be true of touch-based authentication secrets. Nevertheless,

	Vision	**Sound**	**Touch**
Memorability	Dedicated and expert processing of visual and audio information.		No clear benefit over passwords
Accessibility	Accessible to most with advantages for users with difficulties with language as well as those with specific needs.		
Convenience	There are real convenience challenges with all approaches, but could be potentially overcome through further research.		
Security	Concerns surrounding the observability and record-ability of authentication secrets		Resistant to observation

Table 2.5: Comparison and overview of sense-based authentication approaches.

touch-based have strengths in others way, specifically in terms of making authentication solutions more accessible for users with specific needs, e.g. blind users. Similarly, vision and sound-based authentication solution could be more accessible for users who struggle or have difficulties with alphanumerics, e.g. illiterate users. Therefore, sense-based authentication is potentially also more accessible than alphanumeric counterparts.

However, asides from convenience, there are real challenges in terms of security for sense-based authentication. There are concerns that visual and sound-based authentication secrets could be easily observed and recorded by onlookers surrounding an individual when they enter an authentication secret. In fairness, the same problem plagues alphanumeric authentication but at the least the user can wield interaction to shield the authentication secret. This is one area where touch-based authentication has an advantage as the user can authenticate, in view of others. Moreover, while tactile solutions may take longer than comparable graphical solutions, users value and perceive the sense of increased security from observation [159].

However, while sense-based authentication may be more accessible for some, it may not offer improvements for all. The reality is that sense-based authentication has complex and cumbersome interaction that takes time to use. Nevertheless, many of the convenience concerns could be connected to the relatively nascent state of alternative authentication solutions compared to the far more entrench alphanumeric authentication solution. Moreover, specific sense-based solutions may be more convenient in certain contexts.

Therefore, each approach clearly has strengths and weakness. Furthermore, weakness can be potentially addressed through further research and investigation. Nevertheless, a sense-based, pseudo silver-bullet solution that can replace passwords whole-sale across a range of tasks seems highlight unlikely. However, while passwords are still popular, they are arguably not optimal or even sensible on modern consumer devices. Password entry on the Nintendo Wii or Apple TV requires individuals to enter their authentication secret using a cumbersome on-screen keyboard, leaking the authentication secret to every individual in the same room.

Consequently, it seems unlikely passwords will persist as the only authentication solution for consumers. However, it appears equally unlikely that an iterative design or sense-based alternative will simply usurp the password. Therefore, a better strategy is to consider the envisioned context of the authentication mechanism and design a solution to best tackle the scenario. Therefore, using the aforementioned example of the Nintendo Wii and televisions, a graphical-based solution is envisioned as the best direction. The following reasons outline the motivation for selecting a graphical authentication solution, as follows:

- An alphanumeric solution seems unwise for television as the user must rely on an on-screen keyboard that is difficult or navigate and easily observed by onlookers.

- An audio-based authentication solution could be created but would require the user to wear a pair of headphones. Moreover, the authentication process

could be lengthy as the individual would need to listen to segments of music. A process that could prove frustrating to others in the same room.

- Graphical authentication would harness the key asset of modern television, i.e. a big, bright, beautiful and vibrant display.
- The superior encoding of images ensures they are more memorable than text [188].

Furthermore, graphical authentication may be the best direction for most modern consumer devices, such as smartphones and tablets, as they mimic television in matter ways. The modern smartphone and tablet do not rely on physical keyboard and possess detail and vibrant displays. Consequently, a graphical authentication, while not a catch-all solution, could still be applicable to a range of contexts and strong alternative to passwords in many settings.

Consequently, research focused on graphical authentication. Therefore, the aim was to create an alternative knowledge-based graphical authentication mechanism. The next step was to investigate and assess the various graphical authentication solutions. The goal was to determine how the graphical authentication mechanism would be framed, i.e. as either a recall or recognition memory task.

2.4 Taxonomy of Graphical Authentication Approaches

The primary advantage of graphical authentication is memorability. However, the encoding and storage of an authentication secret is only one aspect, another is how the authentication mechanism extracts it from an individual. An individual's memory can be accessed in three ways, as follows:

- *Recall*
  Information is extracted from memory when requested.
- *Response*
  Information is extracted from memory when cued.
- *Recognition*
  Information is extracted from memory when presented.

The information extracted is based on some stimulus. The end result is the same, *hopefully*, in that something is extracted from memory. Therefore, when considering alphanumeric authentication - individuals are typically requested to *recall* a password. Framed as a *response* task, similar to Zviran & Haga's associative passwords [304], a cue word would be presented and associated password would be expected. As a *recognition* task, a password would be presented and the user would be asked, if this was their password or not.

Similarly, the numerous graphical authentication approach proposed by researchers can be classified in terms of how they probe memory. There are several recall, response and recognition-based graphical authentication mechanisms. The proposed authentication mechanisms are presented in the ensuing sections.

2.4.1 Recall

The dominant password approach is framed as a recall memory task. An individual is presented a blinking cursor and is expect to recall the appropriate password string and enter it. Similarly, there are several graphical authentication mechanisms that adopt the same approach.

Naturally, recall-based graphical authentication approaches follow a similar structure, in that they expect users to recall an image and sketch it on a canvas. The authentication approach can heavily restrict drawing, e.g. connect the dots, or it can be incredibly expressive, e.g. freehand drawing. Moreover, the canvas itself can be blank or image to inspire the imagination

Draw-a-secret (DAS)

Draw-a-secret or DAS, proposed by Jermyn et al., is a recall-based graphical authentication mechanism [132]. The approach expects an individual to draw their authentication secret to access an application. DAS is designed for a stylus and touchscreen, although a mouse and monitor can substitute. The approach mirrors a paintbrush and easel, inspiring individuals to be creative and unrestrained when authenticating. An individual is not required to recreate the same drawing, every time they authenticate.

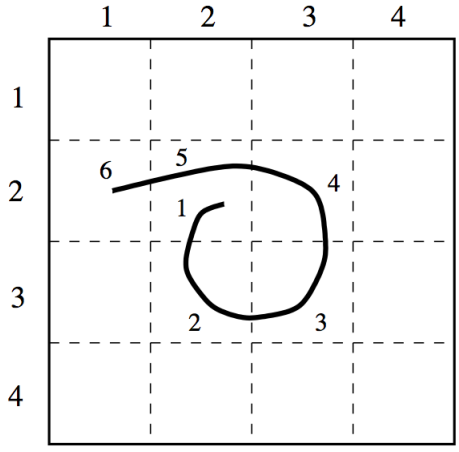

Figure 2.2: An example DAS secret Jermyn et al. [132, p. 5]

DAS does not rely on drawings, from a semantic perspective, but on the underlying grid sectors used to compose them. **Figure 2.2** illustrates a DAS secret, the system is interested in how the user drew the doodle, not the doodle itself. **Table 2.6** represents the encoding of the doodle, illustrated in Figure 2.2. The first six elements are the grid sectors used to create the doodle. The last element is a 'pen-up' event, the end of the drawing. An individual can draw as many doodles as they desire, each will be separated with a 'pen-up' event.

The important point is that the actual doodle is irrelevant, an individual only needs to recreate the sequence of grid sectors and 'pen-up' events. If a user can

achieve the same sequence with a different doodle, then they will be able to authenticate. Jermyn et al. argues this separation, of *drawing* and *encoding*, affords a much larger, practical password-space, than alphanumeric authentication. This practical password-space could be employed to tackle any number of traditional password problems, such as a reply or denial service attacks [158]. Alphanumeric authentication may have a large password-space but is undermined by the fact, that individuals are unable to remember many potential passwords. Jermyn et al. states:

> "We take as a main criterion the need to evaluate graphical passwords' security relative to that of textual passwords."
> **Jermyn et al. [132, p. 2]**

Jermyn et al. claims DAS is more secure than alphanumeric authentication, as there many more memorable authentication secrets. There is a larger *practical* password-space, than alphanumeric authentication. However, Jermyn et al. discusses, designs and evaluates DAS in the context of alphanumeric authentication, rather than terms of an authentication scenario.

	1	**2**	**3**	**4**	**5**	**6**	**Pen-up**
Secret	(2,2)	(3,2)	(3,3)	(2,3)	(2,2)	(2,1)	(5,5)

Table 2.6: The encoding of the drawing illustrated in Figure 2.2

An authentication process is part of a system or workflow. The envisioned workflow and end-user of DAS is rarely discussed. Jermyn et al. do detail a prototype of DAS as part of an encrypted memo application, on a personal digital assistant. However, they do not estimate the time taken to authenticate or outline any of the challenges faced in such a scenario. Jermyn et al. do not outline the expected user of the application or problems they may encounter. They do not evaluate the application or even the approach, with any actual users.

Unfortunately, Jermyn et al. does not provide any estimations or evidence of how DAS actually performs as part of a note taking/reviewing workflow. Furthermore, although Jermyn et al. constructs a prototype there is no evidence of how the mechanism performs on a personal digital assistant. Jermyn et al. states potential devices as the Palm Pilot, Apple Newton or Casio Cassiopeia E-10, none of these products were particularly successfully. Users simply did not adopt them, suggesting they did not deliver in terms of experience.

Apple ceased production of the Newton in 1998, a year before publication of DAS. Casio's Cassiopeia appears to have been poorly received, at least one reviewer stated "I think the E-10's screen just isnt big enough" and "Using the stylus to select and edit text is very inconvenient" [100]. The screen-size and stylus are important as Jermyn et al. states a potential drawback with DAS is a user may, unwittingly, cross grid-sectors when authenticating. Therefore, although an individual may recreate a drawing, the sequence may have a different encoding as strokes could graze other grid-sectors, a real possibility with a small-screen and awkward stylus.

In fairness, Jermyn et al. concedes the power of DAS is all theoretical. Without any field investigations or user studies, there is no evidence to state DAS is a practical authentication approach for the envisioned scenario or any others. The approach could be slow and awkward. Furthermore, although a larger, memorable password-space is possible in theory, it may not exist in practice. The password-space could be reduced, for example, if memorable DAS secrets exhibit a pattern.

Thorpe and Van Oorschot postulate that memorable DAS authentication secrets do exhibit such a pattern [269]. They argue *symmetrical graphical secrets* are a real concern, as symmetrical drawings have superior recall. Thorpe and Van Oorschot state an attacker could craft a dictionary of symmetrical secrets and use it to compromise DAS. They claim that while it may take an attacker 540 years to explore the entire DAS password-space, it would take only 6 days to explore the symmetric password-space. Thorpe and Van Oorschot argue that if users do create large numbers of symmetric passwords, DAS has a much smaller, practical password-space. They speculate users may create many such secrets, given that 5 of the 8 DAS examples Jermyn et al. presents, exhibit symmetry [132, 269]. However, Thorpe and Van Oorschot did not assess the mechanism with actual end-users, it is not clear if users would actually create predictable DAS secrets in practice.

Therefore, Nali and Thorpe conducted an informal user-study with 16 individuals to determine if user-generated DAS secrets exhibited any patterns. The participants were asked to draw 6 *doodles* and 6 *logos* on printed grids, using a pen. Nali and Thorpe discovered that users do exhibit patterns, 45% create symmetrical drawings, 56% of the drawings were centred and 80% of the drawings used less than 3 strokes. Nali and Thorpe argue if these results are symptomatic of a larger user-base, then DAS has a much smaller, practical password-space. Furthermore, 29% of drawings run along grid-lines, potentially creating difficulties in encoding and later recreation. Nali and Thorpe argue this could represent a real usability flaw in DAS as many DAS secrets would be rejected, even though the user believes them to be accurate. However, Nali and Thorpe did not actually inform users they were actually creating authentication secrets. The mechanism was not part of a workflow or performed on an actual device. Furthermore, individuals did not have to recreate any drawing and the time taken is not reported. Therefore, users could have behave differently, if they knew the drawings were used to protect personal notes.

However, Thorpe and Van Oorschot argue that although the study may have limitations, it clearly illustrates that users have a penchant for DAS secrets with fewer strokes [185, 270]. They suggest users favour such secrets, as they have to recall a start- and end- point for each stroke. Therefore, users cannot be relied upon to generate complex lengthy DAS secrets. Naturally, many designers would solve this problem by using bigger grids, with would offer more grid sectors. However, Thorpe and Van Oorschot deem such increases pointless as they do not deliver effective results, certainly not enough to justify such increases. Instead they propose a grid selection method, essentially a grid of grids. An individual initially selects a *drawing grid* from a larger *selection grid*. Thorpe and Van Oorschot argue this simple, additional step potentially strengthens the secret without burdening an individual with additional complexities.

Thorpe and Van Oorschot argue larger grids are a greater burden to end-users and their *selection grid* method is, potentially, superior. However, they do not assess the performance of either option with any actual end-users. The mechanism is not evaluated at all with users, either on its own or part of a workflow.

In fairness, Thorpe and Van Oorschot do argue that better security could be achieved through creating systems that offer a healthy distribution of password in *practice* rather than a purely academic distribution. Thorpe and Van Oorschot argue user-studies are necessary to determine the properties that need to be enhanced or relaxed to develop a mechanism that can perform well, practically.

Multi-grid Draw-a-secret (MGDAS)

Chalkias et al. proposed a method similar to Thorpe and Van Oorschot's *selection grid* but assessed with users [49]. Multi-Grid DAS (MGDAS) is designed to tackle the problem, identified by Nali and Thorpe, of individuals centring their drawing within the grid. This approach favours the amalgamation of several grids, with varying sectors over the traditional singular grid used in DAS. MGDAS has been designed to inspire creativity in the generation of secrets, reducing the number of individuals starting from the same cell or gravitating to the same sectors. Another aim of MGDAS is to reduce the number of *ordering errors* and *shift errors*, i.e. entering the secret in the wrong sequence and/or in sectors adjacent to the correct ones, respectively. Therefore, these additions should increase the probably password space of DAS without an excessive burden.

Chalkias et al. evaluated the mechanism with 15 children and 15 students, using a pen and paper. Initially, individuals endured a 25-minute 'short-course' on passwords and the MGDAS approach. Then, individuals were requested to create a traditional password, a DAS secret and a MGDAS secret. Then after one-hour, individuals were requested to recreate their graphical secrets. Chalkias et al. found that while MGDAS did not prevent ordering errors, it did reduce shift-errors. Furthermore, Chalkias et al. claim, less than 50% of MGDAS secrets exhibit the centring effect.

However, Chalkias et al. did not assess the mechanism as part of a workflow or on an actual device. Individuals did have to recreate the authentication secret, although this was only an hour after they created it. Furthermore, Chalkias et al. utilised children, aged between 6-11 years old, to act as non-technical users. It is not clear if this was a suitable substitute. Children likely have a different comprehension of the importance of an authentication secret. Chalkias et al. did not state if they envisioned MGDAS approach was designed for an application for children or young adults. Therefore, it is difficult to determine the applicability of a child's performance of MGDAS, with that of a non-technical adult. However, an encouraging aspect is that 60% of the children were able to authenticate with MGDAS, higher than DAS but lower than traditional passwords [49].

Overall, Chalkias et al. deem the results encouraging and state the approach could be refined by adding an background image. The background image would aid users in the creation and recall of their secret drawing.

Background Draw-a-secret (BDAS)

Dunphy and Yan propose such a mechanism, titled Background-image DAS or BDAS [74]. Traditionally DAS has been implemented with a white background but BDAS allows an individual to select their own background image. Dunphy and Yan argue this background image aides an individual in not only creating far more complex secrets but in their ability to recall such secrets.

Dunphy and Yan evaluated the performance of BDAS in two user studies, with 21 and 46 participants, respectively. The first user-study separated individuals into two groups. One group created DAS secrets, the other created BDAS secrets. Individuals in the BDAS group were asked to select one of five printed background images. The individuals then overlaid a printed 5x5 grid transparency. Individuals then created their secrets. The individuals were then asked to recreate their secret five minutes later with 8 individuals reached to recall their secret a week later. The second user study was similar in structure and procedure. The main difference was all individuals were recorded using a video-camera and all users drew on transparencies.

Dunphy and Yan reported favourable results in the addition of a user-selected background image. Individuals using BDAS typically created more complex passwords and retained them longer than individuals using traditional DAS.

However, some individuals did generate weak secrets using BDAS, so the approach does not completely remove simple secrets. Nevertheless, Dunphy and Yan never assumed the mechanism would ensure the removal of such secrets. They argue password checking should be used to steer individuals into creating superior secrets. Furthermore, some interesting minor recall errors occurred. A typical occurrence was recreating the secret perfectly but out of sequence. Another example includes an individual producing a near perfect replication of their secret but with a single element was reversed.

These outcomes could be particularly confusing for an individual and does highlight a nuance of DAS. In that the secret is the drawing of the authentication secret, rather than the outcome. The individual is required to recall how to they drew their authentication secret. Furthermore, slight recall errors such as reversing an element are likely to go unnoticed by an individual. Therefore, when an individual's attempt is rejected they may be able to recover from it.

However, probably the most interesting aspect uncovered by Dunphy and Yan about DAS is that the generation of secrets relies heavily on the artist ability of the creator. Those individuals which claimed to lack any real artist ability typically generated simplistic secrets or drew familiar characters in simplistic patterns. Those inspired slightly by the freedom to draw whatever they wanted, still opted from simplistic, familiar objects, such as a cup. Indeed, Dunphy and Yan reported that some individuals who felt uncomfortable with their artistic ability stated they would rather use alphanumeric alternatives. While those comfortable in their abilities had no real preference for DAS or BDAS. Although Dunphy and Yan claimed BDAS was more stimulating to use than DAS, suggesting it may aide those uninspired by a blank background.

Dunphy and Yan also offered some excellent insight in that they claimed some individuals had a great deal of trouble comprehending the BDAS process.

Therefore, they suggest that security designers consider these 'non-technical' individuals when crafting new mechanisms. This is a salient point, as we should not assume that individuals have the same familiarity with novel interaction mechanisms and objects, as they do, say, with keyboards and characters.

However, although BDAS appeared promising the approach was not evaluated as part of a workflow or on an actual device. In a recurring theme with all DAS descendants, the mechanism was evaluated out of context, using paper and pens. Users did not have to create an authentication secret and use it, as part of task. The impact of a workflow could change the secrets generated by users. If users have to authenticate regularly, they may favour simple secrets purely to complete a task faster. Therefore, it is difficult to estimate the relevancy of results as they refer only to a mechanism used in a laboratory, without an accompanying task.

In fairness, Dunphy and Yan recognise the limitations and concede that many new concepts work well an laboratory setting, only to fail when deployed to an actual user-base. Dunphy and Yan argue crafting an application to a specific scenario that actually impacts on the user would advance research in the usability of passwords. Dunphy and Yan state a drawback of the BDAS evaluation was that users have no stake in the mechanism. There was no incentive to successfully authenticate, users experienced no impact from their failure to recall. Dunphy and Yan argue that deployed in the field, as part of an application, BDAS would have greater ecological validity and allow for the exploration of other problems, such as shoulder-surfing.

Shoulder-surfing is when an attacker attempts to observe entry of an authentication secret. The majority of authentication mechanisms implicitly assumes a closed context, i.e. an individual is alone. This is an acceptable academic assumption but one that might not prove true, practically.

Qualitative Draw-a-secret (QDAS)

The reality is that a DAS secret is easily exposed to onlookers. If an attacker is able to observe entry of a DAS secret, they may be able to authenticate using the same drawing. Lin et al. proposes Qualitative DAS (QDAS) to tackle the problem of observation [170].

The key difference between DAS and QDAS is that stroke encoding evolves from a sequence of coordinates to a sequence of direction changes. Another addition is on-going grid transformations which are used to conceal the entry of secrets. The changes in encoding are aimed at improving the retention of secrets, effectively increasing the practicability of QDAS. While grid-transformation are used to extend the reach of DAS into new contexts, such as public places, where previously, it would be practically useless.

Lin et al. evaluated QDAS with 20 computing science students, in two groups of ten, through three stages. The first-stage of the evaluation requested students to create an authentication secret and recreate it, using pen and paper. The second-stage requested users to watch a video of an individual entering a DAS and QDAS secret. The final-stage expected users to recreate their secret, one week after creating it.

Lin et al. found QDAS more resistant to observation than DAS. The first-part

of the study revealed that all ten QDAS users could authenticate and only 1 DAS user could not recreate their drawing. The second-part of the study revealed that no user could steal a QDAS secret through observation, while 7 users were able steal a DAS secret through observation. The final-stage of the study, revealed that 6 DAS users could recall their drawing and 5 QDAS users could recall their authentication. However, one QDAS user withdrew from the evaluation in the final-stage.

Therefore, QDAS shows promise and could be a potential shoulder-surfing solution to DAS. However, Lin et al. did not evaluate the mechanism as part of a workflow or on an actual device. Therefore, it is difficult to determine if individuals can use the mechanism on a target device. Lin et al. do not indicate the target platform but the assumption is that QDAS, as with DAS, is designed for personal digital assistants. However, QDAS on a personal digital assistant could be awkward and difficult to use. There is no way to determine if the mechanism will be successful on mobile devices, without deploying it to them. In fairness, Lin et al. acknowledge the lack of context and state future focus will be on improving the ecological validity of QDAS evaluations. Therefore, although QDAS may be resilient to shoulder-surfing, until users assess it, there is no way to know if users would embrace it or not.

Chakrabarti et al. argues that rotating the canvas the user draws on could improve the resiliency of DAS to observation [48].

Rotation Draw-a-secret (RDAS)

Chakrabarti et al. propose RDAS, illustrated in Figure 2.3, an approach that allows users to rotate the DAS canvas. The separation of the doodle and the encoding is still true with RDAS, except the angle at which an individual rotates the canvas is appended to the encoding.

A vertical slider is positioned adjacent to the right of the DAS canvas, in the RDAS prototype. The user pulls the slider up and down to rotate the canvas. Therefore, a user could draw stroke, rotate the canvas, draw another stroke and

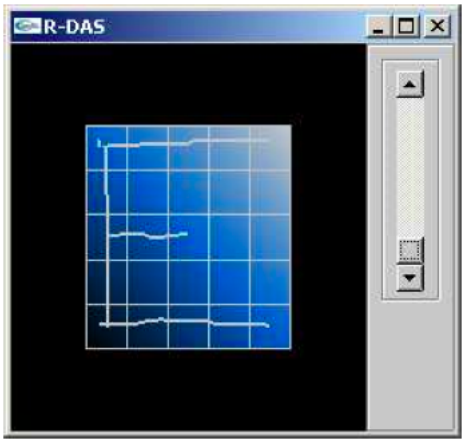

Figure 2.3: Chakrabarti et al. proposed RDAS approach [48, p. 564]

rotate it again. The rotation value is added, even if the user does not add a stroke. Chakrabarti et al. argues the addition of rotation not only increases the overall password-space of DAS but improves the practical password-space.

Chakrabarti et al. claims that RDAS tackles problems, such as symmetrical secrets, as the encodings for RDAS secrets would include rotation values. An attacker would have to have knowledge of the rotation values, as well as the secret, to compromise the system. Chakrabarti et al. argues rotation not only makes DAS more secure but more user-friendly.

However, Chakrabarti et al. does not actually evaluate R-DAS with users. Furthermore, Chakrabarti et al. does not outline the envisioned workflow or device that would incorporate R-DAS. There is no way to determine if R-DAS would be usable on certain devices. Moreover, there is no guarantee that users will actually rotate the canvas, as assumed by Chakrabarti et al.. If users do rotate the canvas, there is nothing to say that they will not do so in a predictable fashion.

In fairness, Chakrabarti et al. concedes that R-DAS would need to be deployed as part of a wider user study. Chakrabarti et al. states knowledge-based mechanism are heavily dependent on users and the mechanisms would need to be assessed with actual users to determine real world performance.

Yet Another Graphical Password

Gao et al. proposes Yet Another Graphical Password (YAGP). A DAS-inspired scheme with a design that afford users more expression and freedom [92]. Gao et al. argues the central drawback of DAS is that it has too many constraints, e.g. an individual is required to position the secret properly and avoid the edge of grid-sectors.

YAGP relies on a finer-grid than DAS as well as a different encoding and assessment approach. YAGP secrets are represented as pen-up, pen-moving and pen-down events. The pen-moving event is encoded as the position of the current cell in relation to the previous cell. These pen-moving events are book-ended with pen-up and down events. The entire encoding is considered a stroke. A YAGP secret can comprise one or more strokes.

The encoding produced during authentication is compared to the encoding generated during registration. The encoding is either a complete, partial or miss match of the original. The matching of encodings is achieved by analysing the stroke, its direction and by computing the Levenshtein Distance.

Gao et al. evaluated YAGP with 30 participants, all seasoned computer users. Participants interacted with the C++ prototype on a traditional personal computer. The grid was designed for a 3.5" inch screen, the same screen-size as the original Apple iPhone and iPod touch. Gao et al. investigated five different grid-granularities with individuals. The participants were initially given ten minutes to familiarise themselves with the mechanism before creating an authentication secret. They were then requested to make an authentication attempt. Participants were then asked to recreate their authentication secret, two days later.

Gao et al. reports that fine grids are better for YAGP authentication, as users do not need to concern themselves with a specific grid starting point, as is the case with DAS. Furthermore, Gao et al. stated that user personality is infused in

a drawn secret and consequently, makes it hard for an attacker to impersonate. However, Gao et al. highlight the fact that users are unlikely to replicate a YAGP secret well, every time. They argue this leads to the possibility that an individual is not guaranteed access to a system in every instance. It is not clear a user would favour such odds.

Furthermore, although Gao et al. evaluated the approach with actual users, on a device, they did not incorporate it into a workflow or on a device, similar to the one envisioned. Gao et al. envisions YAGP being used on personal digital assistants, devices that rely on a capacitive touch-screen and potentially a stylus. Gao et al. evaluated YAGP with a mouse and monitor. It is not clear how these interactions are comparable. An individual will not authenticate with YAGP at a bus-stop, on a rainy morning when the wind is blowing. They can do this with a modern-day smartphone or personal digital assistant.

YAGP may perform dramatically different in these settings. Furthermore, Gao et al. did not incorporate YAGP into an actual workflow. Therefore, there is no way to determine how users might feel about not accessing a system. YAGP may be better suited for systems a user interacts with rarely, where access is not crucial. However, Gao et al. never outlines the envisioned workflow of YAGP.

Passdoodle

Passdoodle, proposed by Goldberg et al., is an approach similar to DAS, except there is no grid. Users draw secrets on a blank canvas [102]. Furthermore, users can use various stroke colours and create a secret by drawing it four times. However, the secret is still encoded and assessed in a similar fashion to DAS. Once again, the process of drawing is assessed to gain access, not the doodle itself. The direction, order and number of strokes are all used to assess a doodle.

Goldberg et al. evaluated Passdoodle with 13 participants using paper and pens. Individuals were presented with three different scenarios, although only two scenarios were detailed. Goldberg et al. requested individuals to create a password and Passdoodle secret for an online bank and retailer, that users should imagine were accessed through a traditional personal computer.

Unfortunately, Goldberg et al. reported that although individuals could retain doodles just as well as alphanumeric passwords, they did not retain the process of creation. Therefore, while individuals were able recreate the doodle, the drawing process was not always the same. Therefore, although users could create the same doodle, if they deviated from the drawing process used to create it, the Passdoodle secret was rejected. Nevertheless, Goldberg et al. argues Passdoodle shows promise as users felt the doodle was easier to remember than passwords.

However, Goldberg et al. did not evaluate Passdoodle as part of a workflow or on actual device. Furthermore, although Goldberg et al. states Passdoodle is envisioned for use on a personal digital assistant and that a prototype has been constructed, they then evaluated it in terms of traditional personal computers. In fairness, Goldberg et al. concedes that although they tried to closely model a real application, the next step would be to evaluate Passdoodle as part of an actual system.

Goldberg et al. argues Passdoodle does show promise, as users found doodles

easier to recall than traditional passwords. The problems with the approach appear to stem from the fact the authentication secret is the process, not the product. Therefore, an approach based on the actual outcome of the drawing process, rather than the process itself, may make for a better authentication approach.

Scribble-a-Secret

Scribble-a-Secret (SAS), proposed by Oka et al., is such an authentication approach, relying on the user's actual sketch, rather than the process used to create it [195]. Figure 2.6 illustrates

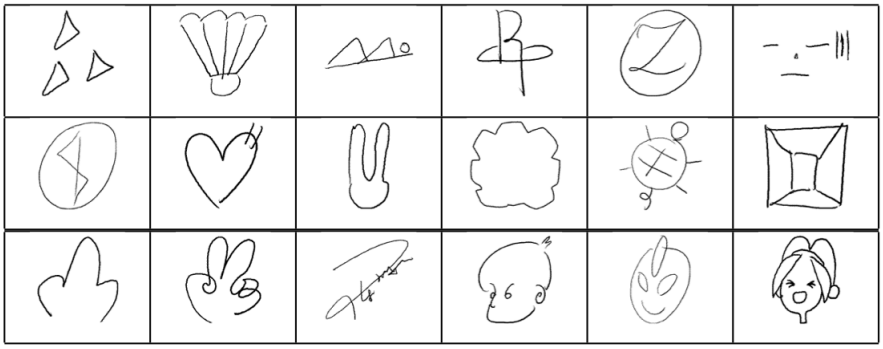

Figure 2.4: A selection of sketches produced by individuals using SAS [195, p. 3].

The approach expects a user to create a sketch during registration that acts as their SAS secret. Users then recreate this sketch to authenticate. If the recreated sketch is *similar* enough to the SAS secret, authentication is successful.

SAS determines similarity of sketches using *edge orientation patterns* which Oka et al. argue are vigorous in detecting drawn objects. The registration process requires an individual to sketch a drawing and then repeat the process several times. This allows the system to *learn* the sketch. In actuality the system is extracting various edge orientation patterns from the numerous sketches in an attempt to make a far more robust record of the sketch. This affords the user some freedom in that they can vary the sketch in terms of position and orientation, i.e. it does not need to be a perfect replica.

Oka et al. emphasise the learning phase is crucial if SAS is to be successful in correctly identifying sketches. The proposal itself is commendable and is incredibly easy to understand and easily explained to others. The biggest advantage of SAS is the utilisation of previous learning and not burdening users with novel interaction approaches. Users know how to sketch, they do not need to learn a new interaction method or memorise a sequence.

Oka et al. evaluated SAS with 87 individuals, using a Tablet PC. Users were asked to sketch their secret 10 times. The investigation primarily focused on assessing if edge orientation patterns could effectively distinguish between different sketches. Oka et al. reported the approach was met with success and that the

system did not have any significant problems in distinguishing sketches produced by different users.

However, although Oka et al. evaluated the mechanism on an actual device, they do not report any information regarding usability, e.g. time to sketch. Furthermore, SAS is never used as part of a workflow or actual system. Oka et al. do not indicate an envisioned application. Therefore, it is not clear if users would endure a lengthy registration process, i.e. sketching a secret 10 times, to authenticate for a simple task, e.g. browsing photographs.

Furthermore, users may not feel comfortable using the mechanism. Dunphy and Yan stated certain users were uncomfortable using BDAS because they felt they had no real artist ability [74]. Moreover, SAS is easily susceptible to shoulder-surfing, an attacker may observe a SAS sketch and be able to recreate it. In fairness, Oka et al. state this is a potential weakness of the approach and that further investigation will be required to determine the best way to address it.

PassShapes

Weiss and De Luca propose PassShapes as a mechanism which offers a memorable secret which can be conveyed quickly [289]. The approach is simple, an individual constructs secrets using eight different directional strokes, best illustrated as eight cardinal points. The secret is encoded using these direction strokes and a secret can be composed of several elements all separated by a pen-up event. Weiss and De Luca evaluated the usability and convenience of PassShapes with a user-study. Weiss and De Luca investigated two aspects, authenticating and password regeneration, with twelve users. Individuals used a Java-based prototype executing on a Tablet-PC. The users were asked to authenticate once and to change their authentication secret once. Weiss and De Luca reported that not only are PassShapes memorable but they are quicker to enter than other graphical authentication mechanism, although not as fast as PIN entry.

However, Weiss and De Luca are keen to stress that additional research and investigation is required before PassShapes can be confirmed as being quicker than all competing graphical mechanisms. Furthermore, although Weiss and De Luca assessed the approach with users, the mechanism was not deployed as part of a workflow or actual system. Therefore, it is not clear how users may respond to the mechanism over time. In fairness, Weiss and De Luca acknowledges PassShapes would need to be deployed in a longitudinal study to determine if the mechanism will be embraced by users.

Pass-Go

Tao and Adams propose an authentication mechanism inspired by DAS and the ancient Chinese board-game, Go [262]. Tao and Adams argue users struggle to recreate DAS secrets where strokes cross corners or run near the edge of grid-sectors. Therefore, Pass-Go embraces corners rather than instructing users to avoid them. Pass-Go expects users to create a secret by selecting corners, or rather intersections. The user's secret compromises of a series of intersections and

is encoded the same as a DAS secret. The intersections selected by an individual are connected by a line. Therefore, a user would select a initial intersection and then select another intersection point. A line is drawn between the two points. The secret follows this path until the user performs a pen-up event.

Tao and Adams evaluated a user-study with 167 participants, mostly students, over the fall term. A Java-based Pass-Go prototype protected a web-based teaching tool for science and engineering courses. The tool allowed students to access and view grades and course materials. The system also allowed teachers to post grades and course materials. Tao and Adams state all these accounts, including the teacher accounts, were protected using Pass-Go secrets.

Tao presented a tutorial to students using the course. The tutorial covered the mechanism, how to use it, how to improve password length etc. Tao and Adams state that to ensure deployment of the mechanism was to close to real-world use, the tutorial was not marketed to students, the time and place were not communicated to students. Furthermore, they did not record attendance or suggest students attend. However, Tao and Adams report that 80% of students did attend the tutorial.

Tao and Adams deem the mechanism was a success with 78% of authentication attempts successfully. They argued the main reason for performing the user study was to determine the convenience of the mechanism as well as learn the secrets individuals generate with Pass-Go. Tao and Adams report that users felt the mechanism was difficult to use of laptop using a trackpad. Furthermore, users stated that some mice in the computer laboratory, due to age or quality, made it difficult to use the mechanism.

However, Tao and Adams do not report the time it takes to authenticate with the mechanism. Tao and Adams do not discuss authentication time, at all. Therefore, it is difficult to determine aspects of convenience. Furthermore, although they state that Pass-Go is not limited to just mobile devices but anything with a web-browser connected to the Internet, they do not discuss the range of devices that access the tool. Furthermore, it is not clear what devices could actually interact with the Java-based prototype. Moreover, Pass-Go was assessed with highly competent computer users, students enrolled in science and engineering courses. The experience of non-technical students may be very different. Furthermore, Tao and Adams do not outline any alternative systems that allow individuals to assess grades and content, if they do not want to use the Pass-Go system. Students are practically held to ransom to use the mechanism. There is no alternative.

Nevertheless, Tao and Adams should be commended for actually assessing the authentication mechanism with a user-base. However, a concerning aspect of the user-study is that Pass-Go was used to view and post grades. Tao and Adams may want to create a realistic setting in which to evaluate the authentication approach. However, using an experimental authentication mechanism seems unwise. The mechanism could contain any number of flaws and bugs. The authentication approach could contain loop-holes or design-flaws, that an attacker could easily compromise.

If, for example, a student compromised a teaching account and interfered with grades, it could cause upset and distress among other members within the class.

Tao and Adams may have considered this eventuality and have control methods in place to deal with potential security breaches. Furthermore, it may be the scenario was the lowest risk of all considered. However, Tao and Adams do not detail the design process of the application. They do not detail the authentication scenarios considered and they do not report any risk evaluations. Therefore, although Tao and Adams have assessed the mechanism with a user-base, they have not detailed the envisioned use of the mechanism and if the user-study reflects this intended use.

However, aside from these concerns Tao and Adams have created a mechanism, that much like Jermyn et al.'s DAS, has inspired other researchers to refine and improve it.

Background Pass-go (BPG)

Por and Lim propose Background Pass-Go (BPG) [210]. The approach is inspired by BDAS and Pass-Go. The main difference between BPG and Pass-Go is that a user is able to select a background image. [210] argue the motivation behind affording the user the option to select a background image is that it may help in the recall of BPG secret. Furthermore, it may aide users in the creation of a BPG secret by affording user the option to decorate or focus on a particular part of an image.

Multi-grid Background Pass-go (MGBPG)

Por et al. refine BPG further with Multi-Grid Background Pass-Go (MGBPG) [211]. Por et al. state that MGBPG is inspired by MGDAS, BDAS and Pass-Go. The primary difference between the BPG approach and MGBPG is that users are able to vary the size of grid-sectors. The grid canvas has slider controls to the left and bottom of it. Users can vary the size of the grid-sector by moving sliders.

Por et al. claims that both BPG and MGBPG provide superior interaction and memorability to their descendants. However, Por et al. do not evaluate either mechanism with any actual users. Furthermore, Por et al. do not detail an envisioned system or how either mechanism would be incorporated into a workflow.

2.4.2 Cued-Recall

Framing authentication as a recall memory task could arguably be considered the strongest authentication solution as the user is provided no hints. Nevertheless, the concern is that users will merely record authentication secrets, motived by the prospect that they will not be able to recall them. Consequently, some researchers favour framing authentication as a cued-recall memory task, as the user is supported with a cue or hint.

There are several proposed graphical authentication mechanisms that are framed as a cued-recall memory task. Similar to the situation present in recall approaches with DAS, cued-recall approaches all follow a similar structure to

the PassPoints approach.

PassPoints

PassPoints, proposed by Wiedenbeck et al. [292], is a Blonder-esque mechanism. PassPoints expects an individual to select five pixel points from an image. The sequence of pixel points is the authentication secret, each pixel has a small tolerance radius as perfect replication is not expected. The aim of the approach is to resolve the problems of the recall-based password approach.

Wiedenbeck et al. argue that the inferiority of alphanumeric authentication stems from two contradictory rules which embody the process, they are, as follows:

1. Passwords should be memorable.
2. Passwords should be secure.

Secure passwords are authentication secrets that are unique, lack information and above all reside in the mind of the user. This requirement is in direct conflict with the first rule. Memorable passwords contain information, structure, pattern. Users creating secure passwords will resort to insecure behaviours to memorise

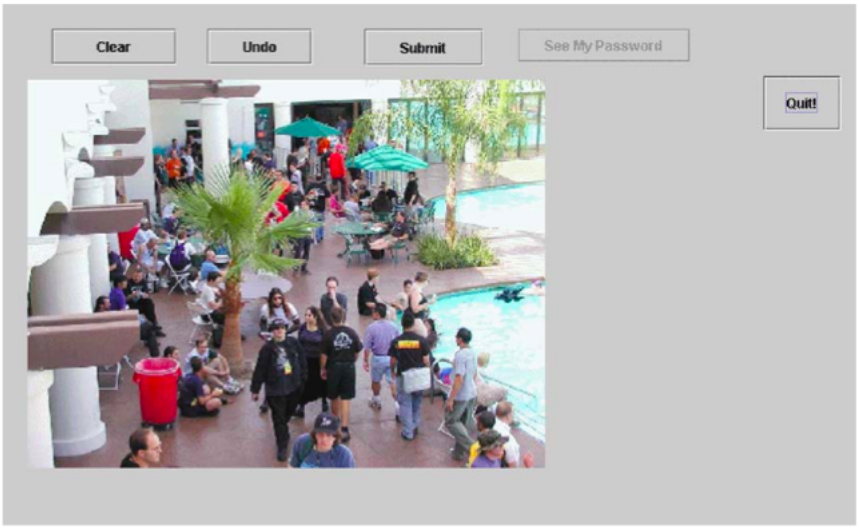

Figure 2.5: PassPoints prototype user interface [292, p. 113].

them, such as reuse or recording. Wiedenbeck et al. argue this is *the* fundamental problem of alphanumeric authentication, a problem they label, as:

> "the password problem".
> **Wiedenbeck et al. [292, p. 104]**

Wiedenbeck et al. claim PassPoints is the solution to this problem. The password-space of PassPoint is vast, even with the addition of the tolerance radius, as a single image can contain millions of pixels. The image used can be selected from a library or provided by the user, the only requirement is the image is complex enough to inspire users and protect the secret.

Wiedenbeck et al. evaluated PassPoints with 40 proficient computer users. The participants were separated into two separate groups, one using PassPoints the other using passwords, over a period of six weeks. The participants attended three sessions, one in the first week, one in the second week and one in the final week. The PassPoints prototype was Java-based and executed on a traditional personal computer, using a mouse and monitor.

Wiedenbeck et al. report that individuals found the creation of a PassPoints secret relatively straight-forward and that they discovered no real usability issues. However, participants did spend a significant amount of time entering their graphical secret when compared to those entering traditional alphanumeric passwords. The hope that this may subside over-time was also quashed when graphical secrets still took longer to enter than their counterparts, six-weeks on. Furthermore, retention was similar to that of alphanumeric authentication, there was no striking difference.

Indeed, during the learning phase, one-fifth of individuals using graphical passwords made around 17 to 20 incorrect entries. However, Wiedenbeck et al. argue this should be considered from another perspective, that graphical secrets are incredibly rare and few individuals will be familiar with them when compared to passwords. Furthermore, Wiedenbeck et al. state the majority of errors came from participants clicking near a PassPoint but not near enough, i.e. not within the tolerance area. They argue memorability could be improved by thoroughly considering the image used with PassPoints.

Therefore, Wiedenbeck et al. investigated PassPoints further with different tolerances and images [291]. Wiedenbeck et al. explored two different tolerances, 10x10 and 14x14 pixels, with 32 undergraduate students, all proficient computer users, in two separate groups. The participants were asked to create a PassPoints secret and authenticate, immediately after creation and then one week later. Wiedenbeck et al. report that no real difficulty was experience by any group in recreating the PassPoint secret after creation. However, they do state that participants experiencing the 10x10 tolerance condition made significantly more errors.

Wiedenbeck et al. argue that the errors not stem from poor interaction or usability errors but because a user is unable to precisely recall the location of their PassPoints. Furthermore, although Wiedenbeck et al. investigate the use of different images they did not find any significant difference between images. However, Wiedenbeck et al. state that they did observe specific images proving less inspiring than others. Furthermore, Wiedenbeck et al. claim that asides from tolerance concerns, users were able to grasp PassPoints quickly and able to authenticate with it. The retention rates and reduction of errors after learning suggests PassPoints may be able to tackle the *the password problem*.

However, Wiedenbeck et al. do not actually assess the mechanism as part of a workflow or a system. Furthermore, although Wiedenbeck et al. state PassPoints may be suitable for use on personal digital assistants and have constructed a prototype for such a platform. They do not assess the mechanism with actual users on the device. Users may not perform the same on a mobile device, on a noisy street as they do with a mouse and monitor in a quiet room. Furthermore, Wiedenbeck et al. do not have a real sense of the actual authentication secrets

individuals create with PassPoints. The mechanism may have a large password-space in theory, not necessarily in practice.

The apparent strength of the PassPoints approach is the large password-space afforded by the pixel-rich images. Thorpe and van Oorschot argue the practical password space of PassPoints is reduced because of 'hot-spots', i.e. popular pixels, as well as patterns within secret generation [271]. Thorpe and van Oorschot investigated both human-based attacks and automated attacks. They investigated two highly-detailed images for popular pixels. They discovered that 5 points in both images proved popular with individuals, between 24-31% for the first image and 20-24% for the second. Similarly, Dirik et al. developed a model that they claim can identify popular regions for points, they cautiously report that they were able to extract 70 - 80 % of points [72]. Furthermore, Thorpe and van Oorschot also suggest that patterns existed in sequence selection, these patterns are as follows:

- Right to left
- Left to right
- Top to bottom
- Bottom to top
- Clockwise
- Counter Clockwise
- Diagonal

Applying the knowledge of popular pixels and possible patterns could significantly reduce the strength of PassPoints. This is similar to using a word in alphanumeric authentication. If an attacker determines the first character of the word the number of possibilities is dramatically reduced.

Thorpe and van Oorschot then take the next logical step, which is to develop an automated attack, for this they need to determine *candidate click-points*. They define these points as (1) being identifiable and (2) distinguishable. Thorpe and van Oorschot argue that the fewer candidate click-points contains, the higher the probability of the image being undermined. A variety of image processing techniques were employed by Thorpe and van Oorschot to extract a list of these candidate click-points. Thorpe and van Oorschot report that 9.1% of the PassPoints secrets were extracted from one of the test images.

Salehi-Abari et al. argue this is because individuals have a penchant for specific points [236]. They argue that our attention is drawn to specific points, either through *top-down* or *bottom-up* visual processing. Top-down processing requires knowledge about a visual object in advance, i.e. where is wally, while bottom-up processing focuses on objects that are distinct or salient. However, when investigated Salehi-Abari et al. claim that although bottom-up processing may play a part in point selection for some individuals, it does not for all.

Interestingly, Salehi-Abari et al. did have some success, determining 48% of secrets for a test image. The same image was used by Thorpe and van Oorschot, for which they extracted 9.1% of secrets. Salehi-Abari et al. achieved this not through using visual attention but through image-independent guessing patterns.

Therefore, while PassPoint may have a large, theoretical password-space, it may not have a large, practical password-space. Users could behave in a predictable way when creating authentication secrets, that undermines the strength of the mechanism. Wiedenbeck et al. may have found PassPoints promising in the laboratory but the performance of the mechanism may be different in the field. Chiasson et al. conducted two user studies, one in the laboratory and another in the field, to contrast and compare the performance of PassPoints, in and out of the laboratory [51].

Chiasson et al. evaluated PassPoints, initially in a laboratory with 41 students. These participants were all proficient computer users, asked to use a PHP prototype on a traditional computer, using a mouse and monitor. The participants were given a presentation on the mechanism and were shown how to create a secret and authenticate. The students were advised to create an authentication secret, as if it was guarding their bank account. A session compromised of five stages. The first stage focused on creating a PassPoints authentication secret. The second stage was confirmation of the authentication secret. The third stage required users to answer two questions. This was followed by a stage that expected users to complete a puzzle. The fifth and final stage required individuals to authenticate with the CCP secret. Lastly, participants were asked to complete a questionnaire. Students could complete as many of these five-stages stages as they wanted within the hour, with the first two being discounted as training sessions.

Chiasson et al. then evaluated PassPoints, in the field, with 191 computing science students in the fall term. These participants used PassPoint to access a web application, containing course notes. Students were educated about the mechanism through tutorial demonstrations, emailed instructions and a support website. Participants within the field-study experience one of two conditions, either PassPoints with a 13x13 tolerance or a 19x19 tolerance.

Chiasson et al. report dramatic differences between the laboratory investigation of Passwords and the field-study. Chiasson et al. state the differences they observed call into validity of using laboratory investigation to graphical authentication mechanisms. The differences between the two studies are detailed below, they are, as follows:

- The laboratory investigation of PassPoints produced far more positive results than the field-study.
- The participants in the laboratory investigation were more successful, in terms of authentication attempts and accuracy, than participants in the field-study.
- The participants in the field-study had shorter authentication times, than those in the laboratory investigation.

- Participants in the laboratory investigation had a more positive attitude towards PassPoints. Participants were positive about the time taken to authenticate, even though they were slower than those in the field.

Chiasson et al. postulates the following explanations as to why these differences might have occurred between the laboratory investigation and field-study. They are, as follows:

- The participants in the laboratory investigation had highly concentrated sessions with PassPoints. Users created and authenticated with several PassPoint secrets within a single hour.

- The participants in the laboratory investigation had practice sessions and experts on-hand to answer any questions about the mechanism. The participants in the field-study had a brief explanation and emailed instructions.

- Authentication attempts immediately followed creation in the laboratory investigation, whereas field participants had large gaps of time between authentication attempts.

- The **primary focus** of the laboratory investigation was the PassPoint authentication secret. The **primary focus** of the field-investigation was access to course-notes. The PassPoints authentication secret was **secondary**, a means to an end.

Chiasson et al. argue that while the laboratory investigation of PassPoints showed it to be promising, when it was deployed in the field, it was not as well received. The mechanism may not have performed as well in the field but Chiasson et al. argue it still showed that PassPoints could work in practice.

However, Chiasson et al. argue the difference between the two studies emphasises the need to consider the evaluation of graphical authentication mechanisms. Chiasson et al. state the majority of evaluation is in a controlled, laboratory setting, rather than in the field. The differences between Chiasson et al. studies suggests laboratory settings do not provide accurate insight into the usability of graphical authentication mechanisms.

However, although there were differences between the studies. They both show that popular pixels or hotspots may be a problem for the PassPoint authentication mechanism.

Cued Click Points

Chiasson et al. proposes Cued Click Points or CCP, a cued-based graphical authentication mechanism [52]. The approach is a variation on PassPoints, in the sense that an individual selects a pixel from an image. However, the main difference is that an individual is required to repeat this action over several images. Therefore, the secret is a sequence of pixels selected from a series of images. The images are intended as cues, hopefully aiding recall by helping the individual remember the pixel point they select, if any at all.

The process begins with an individual registering a secret. During the authentication phase the individual is initially presented with a image that forms part of their secret. If they successfully select the correct pixel from that image, then the individual is presented with another image from their secret and so on. If an individual is unsuccessful in selecting the correct pixel point, the next image will not form part of their sequence and this process continues.

Therefore, unlike PassPoints, subtle feedback is provided users at every step, not just at the end. Furthermore, any attackers are taken through a false path, shown images which do not form part of the individual's secret. Chiasson et al. suggest that this process is not only an overall improvement on PassPoints but should also require less of an individual, as they will not be required to recall a sequence of pixels for one image. An individual using CCP is only required to remember the pixel they have selected in a image, the sequence is not required. This is because the images are randomly presented as cues.

Chiasson et al. evaluated CCP with 24 computing science students, all enrolled on a security module. The participants were given a presentation on the mechanism and were shown how to create a secret and authenticate. The tolerance used was 19x19 pixels. Participants used a J# prototype executing on a traditional personal computer, using a monitor and mouse. The procedure followed was the same, as in [52], except towards the end of the hour, participants were asked what their personal preference was, between CCP and PassPoints.

Chiasson et al. reports that individuals strongly prefer CCP over PassPoints but cautioned that greater investigation is required. Furthermore, Chiasson et al. argue that CCP is far more usable as users are able to authenticate rapidly and accurately. They state that users valued the feedback offered by CCP, in that inaccurate pixel selection would result in an unknown subsequent image. Moreover, Chiasson et al. state that participants become more versed at the mechanism as they use it. However, concerns still remain about CCP, primarily those inherited from PassPoints, such as popular pixels.

Persuasive Cued Click Points

Chiasson et al. tackles this specific problem with Persuasive Cued Click Points or PCCP [53]. This approach uses CCP with the addition of a *persuasive viewport*. During the registration phase a viewport is randomly positioned over the image. The viewport is emphasised by reducing the brightness of the rest of the image. The individual is only allowed to select a pixel point from within the viewport, which they can shuffle if they do not like the position. The viewport itself is positioned randomly rather to any real heuristic, as this would simply generate new hotspots.

However, PCCP is not necessarily persuasive but more forced. This is because individuals have to select a pixel from within the viewport. Furthermore, whatever aspects of a pixel influence an individual to select it, are somewhat removed as individuals are required to make an assessment based on far fewer pixels. However, the benefit is arguably a far stronger secret.

Chiasson et al. evaluated the PCCP with 39 students, all proficient computer users. Participants used a J# prototype, executing on a traditional personal

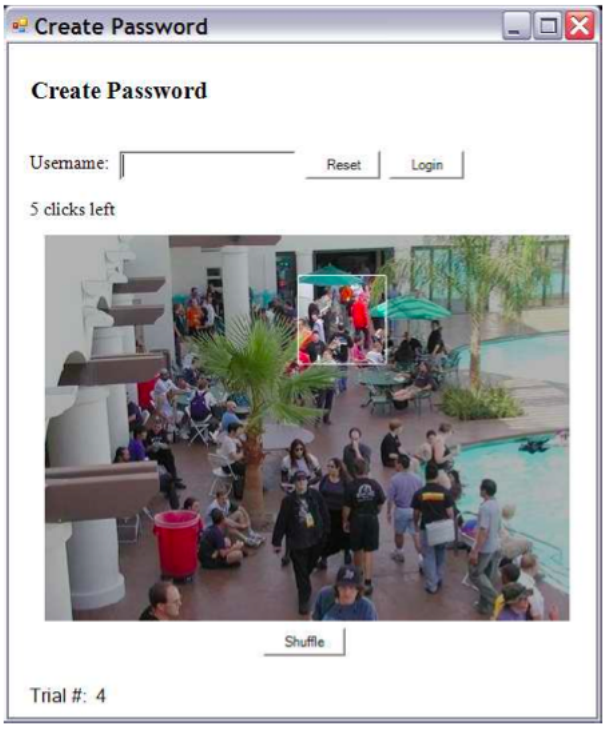

Figure 2.6: PassPoints prototype user interface [53, p. 124].

computer using a mouse and monitor. Students followed, essentially, the same procedure previously used in as in [52].

Chiasson et al. report that the viewport is successful in reducing hotspots and the spread of points. Furthermore, individuals claim they feel their PCCP secret is far more secure than a CCP secret. However, success in reproducing the secret was no better. Therefore, given the security benefits, the PCCP approach would be a step forward. However, it's clear far more investigation will be required to determine if the viewport has any negative impact.

However, Chiasson et al. did not actually evaluate the mechanism as part of a workflow or an actual system. Furthermore, they do not detail the intended use of the mechanism. If the mechanism was intended for use on mobile-phones or televisions, for example, then evaluating it on a traditional computer may not be particularly relevant. Users may act one way in a laboratory, with say, their lecturer, than they would with their friends at the local coffee shop.

2.4.3 Recognition

The concern with pure recall and cued-recall authentication approaches is that individuals may be unable to remember the authentication secret, precisely. The user may have memorised the authentication secret but may not be able to articulate it, as expected. Therefore, some researchers favour recognition-based authentication approaches, as users are merely presented the authentication secret.

Nevertheless, such an approach presents challenges to designers as the authentication approach must still remain secure. Consequently, recognition-based approach follow a similar pattern of presenting users target objects mixed with distractor objects. The authentication secret is represented as a single or series of target objects, mixed with several similar distractor objects. The distractor objects are positioned to dissuade and confuse attackers. The PassFaces approach is a common example of such a design, relying on targets and distractor images.

PassFaces

PassFaces is one of the few commercial graphical authentication mechanisms [266]. The authentication approach assigns an individual a collection of faces as their authentication secret. The user is then presented a series of challenge stages that compromises of nine images within a grid. The eight images in the grid that are not part of the user's authentication secret are termed *distractors* while the single image from the user's authentication secret is termed the *target* image. The user completes a challenge stage for each image in their authentication collection.

Story

Davis et al. proposes an authentication approach that requires users to create a 'story' [62]. The authentication mechanism is similar to PassFaces in that a user is presented a series of grids containing images. However, the images are not restricted to faces but include other objects, such as vehicles and animals. The user is expected to create a story using the images. The story is the authentication secret. An example authentication secret is: a couple have lost the keys for their car. Therefore, the resulting image sequence is: a male face, a female face, a set of keys and a car.

Davis et al. evaluated the scheme, alongside PassFaces, with participants using several categories, including vehicles, animals, food, vistas, household objects and models. Davis et al. reports that users found the authentication secrets far less memorable than those used in PassFaces. Davis et al. states they are not surprised as faces have been shown to have superior recognition performance. However, Davis et al. does state performance with story was impacted, for the following two reasons:

1. *Users ignored instructions*
   Users did not create a story with the images they selected. Davis et al. states that instead users selected image they felt were memorable, distinct or attractive.

2. *Users struggled with sequence*
   Users were able to recognise images but were unable to remember the necessary sequence to enter them.

Therefore, Davis et al. states if sequence is important in any authentication approach it should be emphasised to the user.

Déjà Vu

Dhamija and Perrig propose Déjà Vu, a recognition-based graphical authentication mechanism [68]. Dhamija and Perrig argue a recognition-based approach will require less of individuals than a similar recall-based approach. Déjà Vu consist of three stages: (1) secret creation, (2) training and (3) authentication.

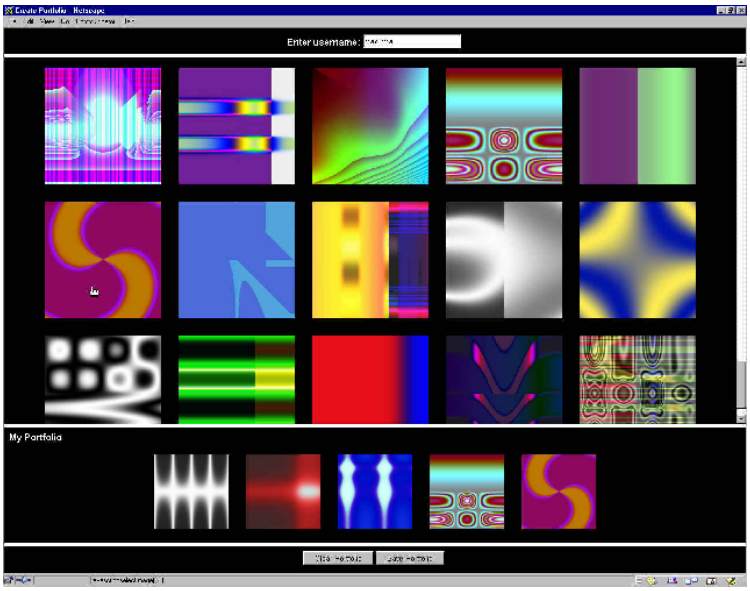

Figure 2.7: Déjà Vu portfolio creation [68, p. 5].

Secret creation with Déjà Vu requires an individual to select a set of images from a large collection, as illustrated in Figure 2.7. Each image is abstract in nature and the collection is generated using a mathematical formula, the output depends on an initial seed. The beauty of this design is that the actual images do not need to be stored, just the small initial seed.

The set of images selected during the creation stage constitutes the individual's secret. The next step is to identify these images amongst other decoys in a training stage. The individual is not presented with every image in their secret set, just a few of them. Furthermore, there is no sequence involved in selection, all the user has to do is identify the images extracted from their secret set. The purpose of the training stage is to strengthen retention of the images within the user's secret set.

The authentication stage essentially mirrors that of the training stage. Dhamija & Perrig evaluated Déjà Vu against other competing systems and assessed it using two different image-types, i.e. abstract against semantic. Déjà Vu performed well against competing recall-based approaches such as passwords and PINs. Indeed, Dhamija & Perrig reported that more individuals were unable to recall their username than recognition the images within their secret sets. Individuals using Déjà Vu felt that it was overall easier to use but at the expense of time and security.

However, Dhamija & Perrig report an interesting insight in regards the image-type used in Déjà Vu. When using semantic images, i.e. photographic scenes,

some individuals selected the same images. One specific image was selected by 9 out of 20 individuals. Furthermore, these images are far easier to explain and describe, thus, as a consequence an individual's secret set of images is easier to convey to someone else. For example, the aforementioned popular semantic image contained the Golden Gate bridge.

Conversely, abstract images rarely overlapped and descriptions of them rarely, if ever, matched. This in theory strengthened the probable password space of the approach as there was no real pattern or popular images. Naturally, further investigation will be required but Dhamija & Perrig research does highlight the impact the image-type can have on an approach.

Photographic Authentication

Pering et al. propose Photographic Authentication, an authentication approach for use on public machines [203]. The approach was designed for use untrusted computer systems, such as those found in libraries and airports. Pering et al. state the mechanism relies on a home or personal server that manages authentication. The personal server manages authentication with the user, not the untrusted system. The server indicates to an external system the result of authentication.

The process presents a series of challenge grids, containing four images. Three of the these are decoy images sourced from other users, alongside a target image sourced from a user's personal collection. The user must indicate the target image in ten challenge grids to authenticate. Pering et al. evaluated the mechanism with 8 participants, using a web-based prototype of the mechanism.

Pering et al. report that participants were able to quickly determine target images from decoy images. They state that when a user has a large collection of personal images, authentication is typically slower. Pering et al. indicate two users had between 1200 and 1800 images, while the rest had between 40 and 500. Furthermore, while some participants felt the mechanism was stronger than passwords, others were not convinced.

Pering et al. also state that another concern is the fact that attacker does not need to identify a target image to compromise the system. The attacker simply needs to identify the images that do not belong to a user. Pering et al. argue the mechanism could tackle this concern by using the target images of other users. However, Pering et al. concede they are not sure how this would scale to a large group of users.

Tiles

Nicholson et al. propose Tiles, an authentication approach that requires users to simply memorise a *single* image [189]. The image is segmented into nine equal tiles. The mechanism then presents a single segment or tile from the image in a grid, alongside eight others that have been randomly selected from decoy images. The user needs to complete four of these challenges grids to access a system.

Nicholson et al. argue that relying on a single image, instead of four, reduces the burden placed on the user. However, they do emphasise that being dependent on a single image may make the mechanism more vulnerable to (a) sharing and

(b) observation. Therefore, they evaluated these two specific attacks with two user, or attacker, studies.

The mechanism was exposed to 60 students, acting as attackers. The participants were presented nine different challenge grids, all using a different target image. The participants were asked to identify the segment belonging to the target image in each grid. The decoy images used are categorised as either similar, medium or dissimilar to the target image. The target images were not presented to participants. The participants were either presented (a) a description of the target image or (b) a segment of it.

Nicholson et al. reported that decoy images, similar to the target, proved the most resilient to both attacks. Furthermore, they state that specific target images appeared more resilient than others. Therefore, Nicholson et al. suggest careful consideration of the target image could strengthen the mechanism against both of these attacks. However, while Nicholson et al. emphasise the focus is on the ability of attackers to compromise the system, they do not perform any assessment of the performance of the mechanism with any actual users. Therefore, while certain images be more resilient to attackers they may also prove more difficult for users. That may result in users adopting unknown coping strategies.

2.4.4 Comparison of Approaches

Section 2.2 outlined and classified various graphical authentication mechanisms as either recall, cued-recall or recognition memory tasks. The alphanumeric authentication approach can be considered a recall task, as the user is required to enter a password from memory. An alternative authentication approach, such as DAS, is also classified as a recall task, as the user is presented a canvas and required to draw an authentication secret from memory.

An alternative authentication approach, such as Story, is classified as a recognition task, as the user is presented with images of objects and is required to indicate if they recognise them or not. The interesting aspect of these authentication approaches is whether or not framing an authentication approach as a recall or recognition memory task improves authentication. The intention is to craft an authentication approach that is convenient and painless for the user. If a specific memory task is classed as more convenient or less intensive than another, then it should form the foundation of an alternative authentication approach.

Recall memory tasks provide few, if any, prompts to aid memory retrieval, while recognition memory tasks present actual elements to prompt retrieval. Tulving and Watkins states that retrieval cues interact with memorised information [275]. The combination of the cue and stored information allows an individual to extract memorised information. However, a cue may not be particularly well suited or carefully crafted for retrieval. Furthermore, cues themselves may be processed and encoded differently to the information sought. An example would be using smell to retrieve a drawing. Consequently, Tulving and Watkins state recognition is essentially an easier memory task, than recall, as the retrieval cues are more pertinent. The element being sought is essentially the element presented.

Therefore, the decision was taken to frame the alternative authentication mechanism as a recognition task. The expectation was an recognition-based ap-

proach would ease retrieval of an authentication secret. Consequently, the next step was to determine an appropriate image-type for use in authentication.

2.5 Taxonomy of Image-types for Recognition-based Authentication

Standing argues that recognition potential for images, when considered under certain conditions, is almost limitless [257]. Similarly, Nickerson states that humans have strong recognition abilities for complex images [190, 191]. Nickerson investigated the recognition performance of images, exposing individuals to 600 images. The participants were initially presented 200 images, immediately followed by another set of 400. The second set was a mixture of the previous set and 200 unseen images. The participants were asked to determine if an image in the second set was new or old, i.e. unseen or seen. Nickerson reports strong recognition performance at 95% accuracy. The lowest performing participants still exhibited above 80% accuracy.

Nevertheless, while individuals may exhibit strong recognition performance for images, not all images necessarily elicit the same performance [125]. Furthermore, not all images that an individual can recognise are necessary suitable or realistic for use in an authentication mechanism. Consequently, the image-type selected for use in a recognition-based graphical authentication mechanism must represent a balance between usability and security.

There is not necessarily a single image-type, researchers and designers have taken several different directions in designing alternative authentication mechanisms. Section 2.4.3 outlined and discussed several recognition-based graphical authentication mechanisms. The image-types used in each of these mechanisms are:

- **Scenes**
  Tiles [189] and Photographic Authentication [203] both rely on images of scenes. The image can comprise of objects, individuals and/or landscapes.

- **Objects**
  Story [62] relies on images of objects of general items as well as faces.

- **Faces**
  Passfaces [266] relies on images of faces.

The aforementioned image-types are not the only options for recognition-based graphical authentication mechanisms, another potential image-type is:

- **Caricatures**
  Faces with emphasises on distinctive facial features, e.g. exaggerated chin or nose.

However, there are other images, such as "random art" images [68], graphical icons [293] and mikons [223]. Nevertheless, "random art", graphical icons and mikons are essentially synthetic images that are favoured as they offer practical

and technical advantages. The synthetic images are not selected as they represent a usability and security balance. The primary advantage of "random art" images is that they can be systematically generated, affording a wide password space and avoiding complex copyright issues. Similarly, graphical icons without legal constraints are numerous and relatively easily sourced. Nevertheless, the usability advantages of either approach is not particularly clear.

Graphical icons were not designed to be used in recognition-based graphical authentication but instead serve to afford users the ability navigate operating systems effectively and effectively. Furthermore, graphical icons may simply not be distinctive enough for an individual to recognise rapidly among a see of others [293]. Similarly, the unusual and novel shapes, colours and composition of "random art" may not be particularly memorable or distinctive for users to rapidly process in an authentication approach.

However, while the aforementioned image-types of *scenes*, *objects* and *faces* represent a potentially better balance, they are not without concerns. The images of all have their advantage and disadvantages in terms of use in an authentication approach, these outlined in the ensuing sections.

2.5.1 Scenes

Oliva states that individuals can capture the gist of a scene with a single glance [196]. Oliva points to 'channel-surfing' as an example of such sophistication, i.e. an individual progresses through television channels, sparing seconds on scenes, seeking content they want to watch. An individual does not linger for long in their quest for entertainment.

The speed and sophistication of individuals in processing scenes suggests such images may be acceptable for use in authentication. An individual is not merely processing colours and compositions. Biederman states comprehension of scenes is equally important, objects and faces do not appear in a vacuum [30]. The elements within a scene are connected and when processing the scene, an individual is assessing the semantics of it.

Consequently, the contents and composition of a scene may impact on the memorability of it. Parikh et al. states that scenes that comprise of an enclosed environments and faces are more memorable than picturesque landscapes [200]. Furthermore, Parikh et al. states novel or visually pleasant scenes are not more memorable, despite common assumptions [200] .

Furthermore, Xiao et al. argues that scenes have an intrinsic memorability value that is consistent across individuals, regardless of experience and knowledge [297]. Therefore, the speed and sophistication of the human visual system in processing and storing scenes suggests such images are suitable for use in an authentication mechanism.

Nevertheless, the aspects of scenes that contribute to the memorability of an image could be harnessed by a system to classify and assess scenes. Xiao et al. demonstrates such a system that is able to rapidly process and assess the content of a scene and the potential memorability of it. Therefore, given an individual may favour memorable images, users may make predictable choices that another system could easily determine. However, Parikh et al. argues while some scenes

are more memorable than others, users are not necessarily good at determining such scenes [200].

Unfortunately, while this may address the problem of predictable choices, in terms of memory, it raises concerns about the memorability of scenes that users may select for use in authentication. However, users could merely be issued scenes than select them or users could be presented vetted scenes, that already are deemed to have high memorability. Consequently, the design of the authentication mechanism is crucial in extracting the maximum value from images of scenes.

However, given the complexity of scenes and space required to display them, there are concerns on how scenes could be presented to users. If users were presented target scenes alongside distractors, arguably considerable space would be needed. Furthermore, it is not clear such space would be available on many devices. Moreover, even with increased screen spaces, users may need to have a short viewing distance, as well as a high-quality screen, to process small and specific details of a scene. Nevertheless, the design of the authentication approach could arguably tackle such problems. Tiles, outlined and discussed in §2.4.3, tackled the problem of scene scale by segmenting scenes into small sections. Users are presented a single target segment alongside several distractor segments, over several stages [189].

Nonetheless, Nicholson et al. did not assess the approach with actual users, consequently, it is difficult to determine if the approach is practical. Furthermore, arguably segmenting scenes into sections is little different from using a single object or face from a scene. Therefore, objects and faces could also be potential images-types for use in authentication.

2.5.2 Objects

Individuals exhibit speed and skill in observing and processing objects, unsurprising as they encounter them everyday. Therefore, images of objects may represent a strong image-type for an authentication mechanism. Jolicoeur et al. states that recognition of objects represents the intersection of perception and memory [137]. Individuals do not simply observe an object they attempt to comprehend it, its reasons for existence, its purpose.

Consequently, Jolicoeur et al. argues the way an individual processes an object could be dependent on the level of knowledge and experience an individual has of the object [137]. Furthermore, Konkle et al. state that memorability of an object is very much associated with experience, in that objects will have increased memorability, if an individual has some prior knowledge or experience with it, upon encountering it [156]. Moreover, individuals can potentially be trained and taught to process and distinguish even subtle changes between objects [31]. Therefore, images of familiar objects, even particularly unusual or esoteric objects could form the basis of an authentication mechanism.

However, Biederman et al. states that while individuals display prowess for object recognition, it is not without expense [32]. Biederman et al. evaluated the cost of object recognition by presenting individuals an increasing number of objects in a clock-face composition [32]. Biederman et al. states that there was notable decrease in detectability as more objects were presented, suggesting

object recognition is an "attention-demanding" process [32]. Therefore, there are constraints on the design of the authentication mechanism as to ensure the mechanism is not too demanding of users, in terms of the number of objects.

Furthermore, there are concerns that users may make predictable image choices and exhibit patterns in their object selection. The patterns could be exploited by an attacker and result in the compromise of the authentication mechanism. Spain and Perona argue that objects are indeed not all equal, individuals assign different levels of importance to objects [254]. Consequently, there are real concerns that users may make predictable choices based on such importance levels when selecting images for use in authentication and an attacker could utilise such patterns.

However, Davis et al. states that users of the Story scheme, outlined and discussed in §2.4.3, do not appear to make such predictable choices [62]. Therefore, images of objects may still be a suitable image-type for use in an authentication mechanism. Nevertheless, there are still concerns surrounding the sourcing of object images. The images would need to assessed and primed to ensure they are distinct enough so users can rapidly detect and select target images. However, users may fair better with a single prototype object that varies between images, such as faces.

2.5.3 Faces

The human vision system is able to perceive and process faces almost instantaneously [273] with recognition for faces superior to other objects [103]. The speed and accuracy of the system, as with many others, depends on a number of factors [303]. The majority of neuroscience and psychophsyic research agrees individuals have a penchant for people and are incredibly efficient at processing faces. An ability that could form the strong foundation of an alternative authentication mechanism.

Ellis states that face processing is a dedicate process [79] as recognition of a face is affected by inversion [301] and prosopagnosia suffers are able to recognise other objects [36] as well as the preference of newborns for "face-like" patterns [105]. These discoveries may indeed be indicative of a dedicated process. However, many researchers argue that conclusions drawn from such discoveries are not necessarily the only explanation.

Valentine states that while inversion certainly does effect the processing of faces so does short presentation time [279]. Valentine argues neither are particularly indicative of a dedicated process. Furthermore, Diamond and Carey demonstrate the same 'inversion-effect' in canine specialists' recognition of dogs [70], suggesting such a transformation is not exclusive to face processing. Furthermore, there are also counter arguments and explanations for the face processing abilities of prosopagnosia patients and newborns.

The reality is that prosopagnosia patients do not exhibit the same symptoms. Malone et al. demonstrated the ability of one prosopagnosia to match some famous faces but not unfamiliar faces, while another prosopagnosia patient was able to match unfamiliar faces but few famous faces [178]. Moreover, Bruyer argues the "information processing system" of a prosopagnosia suffer may recognise

and process a face while the patient does not [43]. Consequently, prosopagnosia suffers may not be indicative of a dedicated face process.

Therefore, the preference of newborns for face-like patterns may be the only aspect that points to a dedicated process. However, while Johnson et al. was able to confirm the preference for newborns, the preference itself declined after a few months [136]. Consequently, while neonates do exhibit a preference for faces, it does not persist, further questioning the existence of a dedicated process.

Nevertheless, other evidence could suggest the possibility of a dedicated process for face processing. Perrett et al. identifies face specific neurones in rhesus monkeys, arguing the existence of such cells as evidence of specialised areas for faces [205]. Nevertheless, such specialised areas may not be exclusive to faces [206]. Similarly, Kanwisher et al. outlines the fusiform face area, an expert region that is more responsive to faces than other objects [149]. However, yet again, such specific and specialised processing and encoding may not be exclusive to faces [95]. Nonetheless, Tanaka suggest faces are processed differently and argues the entry point for face recognition is different from object recognition [261]. Tanaka argues that object recognition entry point is at a basic level, e.g. 'a human being' while the entry point for faces is at a subordinate level, e.g. 'Barack Obama'. Consequently, there is still much debate about *exactly* how faces are special.

However, there appears consensus or general agreement that the brain reacts to some object differently and can be optimised for certain objects. Therefore, while some individuals may be experts on specific objects and visually process them efficiently, all individuals are face experts without training. Furthermore, individuals exhibit strong retention for faces as well.

Bahrick et al. assessed retention of classmate faces for 392 participants, who had graduated, between 2 weeks and 57 years, from high-school. The participants were asked to complete a picture recognition task. The task consisted of ten stages, each stage involved the presentation of a row of five faces on a single card, one face was the target while the other four were distractors. The recognition performance was 90% for those who recently graduated and appears to remain consistent for approximately 35 years before deteriorating.

Furthermore, distinct and attractive faces appear to improve memorability and recognition. Cross et al. report that in every case where individuals deemed a face as attractive or beautiful, it was subsequently recognised [61]. Similarly, Fleishman et al. found that individuals recognised faces that were initially deemed as either very attractive or least attractive than neutral faces [84]. Moreover, even the expression on a face can impact on an individuals recognition performance [91]. Consequently, distinctive features on a face likely ensure future recognition.

However, such aspects of face processing could prove problematic for use in an authentication mechanism. Davis et al. discovered that individuals make predictable choices when they are required to select images for use in graphical authentication [62]. The following aspects of faces may influence user choices:

- **Attraction**
  The purpose of attraction is propagate desirable genes [268, 267]. The level of attraction associated with face may decrease but remains similar overtime

[264]. Furthermore, no specific feature is responsible, instead perception of attraction is based on symmetry, averageness and secondary sexual traits [81]. Lastly, an attractive face is a memorable face [61]. Consequently, an authentication secret comprising of attractive faces is a memorable authentication secret.

- **Race**
  The reality is that race does impact on the recognition with faces [280, 46, 265] with individuals exhibiting superior memorability for own-race faces, than other-race faces [179, 230]. However, Cross et al. found that African Americans exhibited similar recognition performance for own-race and other-race faces [61]. Cross states performance may stem from the fact that participants had 'experience' of white faces from television programmes and films.

- **Familiarity**
  Goldstein and Chance state faces exhibit superior recognition performance to inkblots and snowflakes [103]. Goldstein and Chance argue familiarity with a class may not only aid in assimilation of new members but also strengthen resilience to decay [103]. Furthermore, Ida Gobbini et al. state familiar faces, belonging to friends and family, exhibit a stronger response than famous or unfamiliar faces. Ida Gobbini et al. state that familiar faces induce neural activity in areas traditionally associated with social knowledge. Moreover, Ida Gobbini et al. argue unfamiliar faces induce activity in areas, hypothesised as a 'social brake', when assessing potential threats [121].

Unfrotunately, these general aspects appear to impact on the use of face images in authentication mechanism. Davis et al. demonstrates that when creating authentication secrets based on images of faces, users tend to select own-race faces [62]. They state that 50% of authentication secrets generated by Asian females display an own-race bias, as do 90% of authentication secrets generated by African Americans. However, image-sets could be vetted and configured to tackle such problems.

Nevertheless, aspects such as attraction could still present a problem. Langlois et al. argues maxims, such as *beauty is in the eye of the beholder*, are myths and there is agreement on attraction within and across cultures [166]. Therefore, individuals could exhibit patterns that an attacker or system could exploit. Eisenthal et al. states that the beauty of a face is a 'universal concept', learnable by a computer [78]. Consequently, any system attempting to introduce faces would need to address these concerns in their design [62].

However, the design of an authentication mechanism can also potentially enhance the use of faces. Dukes and Bevan states retention of faces are improved if observed in multiple poses [73]. Similarly, recognition is improved when individuals encounter distinctive familiar faces [281]. Therefore, the design of an authentication mechanism could display multiple poses of the same face or rely on an image-set of familiar faces, for example. Consequently, the design of an

alternative authentication mechanism is incredibly important as it has the potential to reduce the concerns of using specific image-types as well as enhancing their use.

However, while face images may be suitable for use in authentication, an alternative direction could be caricatures. The caricature image-type may afford many of the benefits of face images but avoid complications such as attraction and race.

2.5.4 Caricatures

Tversky and Baratz argue that caricatures, rather than face images, closely resemble the schematic representation of them in memory, as they they maintain stable features and emphasis distinct features [276]. Benson and Perrett argues that memory may process and encode faces in terms of deviations from a prototype [25]. Consequently, caricatures may harness and stimulate such representations. However, Rhodes et al. argues while line drawing caricatures may have an advantage, photographic caricatures do not [229]. Furthermore, caricatures of familiar faces are identified faster than veridical line drawings and anti-caricatures [228]. Moreover, Rhodes suggests that the preference or accuracy of a caricature depends on the familiarity of the face to the viewer [226].

Therefore, traditional hand-drawn caricatures may be a suitable image-type for an alternative authentication mechanism. Nevertheless, Rhodes and Wooding states that while caricatures may be closer to the mental representation of faces, traditional photographs are recognised faster [227]. Moreover, Lewis argues advantages in using caricatures may apply to any distorted face not simply caricatures [168]. Consequently, there is possibly no real advantage in using traditional caricatures over distorted faces in an authentication mechanism.

Furthermore, caricatures do not necessarily address many of the problems of using face images. Byatt and Rhodes states that there is an own-race bias when recognising caricatures [45]. Therefore, users may make predictable choices in creating an authentication secret that an attacker or system could exploit. Gao et al. demonstrates machine recognition of emotions from caricature faces [94]. Consequently, a system could be potentially devised to process and exploit patterns within caricature-based authentication secrets.

However, arguably the authentication mechanism could be designed to curb the impact of such concerns. Nevertheless, the immediate benefit of caricatures over face images is not clear. Furthermore, sourcing caricatures could be more complex than face images, as there is no readily available source of them when compared to other image types. Nonetheless, while caricatures may have concerns and problems they still represent a suitable image-type for an alternative authentication mechanism.

2.5.5 Comparison of Image-types

The human vision system exhibits speed and prowess in processing, encoding, storing and retrieving images. Moreover, images are more memorable than the words that name them [188]. Therefore, images appear a strong foundation for

a recognition-based graphical authentication mechanism. Nevertheless, not all images are necessarily equal and some image-types may be more superior than others for the purposes of authentication.

The previous section, §2.5, outlined and discussed four images types, namely: scenes, objects, faces and caricatures. The use of scenes of images appears advantageous as individuals can capture the gist of a scene with a single glance [196]. However, constructing an authentication mechanism around images of scenes may be difficult due to limitations on screen space and screen quality. Biederman states comprehension of a scene is key, as individuals do not process scene elements in a vacuum [30]. Consequently, there must be enough space and detail for an individual to comprehend a scene. An aspect that may be difficult to perform on small screens or large low-quality screens.

However, such concerns are becoming irrelevant as screens become increasingly sophisticated and readily available. Nevertheless, even beyond screen quality there are still limitations on screen real estate and how many scene images can be practically presented to users at any one time. Moreover, even if scene images are segmented, comprehension is potentially impacted as individuals do not have access to the entire scene. Consequently, scene images appear awkward, in terms of use for authentication. Nonetheless, scene elements such as objects and faces could be better images-types for use in authentication.

An authentication mechanism based on objects seems possible as users process and encounter them everyday. Nonetheless, object recognition and processing is an attention-demanding process [32]. Furthermore, while individuals encounter objects regularly, there are still concerns surrounding the practicality of them forming the basis of an authentication mechanism. An individual's prior knowledge of an object impacts on the processing and memorability of the object [137, 156]. Consequently, constructing an object image-set could be problematic, unless it comprises of common objects. Furthermore, object would need to be vetted to sufficiently distinct to ensure users do not confuse objects. Moreover, as users assign different levels of importance to objects, they may create predictable authentication secrets.

However, users do not appear to exhibit such patterns when creating authentication secrets comprising of object images [62]. Nonetheless, there are still concerns about sourcing and filtering object images for use in an consumer-level authentication approach. There is no guarantee that users may be familiar with all images or that they can rapidly differentiate between objects. Moreover, any algorithms or pre-processing to vet an image-set may simply introduce complexity that an attacker could exploit to comprise systems. Nevertheless, an authentication approach could be tailored to specific and services to harness the potential of object images. A fan-based canine website, for example, could utilise a graphical authentication approach. The approach itself, could utilise an image-set of particular breed of dog. A fan or expert in dogs would be able to rapidly recognise and process such images as they have experience with them.

Therefore, while there is undeniable potential for object images it is arguably restricted to specific systems and services. Nevertheless, all users have knowledge and experience with certain objects, namely faces. The reality is that almost all individuals are face experts and a graphical authentication approach based

on faces arguably has many advantages. Users are able to retain and recognise faces for decades [13]. The improved recognition ability may be an indication of a dedicated process for faces within the brains of individuals [79]. However, it may simply be indicative of an expert area of the brain, that individuals can harness to efficiently process specific objects [95].

Nevertheless, there are concerns surrounding use of face images in authentication. The primary concern is users making predictable choices based on attraction and race, as such patterns could be exploited by attackers. Moreover, there are lesser concerns such as actually sourcing suitable face images. Consequently, caricatures rather than faces images are often considered as a potential alternative. The assumption is that caricatures retain the advantages of face images but address disadvantages such as predictable users patterns.

Unfortunately, race does impact on the processing of caricatures [45]. Moreover, sourcing is potentially far more complex, although one potential solution is for users to create their own caricatures. However, such a requirement would only balloon the cost of the authentication approach in terms of time and would expect users to have skill and inclination to create such caricatures. Furthermore, traditional photographs are recognised faster than caricatures [227]. Consequently, caricatures do not appear to present any real advantages over face images, in terms of use in authentication.

Therefore, after considering the four specific image-types, the following conclusions can be drawn:

- Scene images have potential but require considerably screen real estate to display to users, limiting the number that can be presented at one time.
- Object images have potential but image-sets would need to be tailored to ensure users have familiarity with images. The concern is that users may struggle with unfamiliar objects.
- Caricatures have potential but there are concerns about sourcing and they arguably offer no benefits over face images.
- Face images can be recognised rapidly and retained for decades but there are concerns users may exhibit predictable patterns when creating authentication secrets.

Therefore, the decision was taken to focus on face images and use them in the alternative authentication mechanism. Nevertheless, while face images may be advantageous for use in authentication, another aspect had to be addressed, namely obfuscation of the authentication secret. The primary concern of using images as part of a recognition-based authentication approach is that onlookers could potentially observe entry of the authentication secret. Therefore, an obfuscation method had to be determine to ensure entry of the authentication secret is shielded from onlookers.

2.6 Taxonomy of Observation-types for Graphical Authentication

The observation of authentication entry by onlookers is a major concern for almost all knowledge-based authentication approaches. Password entry is routinely obfuscated so that other individuals are unable to easily observe what is being typed, which reassures end-users. Unfortunately, Tari et al. discovered that when users type long and obscure passwords, entry is more easily observed by shoulder-surfers than when typing simple and familiar words [263]. Unfortunately, alphanumeric authentication secrets generated by users to protect bank accounts and tax records are likely to exhibit exactly these characteristics, so efforts spent by a user to be "secure" actually backfire. Even so, most users are fairly confident that observers cannot guess their password with any degree of accuracy [288], even though such confidence is probably misplaced [263].

Researchers have proposed a number of different ways of alleviating these problems. Therefore, there are many different strategies for shielding authentication secrets within a shared space. Tan et al. propose a spy-resistant on-screen keyboard specifically designed for kiosks. Unfortunately, users are somewhat uncomfortable using the keyboard [260]. Nevertheless, Tan et al. proposed approach is only one of several and researchers have proposed several solutions that are designed to shield authentication entry from onlookers. The solutions can be primarily divided into the following two categories:

- **Hardware**
  The authentication approach is complimented with dedicated hardware to obfuscate entry of the authentication secret.

- **Software**
  The authentication approach is designed in software to obfuscate entry of the authentication secret.

There have been several hardware and software approaches proposed by researchers, these outlined and discussed over the ensuing sections.

2.6.1 Hardware

There are various hardware solutions to obfuscation of authentication. The proposed solutions vary from utilising existing hardware within devices to designing and constructing bespoke hardware solutions. There are many possible solutions but two prominent directions are:

- **Tactile**
  Orozco et al. outline a graphical authentication approach that relies on almost imperceptible tactile interactions, such as pressure and velocity, during authentication [197]. Similarly, Sasamoto et al. proposes an an approach that relies on tactile interaction that is concealed and shielded from onlookers [237]. Nevertheless, drawbacks for both include the need for specialised hardware and relatively complex user interaction. Consequently, De Luca

et al. propose an approach that relies vibration, as several consumer devices already support such tactile interaction [64]. Similarly, Bianchi et al. outline approach that relies on vibration and audio cues [28]. Nonetheless, both approaches are still relatively complex and this is reflected in lengthier authentication times.

- **Gaze**
  Hoanca and Mock propose the use of eye-tracking systems to determine the direction of the user's gaze [118]. Similarly, Kumar et al. propose gaze-based entry of passwords as a method of concealing authentication interaction from onlookers [162]. The approach requires an individual to essentially gaze at an object, such as a button, and then either *dwell* on the object or push a trigger to indicate selection. Furthermore, the same interaction could be applied beyond alphanumerics to graphical authentication. Dunphy et al. proposes a variant of PassFaces, original outlined and discussed in §2.4.3, reliant on gaze rather than mouse clicks [75]. Similarly, Forget et al. proposes a gaze-based variant of Cued Click Points [87], original outlined and discussed in §2.4.2. Nevertheless, all such approaches require fairly laboured and complex interaction that results in lengthier authentication times, e.g. Forget et al. contrasts performance of 53.5 seconds with 7.5 seconds mean authentications times for Cued Gaze Points and Cued Click Points, respectively. Furthermore, such interaction still requires arguably sophisticated and dedicated hardware that many consumer electronics devices may not support.

The aforementioned design directions are interesting and represent some of the most prominent solutions for tackling observation of authentication. Nevertheless, while interesting and novel, hardware is expensive and limited while software is arguablly inexpensive and versatile. Consequently, several researchers have also proposed software-based solutions for obfuscation of authentication.

2.6.2 Software

There are several advantages to using software to shield the entry of authentication secret. The primary advantage is that the solution can potentially be deployed on several different systems or retroactively added to shipped devices. Moreover, a software solution can be far less expensive to deploy than a bespoke hardware alternative. These advantages have motivated researchers to deploy software-based solutions to shield authentication from onlookers. There are several design directions for software solution but two prominent choices are:

- **Trapdoors**
  Roth et al. argues an authentication mechanism can be presented as game that contains a cognitive trapdoor [233]. Roth et al. outlines a game-like PIN approach, illustrated in Figure 2.8, where the cognitive trapdoor is the PIN itself. The user authenticates by entering the correct sequence of colours, users will struggle if they do not have knowledge of the PIN. Similarly, Hayashi et al. propose a visual trapdoor approach that relies

on knowledge of a pristine image to navigate a series of degrade images [113]. Similarly, Wiedenbeck et al. proposes use of a cognitive trapdoor in graphical authentication by expecting users to create a convex hull using knowledge of secret graphical icons [293]. Moreover, Weinshall outlines an approach that presents a board of images that users navigate using knowledge of secret images [286]. Nevertheless, while interesting, both authentication approaches require training and relatively complex interaction.

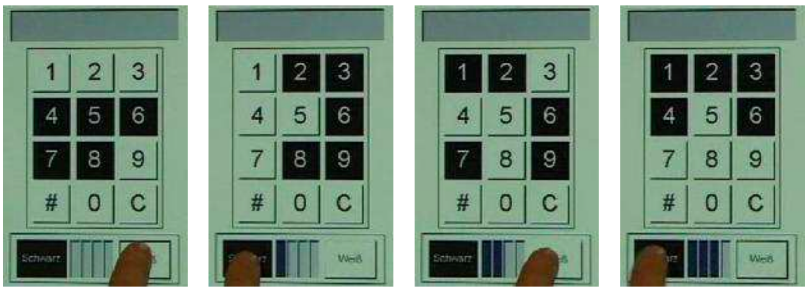

Figure 2.8: An example of an individual entering '3' in the cognitive trapdoor game [233, p. 238]

- **Indirect interaction**
  De Luca et al. argues indirect entry as method of shielding authentication secrets from onlookers during authentication [65]. De Luca et al. outlines a variation on the PIN that requires an individual to associate a colour with each digit. The user is then presented a keypad with coloured characters underneath each digit. The user enters the coloured character that corresponds to the digit and colour combinations of their authentication secret. Similarly, Gao et al. proposes an variant of Story, original outlined and discussed in §2.4.3, that relies on indirect entry [93]. The approach still requires an individual to select images of faces and object the user enters them by drawing over them in sequence. The user can indirectly enter the authentication secret by drawing over as many distractors as possible, as long as the user draws over the target images in sequence they can authenticate. Nonetheless, while users are able to use the authentication approaches, they still exhibit lengthy authentication times.

- **Searchmetric**
  Searchmetric mechanisms are another alternative [225]. Limited disclosure searchmetric mechanisms foil shoulder surfing and key-logging software, since they rely on the use of arrow keys or a mouse to manipulate sets of pictures. Most limited disclosure searchmetric mechanisms have some redundancy so that the observer is not able to deduce the key from casual observation but has either to observe a number of authentications or carry out an error-prone deduction of the key based on a few observations. The

v.Crypt system from Bharosa[1], illustrated in Figure 2.9, requires the user to use arrow keys to line up a shape on the bottom row with an alphanumeric key on the top row or to rotate a dial to line up letters in the same way as a combination lock is operated. This is done for as many letters and numbers as there are in the key.

Figure 2.9: Searchmetric graphical authentication approach.

The aforementioned approaches have several advantages, namely they can argue extend beyond a single device and are inexpensive to implement. Nevertheless, elaborate software obfuscation solutions typically result in lengthy authentication times and elaborate user interaction.

2.6.3 Comparison of Observation-types

There are advantages and disadvantages to both hardware and software obfuscation of authentication entry. Table 2.7 compares and contrasts some of the various obfuscation solutions, outlined and discussed in §2.6.1 and §2.6.2. The

		Approach	Authentication Times	Accuracy	Application	Evaluation
Hardware	Haptic	Undercover Phone lock	32s 28s	74% 90%	ATM Smartphone	Lab
	Tactile	VibraPass TAS	4s 35s	97% 85%	ATM Web	
	Gaze	Eye-password GPF GCP	11s 20s 54s	85% 60% 54%	Web ATM —	
Software	Cog.	PIN Game CHC Query	23s 72s 180s	91% 90% —	ATM Web —	
	Vis.	UYI	12s	100%	ATM	
	Indirect	ColorPIN CDS	14s 14s	77% 97%	ATM PDA	

Table 2.7: Comparison of observation resilient authentication mechanisms

[1]http://www.bharosa.com

aforementioned authentication solutions with observation resilience are almost exclusively designed for use on ATMs or mobile phones. Furthermore, authentication times appear to be comparatively lengthy when contrasted with those for passwords and PINs.

However, the evaluation of each of the observation resilient authentication mechanisms is controlled and performed in the laboratory, i.e. under optimal and ideal conditions. Consequently, the ecological validity of such assessments is questionable and performance is not necessarily representative of actual performance. Nevertheless, obfuscation clearly adds complexity that results in increased authentication times. The complexity costs appears both in hardware and software solutions.

Hardware approaches essentially focus on providing entry methods that others are unable to observe, as well as providing feedback that does not reveal entry of the authentication secret. The ideal hardware solution would be a one where the user simply thinks of the authentication secret as there would be no observable interactions [272]. Unfortunately, such a solution has yet to be developed and deployed. Nevertheless, there are other obfuscation strategies that monitor almost imperceptible actions, such as gaze. However, while such approaches are innovative there are still several questions surrounding eye-tracking. The widespread deployment of dedicated eye-tracking hardware is one limitation as is the concern of detecting a specific gaze in a room full of eyes. Beyond eye-tracking, other obfuscation strategies rely on pressure and vibration.

These strategies are potentially better suited to a remote control, as pressure can be monitored on button presses and vibration easily incorporated. Unfortunately, pressure values exhibited by an individual may not be particularly distinct or repeatable [197]. Furthermore, users appear uncomfortable in changing behaviour to utilise the power of pressure. Vibration, although easier to deploy, is undermined by the noise of the vibration motor [64].

Therefore, hardware-based obfuscation strategies appear unfinished and unpractical. The alternative is software. Software-based obfuscation strategies include indirect entry, visual trapdoors and cognitive trapdoors. Indirect entry is an interesting obfuscation approach being comparatively lightweight to implement. However, it is not clear how indirect entry would function within specific contexts, such as the living room. There is a limited number of buttons, so relying on a keyboard is not an option. Moreover, while indirect entry of characters with characters is doable [65] it may not be optimal with images. A graphical authentication approach would either need to offer mixed images that may interfere with each other, or use a combination of image and text that may not afford optimal performance.

Nevertheless, there are software obfuscation solutions, such as searchmetrics, that can potentially shield entry of purely graphical authentication secrets. There are numerous advantages to searchmetrics in that it not only provides observation resilience to onlookers but also has resistance to attacks such as key-logging. Moreover, an advantage searchmetrics possess over similar alternatives is the potential extensibility of the approach beyond specific scenarios and contexts. There is potential to develop a searchmetric approach that can be deployed across numerous devices, such as televisions and tablets, without dramatic redesign.

Furthermore, couple such advantages with the inherent benefits using software and searchmetric approach seems one the strongest obfuscation design directions. The primary benefit in using software-based approach is that the solution can be added to almost any device, past and present. Moreover, software solutions can be iterated and improved in place without having to request a user to purchase a new accessory or entirely new device. Consequently, a software-based searchmetric solution was selected as the basis of the alternative authentication mechanism, for the following reasons:

- Software-based searchmetric solutions can extend beyond a single device to others without dramatic redesign.
- Software-based searchmetric solutions can be added to existing devices without major expense.
- Software-based searchmetric can be iterated and improved upon in-place.

Therefore, the decision was taken to focus on a software-based searchmetric solution for obfuscation and develop a purely software-based authentication mechanism with observation-resilience.

2.7 Summary

The aforementioned sections, §2.3 to §2.6, outlined and discussed various aspects of several authentication mechanisms. The aim was to form the foundation of an alternative authentication mechanism that would act as alternative to the password in some key contexts.

2.7.1 Proposed Solution

The foundation of the alternative authentication mechanism can be formed from the conclusions of the aforementioned sections as outlined below:

1. The authentication solution is **knowledge-based** rather than token or biometric-based. Section 2.2 outlined and discussed the various options and concluded that knowledge-based solutions are relatively inexpensive and ensures users are actively involved in authentication.
2. The authentication solution is **graphical** rather than alphanumeric or relying on other senses, such as sound and touch. Section 2.3 outlined and discussed various options and concluded that images are more memorable than the words that name them and that modern consumer electronics are well positioned to support graphics.
3. The authentication solution is **recognition-based** rather than recall or cued-recall. Section 2.4 outlined and discussion various options and concluded that s solution based on recognition potentially minuses the use of coping strategies as the authentication secret is actually presented to the user.

4. The authentication solution relies on **face** images. Section 2.5 outlined and discussed various options and concluded that individuals readily retain and recognise face for long periods of time.

5. The authentication solution relies on **software observation resilience** to shield authentication entry. Section 2.6 outlined and discussed various options and concludes that software-based solutions can be deployed to devices past and present and iterated upon in-place.

Nevertheless, while the foundation of the alternative authentication solution was formed, the mechanism itself had to be designed and prototyped. The next chapter details the actualisation of the alternative authentication mechanism.

Chapter 3

Tetrad: An Alternative Authentication Mechanism

Graphical authentication secrets have the potential to replace passwords in some key contexts. The primary reason is that pictures are more memorable than the words that name them [188]. Furthermore, graphical authentication mechanisms can be designed in such a way that an individual merely has to recognise images to authenticate, rather than recall them [266]. Moreover, specific images, such as faces, can be retained and recognised for decades [13].

Consequently, the following chapter outlines the design and evaluation of a graphical authentication mechanism that relies on face images. The ensuing section, §3.1, details the design of the graphical authentication solution, focusing on aspects such as the authentication secret and presentation of images. The outlined aspects are used to actualise the graphical authentication solution, §3.2. Consequently, a prototype is produced and evaluated, §3.3 , with results reported, §3.4, then discussed, §3.5. Lastly, conclusions are drawn, §3.6, then future steps outlined.

3.1 Design

The previous chapter outlined and discussed various aspects of the alternative authentication mechanism, that essentially formed the foundation of the solution. The following aspects still had to be determined:

- **Authentication Secret**
  The actual authentication secret the mechanism will expect users to memorise.
- **Context**
  The envisioned context for prototype.
- **Presentation**
  The presentation or layout of the authentication solution.

The ensuing sections, §3.1.1 to §3.1.3, outline the aforementioned aspects of the alternative authentication solution.

3.1.1 Authentication secret

Section 2.7 detailed that the authentication mechanism relied on face images. Consequently, the authentication secret compromised of a series of face images. The next design decision was to determine the exact number of face images. The assumption is there is a limit to the number of faces an individual can successfully memorise. Miller argues individuals do have finite resources and the capacity of working memory is seven, plus or minus 2, objects or chunks [182]. The argument is that an individual can process a number of objects but as they increase, performance decreases.

Therefore, the limits of individuals had to be considered in the design of the authentication secret. The design had to ensure users could process and encode the authentication secret. John Shepherd-Barron respected such limitations when designing PINs for use in ATMs [21]. Shepherd-Barron originally favoured a six-digit authentication secret, as it equated the same length of his army number, that he could easily recall. However, when outlining the design with his spouse, they favoured a four-digit design. Consequently, Shepherd-Barron revised the design and created the popular four-digit PIN code. Interestingly, Cowan argues a similar figure for the capacity of working memory [57]. Cowan states that four objects or chunks is the capacity of working memory. Nevertheless, regardless of a specific figure on the capacity of working memory, the important aspect is that an individual has limited resources. Consequently, an authentication solution can not expect an individual to memorise an outlandishly large authentication secret.

Therefore, the design decision was taken to limit the authentication secret to **four face images**. The motivation was that individuals have limited resources [182, 57] and millions of them have exhibited their ability to memorise four-digit PIN codes.

3.1.2 Context

Authentication is most often achieved by means of a shared secret. By definition any disclosed secret is no longer secret and thus can no longer serve as an authenticator. Hence authentication context — the task, actors and physical environment within which the user is authenticating — must be taken into consideration when designing the authentication mechanism. There are essentially two environments that need to be considered:

- *Shared Space*
  We are not alone, and we are aware that individuals or devices could be potential threats. Tan et al. point out that large displays have a serious impact on privacy and argue that a solution to this should be sought [260].

- *Secluded Space*
  We are alone and do not need to be concerned that others are observing our actions. Most authentication mechanisms in use today *implicitly assume this context* by requiring the user to provide their entire secret. The only concession to possible observation is obfuscation of the entered text, and

even that has been abandoned on many consumer electronic devices, e.g. iOS briefly displays the entered text so as to enhance usability for the user.

The aforementioned contexts are not necessarily mutually exclusive and users will behave differently in each, no matter how unaware they are of security issues. Customers in the United Kingdom have to enter a PIN when using their credit cards to purchase goods in a card-present transaction. After a number of fraud cases banks are now advising customers routinely to shield their PIN entry. The fact that the banks had officially to issue such advice confirms that many people simply do not understand the security threats they are vulnerable to [288].

However, even the least privacy conscious user will not want private or sensitive information displayed on a large screen for everyone to see in a shared space. Unfortunately, many modern consumer electronics, such as Nintendo Wii and AppleTV, expect users to enter alphanumeric authentication secrets on large displays to play games and watch movies. The task of watching a movie on consumer electronics is unusual when considered. The task is an intensely shared experience only made possible through personal sacrifice. An individual needs to enter a password to confirm purchase of the latest movie or game.

Therefore, there is a need to accommodate the many people who want to be able to share movies, music and photos with others but do not want to give away their authentication secrets in the process. What is needed is a way for people to *prove* knowledge of a secret without revealing the secret to an active observer. The large display is not that of monitor but a television. Users relax with a remote control not a mouse.

Consequently, the decision was taken to outline a working context for the graphical authentication mechanism prototype. The **envisioned context** is a consumer, purchasing a movie through a television display, among friends and family.

3.1.3 Presentation

The last design decision before construction of the prototype was the layout or presentation of the authentication mechanism. Section 2.7 as well as §3.1.1 outline the fundamentals of the alternative authentication mechanism, namely:

1. *Recognition-based memory task*
   The authentication mechanism will be framed as a recognition memory task. The user will be presented elements of the authentication secret and requested to indicate if they recognise the elements.

2. *Face images*
   The authentication mechanism will rely on face images.

3. *Fixed-length for authentication secret*
   The authentication mechanism will use fixed-length authentication secret, limited to 4 elements.

The aforementioned aspects essentially produce a cognometric system [63], e.g. PassFaces. The PassFaces authentication approach is discussed in §2.4.3. Cognometric systems present a collection of images with a single stage, each collection

contain a single target image while the rest are distractor images. The user is required to specify the target image. The user continues completing stages until they have identified each element of the authentication secret.

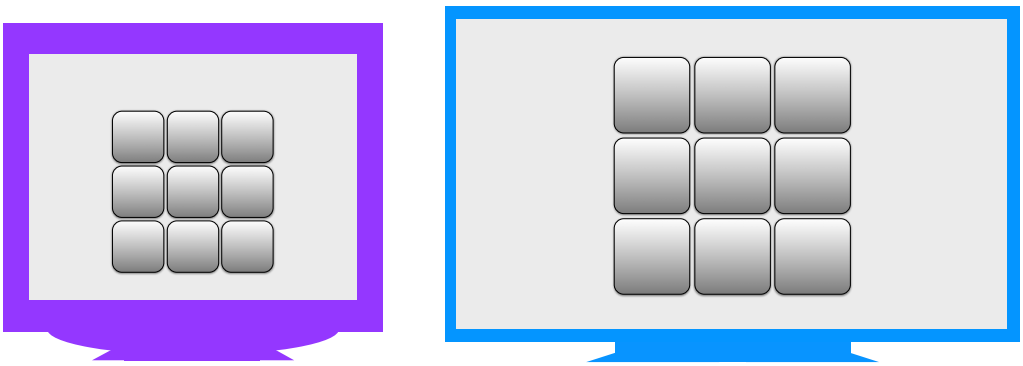

Figure 3.1: Cognometric graphical authentication approach displayed on a traditional computer monitor with an aspect ratio of 4:3 and modern television screen with aspect ratio of 16:9.

Unfortunately, PassFaces and other such cognometric authentication approaches are unsuitable for consumer devices in a shared space context. The approach is not suitable for the living-room, for the following reasons:

- *Limited resilience to observation*
  The user needs to indicate the target face, an interaction that could be observed by onlookers. The user can wield a keyboard to shield entry of the authentication secret. However, physical keyboards are not common in living-rooms, remote-controls are far more prevalent. Nevertheless, directly selecting an authentication element would reveal the target image to others.

- *Poor user experience*
  The authentication mechanism does not translate well from a monitor screen to a television screen. The aspect ratio, resolution, quality and interaction differences between traditional personal computers and televisions are dramatic. Figure 3.1 illustrates the differences between the two screens. The personal computer has excess chrome and additional on-screen controls, while the entire canvas is available on the television screen.

The initial concern can be addressed by redesigning the approach to provide resilience to observation. Section 2.6 outlined the various strategies to obfuscate entry of authentication secrets and concluded the alternative authentication mechanism should be designed as a searchmetric. Consequently, designed as a searchmetric, the alternative authentication solution would be resilient against observation.

Therefore, the next item to address was poor user experience. PassFaces was developed at the turn of the century for displays with a 4:3 aspect ratio. However, modern consumer displays and television screens have widely adopted the 16:9 aspect ratio, as consumers strongly prefer it [209]. Nevertheless, there is no real

evidence or psychological explanation that any aspect ratio is preferable or superior to another [244]. Nonetheless, consumer devices are increasingly adopting widescreen ratios. Consequently, the decision was taken to target such displays. The layout or presentation of the authentication mechanism would be designed as (1) a searchmetric and (2) target widescreen displays. Therefore, a single screen strategy was adopted for presentation, rather than presenting several screens as is the case with some cogonometrics, e.g. PassFaces. Consequently, users do not shift through five screens of nine image, instead they interact with a single screen containing all 45 images. The 45 images are randomly positioned in grid format. Contained within the set are the user's target images, i.e. those images which constitute the image-based authentication secret.

A successful authentication attempt requires the user to re-position columns and rows of images within the grid. The user does not move or select *individual* images; he or she re-positions subsets of the images, i.e. rows and columns. The goal is to align the target images either horizontally, vertically or diagonally. Because the user is moving rows or columns at a time, it is hard for an observer to see exactly which pictures are the focus of the movement. The introduction of an element of redundancy provides the obfuscation which protects the user.

Therefore, given the foundation and layout of the alternative authentication mechanism was determined the next step was to actualise the solution. Consequently, the next step was to construct prototypes of the alternative authentication solution.

3.2 Tetrad

An initial prototype of Tetrad was design for execution through a web browser, discussed in §3.2.1. A subsequent iteration was designed for a shared-space, discussed in §3.2.2.

3.2.1 Web Prototype

The initial Tetrad prototype was built for execution on a web browser, using Javascript to accomplish interaction. Buttons or 'arrow-keys' were positioned at each column and row edge, i.e. to the right of a row would be a 'right arrow-key'. The user would click the arrow-key which would, in turn, execute a Javascript to re-position images. Each click represents one movement in that direction, e.g. click a right arrow-key and all images within the respective row moves one position right within the grid with the rightmost image wrapping around to the left.

Informal testing revealed mixed reactions to the mechanism. This is not unexpected as experimental authentication mechanisms often evoke connotations of unusable and cryptic methods in the mind of a user.

Unfortunately, the early prototype only reinforced this perception.

There were almost as many 'arrow-keys' as there were images, resulting in a cluttered and confusing mess. The interaction was not intuitive, although a 'right arrow-key' may logically communicate the concept of re-positioning images

to the right, it failed visually to communicate this interaction. The re-positioning of images was not animated: they simply appeared in their new grid position, in the blink of an eye. Individuals would frequently click 'arrow-keys' to extract meaning, citing they only did so because they were the only objects on-screen that were not generic images.

Indeed, many individuals clicked target images directly, expecting a response from the system. This not only undermined the main purpose of Tetrad, i.e. resilience to casual observation, but highlighted the interaction flaws.

The interaction initially seemed simple and could scale to a range of *shared space* devices. However, the initial implementation had 28 'arrow-keys'. This vast number of buttons is difficult to navigate using the remote-control of a large-scale display.

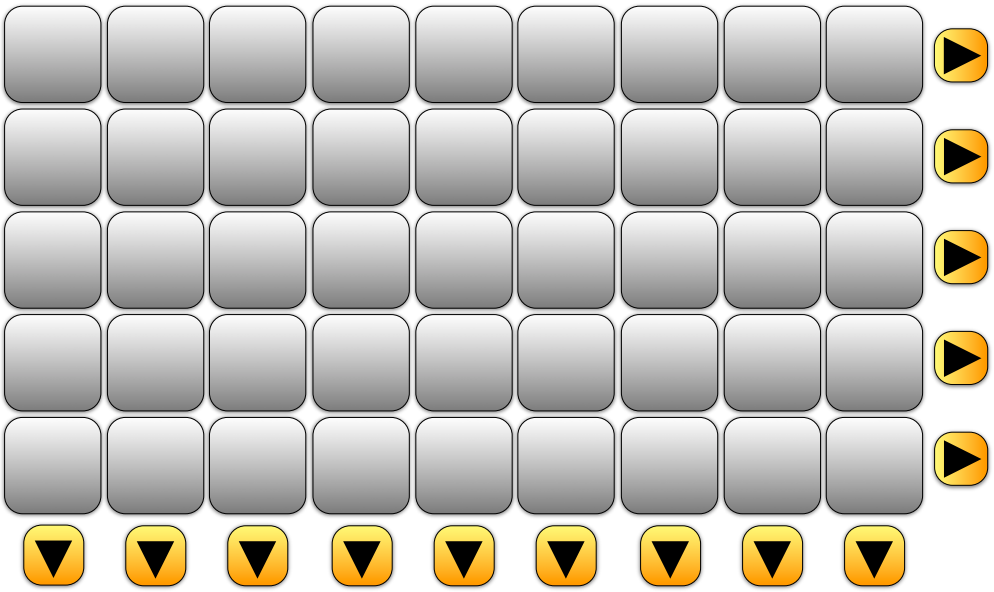

Figure 3.2: The design of an early prototype constructed for execution within a web-browser.

Figure 3.2 illustrates an iteration of the prototype exploring the use of 14 buttons that support circular movement in one direction only, deeming this acceptable since users often simply clicked the same 'arrow-key' to achieve movement, than both. However, 14 buttons is still excessive, and did not resolve any of the aforementioned concerns. Lastly, the generic images themselves also failed adequately to facilitate lightweight recognition [219].

3.2.2 Shared Space Prototype

In developing the *shared space* prototype a fresh perspective was taken: create a somewhat generic approach that could not only scale between shared space devices but also address the concerns of our earlier web prototype.

Furthermore, although images of faces had been determined as the authentication secret, the specific type of such images was not selected. Therefore, a

specific type of face images need to be selected that would facilitate lightweight recognition.

Design & Implementation

The shared space prototype was implemented using Objective-C for use with Apple's OS X. The operating system is utilised across their entire range of Apple devices, in one variant or another, i.e. iMac, MacBook, Apple TV, iPod touch, iPhone and iPad. Therefore, there was scope to explore future prototypes on different devices.

Three main concerns to be addressed in our shared space prototype were:

1. *Visual communication of image movement*
2. *Interaction required to re-position images*
3. *Exposure of target images*

The first concern was addressed by exploring how others had dealt with visually communicating re-positioning of content. Minimising a window is such an example. Historically, windows would simply disappear when minimised and appear in another location, e.g. the task bar. This was potentially confusing to new users as it was unclear where the window had disappeared to, and how to retrieve it. Apple's solution was to animate the window, shrinking it from its current location to its new location, i.e. the 'genie effect'.

Animation, often seen as a frill, serves the purpose in this case of visually communicating to the user the location of the now minimised window. A user consequently knows exactly where the window now resides.

The shared space prototype made extensive use of Apple's Core Animation development framework. The images are animated, moving rather than disappearing and reappearing. The intention behind the decision is to reduce the number of 'exploratory' button pushes by the user, often used to discover how the images moved within the grid.

The next concern was the interaction required to re-position images. Even 14 buttons was still unrealistic. Television manufacturers had tackled the problem of navigating complex programming guides for hundreds of channels without using so many buttons. The solution in almost all cases was a directional-pad complimented with a 'Select' or 'OK' button on the television remote control.

Tetrad's interaction was revisited and efforts were made to map the function of 14 buttons to just five. Our solution was a horizontal and vertical 'selection bar'. Users navigate the grid of images using these bars, rather than a free flowing cursor. The horizontal bar would map to the 'up' and 'down' buttons of the directional-pad while the vertical bar would map to the 'left' and 'right'. Obviously, both bars cannot be on-screen at the same time, so they fade-in and -out in response to directional-pad movements, e.g. while the horizontal bar is on-screen the user can move it up and down — if a user presses 'right' the horizontal bar fades out and the vertical bar fades in. Naturally, the navigation bars remember their position when fading-in and -out, aiding entry.

Once an individual has navigated to the column or row they wish to manipulate, they press the selection button. If a row is highlighted, and the select button is pressed, a circular movement to the right occurs, e.g. all images move right one space, with the last image becoming the first. Similarly, if a column is selected and the selection button is pressed, a circular movement downwards occurs, each image moving down one position, with the last image moving to the top of the column.

The approach also addressed the remaining concern: exposure of target images, as users control only the two selection bars. This means they are able to highlight individual rows and columns but *not* individual images — preventing inadvertent disclosure.

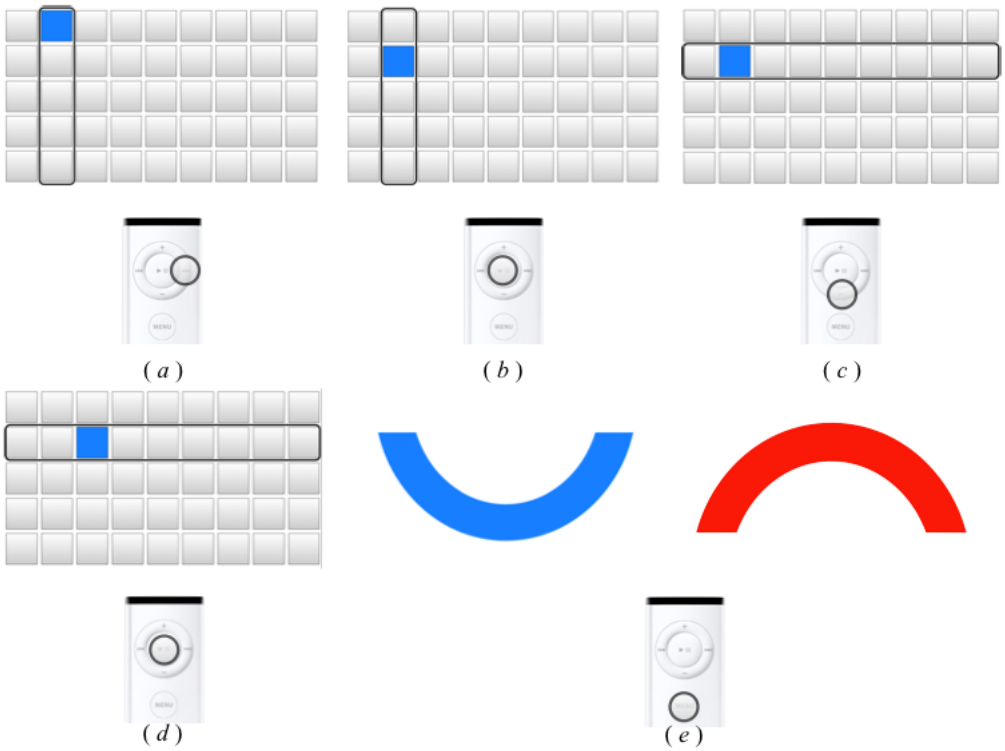

Figure 3.3: Interaction within the shared space prototype

Figure 3.3 illustrates interaction within the shared space prototype. Let us assume that the highlighted square represents an image the user wants to reposition, its intended location being one column along and one row down. The following steps are required:

(a) *Select column* - Using the 'right' navigation button, the user moves the vertical selection bar to highlight the second column.

(b) *Move image downwards* - Using the selection button, the user activates the highlighted column, moving all images within the column downwards, with the last image becoming the first.

(c) *Select row* - Using the 'down' navigation button, the user moves the horizontal selection bar downwards. The vertical selection bar fades out.

(d) *Move image right* - Using the selection button, the user activates the highlighted row, moving all images within the row to the right, with the last image becoming the first.

(e) *Submission* - Using the submission button, i.e. 'Menu' on the Apple remote, the user can submit their efforts for assessment.

   i. *Success* - if successful, a 'smile' is displayed on-screen indicating that access has been granted to the service or system.

   ii. *Failure* - otherwise, a 'frown' is displayed on-screen indicating that access has been denied and that another attempt can be made.

   These simplified facial gestures are generated and animated using the images within Tetrad. The images are repositioned and filled blue for a successful entry, red if otherwise. This approach extends the accessibility and simplicity of feedback while avoiding language.

The sequence of steps represents merely one way of repositioning an image; several other paths could be utilised.

Indeed, such redundancy has the potential to offer flexibility to the user, who, if feeling under threat, could take less obvious routes to reposition images. Furthermore, individuals could perform 'trick-moves', repositioning images not required for authentication to confuse onlookers.

Image Type

The experiment was not intended to test memorability of different kinds of images. Therefore, the most memorable image type was selected so that any observed effects would be easier to attribute to the nature of the mechanism than to the efficacy of the image type used.

Furthermore, sourcing a selection of images can be difficult due to lighting, quality and copyright reasons [76]. Therefore, a collection of celebrity images were generated for use with the prototype mechanism. Figure 3.4 illustrates the shared space prototype with celebrity images.

3.3 Evaluation

The evaluation of Tetrad need to test two aspects: how easy Tetrad was to use, and how easy it was for observers to identify the secret images if they watched someone else authenticating using Tetrad. The first aspect assesses the usability and the second the security of the mechanism. Therefore, participants were asked to engage in three tasks, using a within-subject design.

The three tasks are outlined below:

Figure 3.4: Shared space prototype.

1. *Authenticating with Tetrad.*
   Although participants were familiar with alphanumeric authentication mechanisms and their mechanics and processes, it's unlikely they would be familiar with image-based authentication. Therefore, the first task asked participants to authenticate using Tetrad. This task assessed the usability of the mechanism and also prepared participants for the second task. The cognitive workload of the authentication task was also estimated.

2. *Observing Authentication*
   This task asked the participant to determine the secret key being entered by another user. The participant viewed two videos, of equal length, one showing an unknown individual making an alphanumeric authentication attempt with an on-screen keyboard and the other showing the same individual authenticating using Tetrad. The *independent variable* is the authentication mechanism while the *dependent variable* is the success or failure of the participant determining the password entered. The *experimental hypothesis* is that Tetrad will be more resistant to casual observation than alphanumeric authentication.

3. *Questionnaire*
   The last task asked participants to provide additional information based on participants' thoughts and concerns regarding authentication in shared spaces.

3.3.1 Subjects

Eleven participants were recruited: 6 females and 5 males. Their ages ranged from 20 to 70 and included various backgrounds and professions, e.g. student,

retired, professional etc. The diversity as well as the low number of participants may appear less than ideal but the assumption was that the participant group would be enough to uncover any major usability concerns.

3.3.2 Apparatus & Materials

The system used was an Apple MacBook, Model: MB062LL/A, with 2GB RAM. The MacBook's accompanying Apple Remote was used for interaction.

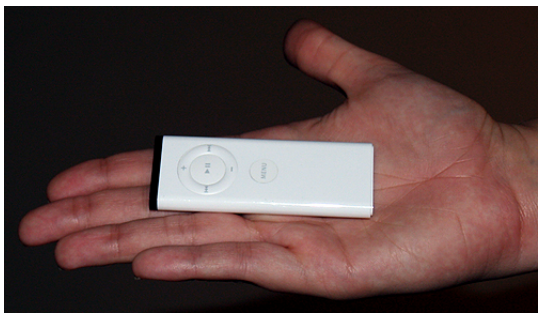

Figure 3.5: Early Apple Remote

Tetrad required two image sets, one for the first task, and one for the second task. A total of 90 face images, 45 for each set, were extracted from the University of Massachusetts LFW database[1].

The videos used in Task 2 were captured using Screenium 1.0 in advanced using our MacBook. The first video required the Nintendo Wii to be connected to our MacBook using Elegato EyeTV Hybrid. The output from the Nintendo Wii was viewed using Elegato EyeTV 3. The captured videos were played full-screen during the trial using QuickTime 7 Pro.

Finally, participants were provided with pens and a handout to complete which included instructions for each task, questions regarding the experiment, cognitive workload assessments and a brief one-page questionnaire.

3.3.3 Procedure

Participants were requested first to read the cover-page of our handout, which outlined the nature of the trial, estimated time to complete, three tasks that participants were expected to complete and our contact details should they have any queries. Lastly, participants indicated consent by signature before the experiment commenced.

Task 1 introduced our image-based authentication mechanism, Tetrad, to the participants and explained the concepts necessary to make a successful authentication attempt. Four images, which represented an image-based password, were printed as part of the instructions. Participants were advised there was no time-limit and that they did not need to memorise any of the images.

[1] http://vis-www.cs.umass.edu/lfw/

Upon completion of an authentication attempt, participants were requested to complete two evaluation procedures which examined cognitive workload. The NASA-Task Load Index or NASA-TLX was used to determine the perceived cognitive workload. The NASA-TLX is an evaluation tool developed at NASA's Ames Research centre and is used to assess the performance of individuals and systems. The process determines perceived cognitive workload using six sub-scales, namely:

- *Mental Demand (MD)*
  The perceived level of mental processing required to perform the task, i.e. users may perceive task as mentally demanding or relatively straightforward.

- *Physical Demand (PD)*
  The perceived level of physical activity required to perform the task, i.e. users may perceive task as physically demanding or requiring little exertion.

- *Temporal Demand (TD)*
  The perceived pace required to perform the task, i.e. users may perceive the pace as overly rapid or slow.

- *Effort (E)*
  The overall level of mental and physical activity to perform the task.

- *Performance (P)*
  The overall assessment of personal performance on the task.

- *Frustration (F)*
  The overall level of stress, anxiety and irritation experienced during the task.

Individuals are expected to produce ratings across the aforementioned scales. Consequently, participants in the evaluation first completed weighting and magnitude ratings for each sub-scale.

Task 2 instructed participants to watch two videos, of equal length. The purpose of viewing the videos was to extract the password entered by an unknown individual. In the case of alphanumeric authentication individuals wrote down the characters in their recalled position, e.g. if the password entered was 'east', a response of 'seat' would result in all characters being correctly identified but only one with the correct position, 't'. Similarly, for the image-based password, participants were requested to select four images from the image-set printed in the handout, as well as identifying that image's position within the password. In both cases, participants were asked to rate their confidence on a scale of 0 to 100, i.e. how confident they felt about their estimations.

The second video had an additional question, which was for the participant to guess the alignment and position of the secret set of images within the image set when the person had completed moving all images around to authenticate. Participants indicated this on a generic template of Tetrad's layout and rated their confidence in their estimation.

Upon completion of Task 2, participants were requested to complete two evaluation procedures which examined cognitive workload for extracting the image-based password. NASA-TLX was used, thus participants first completed weighting then magnitude ratings for each sub-scale.

Lastly, participants were invited to complete a short questionnaire, i.e. Task 3.

3.4 Results

3.4.1 Task 1

Task 1 was completed by all 11 participants, with every attempt being successful. Although time and memorability were not a consideration during this experiment, anecdotal evidence suggests that faces were memorable. Furthermore, it took time to authenticate varied between participants. The evaluation procedure for the first task was completed by all 11 participants, which generated a cognitive workload score. Table 5.4.1 shows the mean, median, minimum and maximum weighted ratings for Task 1.

	Workload Score
Mean	61
Median	60
Min	15
Max	91

Table 3.1: Workload Score for Task 1

The mean weighted rating or workload score for Task 1 is 61. The factors and their respective weights which contribute to the workload score can be seen in Figure 3.6. The factor contributing the most to the workload score is *Effort* with an approximate mean weight of 4. The median, minimum and maximum for *Effort* is 4, 2 and 5 respectively. While the factor contributing the least to the workload score is *Physical Demand* with an approximate mean weight of 1. The median, minimum and maximum for *Physical Demand* is 1, 0 and 4 respectively.

The factor with the highest rating is *Temporal Demand* at approximately 65. The median, minimum and maximum for *Temporal Demand* is 70, 20 and 100 respectively. The lowest rating is *Physical Demand* at approximately 15. The median, minimum and maximum for *Physical Demand* is 10, 0 and 50, respectively.

3.4.2 Task 2

All 11 participants attempted Task 2. After watching the first video all participants successfully extracted the characters within the alphanumeric password and their positions. The mean confidence rating was 95 out of 100, with 75 being the minimum and 100 the maximum confidence rating.

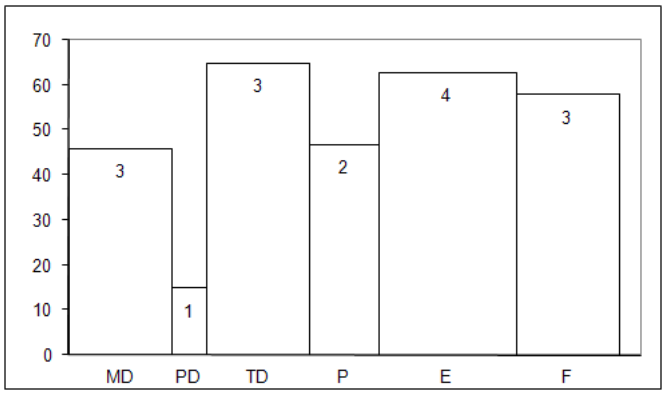

Figure 3.6: Mean Weighted Ratings for Contributing Factors for Task 1

However, after watching the second video, all participants failed to extract any of the images contained within the image-based password. Furthermore, 45% of participants identified at least 4 incorrect images, the mean being 2 images, with 27% of participants not identifying any images. If these participants are removed, the mean increases to approximately 3 images.

Participants did identify similar images, with two images in particular being identified by 45% and 27% of participants, respectively. If those individuals who made no attempt to identify any images, are removed, these values increases to 62.5% and 37.5% respectively.

The mean confidence rating from participants, regarding their identification of images, was approximately 27 out of 100. One of the participants identified only one image (incorrectly) but their confidence rating was 100, confident that above all else the single image they had identified was part of the image-based password. If this outlier is removed, the mean confidence rating drops to approximately 16 out of 100.

Participants were asked an additional question for the second video, which was the alignment of the image-based password. The alignment used within the video was diagonal but none of the participants identified this alignment, 27.2% could not identify the alignment, 27.2% identified horizontal as the alignment while the majority of participants, 45.4%, identified vertical as the alignment.

Participants were asked to rate their confidence in the alignment they had identified, on a scale of 0 to 100, the mean confidence rating was approximately 28.

The participants were asked to complete an evaluation procedure, which assessed workload, for the second video. Table 3.2 shows the mean, median, minimum and maximum weighted rating for the second video.

The mean weighted rating or workload score for identifying the password is 74. The factors and their respective weights which contribute to the workload score can be seen in Figure 3.7. The factors contributing the most to the workload score are *Mental Demand* and *Effort*, with an approximate mean weight of 4.

The factor contributing the least to the workload score is *Physical Demand*, with an approximate mean weight of 0.

	Workload Score
Mean	74
Median	76
Min	23
Max	98

Table 3.2: Workload Score for Task 2, Video 2

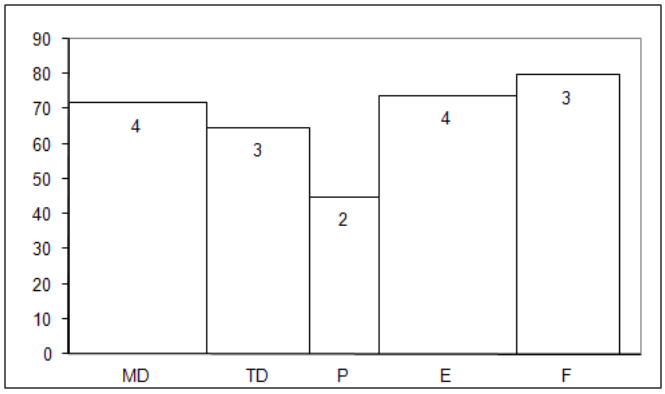

Figure 3.7: Mean Weighted Ratings for Contributing Factors for Task 2

The factor with the highest rating is *Frustration* at approximately 80. The median, minimum and maximum for *Frustration* is 90, 10 and 100 respectively. The factor with the lowest rating is *Physical Demand* at approximately 10. The median, minimum and maximum for *Physical Demand* is 5, 0 and 45 respectively.

3.4.3 Task 3

Lastly, the answers to the questionnaire reveal that 54.5% of participants have purchased on-demand content through their television using their remote control. When asked if they would authenticate when not alone, 90.9% of participants said they would authenticate in front of others, with 70% ranking family as the least threatening and strangers the most.

3.5 Discussion

The evaluation was designed to assess the usability and security of the shared space prototype. The first task assessed the usability of the mechanism. The participants clearly put some effort into authenticating with Tetrad, but at least did not find it physically demanding. Fewer than half indicated that it was mentally demanding. This suggests that the workload score is less than than optimal, and could be improved. However, all participants managed to authenticate successfully in what was their first use of Tetrad, which is encouraging.

Task 2 assessed the observability of Tetrad. The participants were asked to attempt to record the authentication secret after watching someone enter either

their alphanumeric password or their image-based password. It was no suprise that they all correctly observed the alphanumeric password. However, we did not expect that no one would be able to pick out at least one of the images involved in the secret images used in the image-based password. In terms of vulnerability to observation it would appear that Tetrad is as strong as we had hoped.

However, when considering that several individuals mistakenly identified the same faces, it could be that an individuals choice was influenced by attractiveness or race [62]. Thus, Tetrad's interaction redundancy, assumed to increase security, could itself prove *redundant* due to image-type and/or secret-creation. This could be tackled in numerous ways. Whether any such approaches could curb the inherent problems in using faces for authentication is another question.

It is interesting to note that the workload score for users attempting to uncover the image-based password was higher than that of the workload score for Task 1, indicating that authenticating with Tetrad requires less effort than observing someone else authenticating with Tetrad with a view to extracting an authentication secret. However, further investigation will be required to determine strength outside the realm of shared space.

Furthermore, the extra effort perceived by our participants needs to be addressed if the authentication mechanism is to succeed. Even though people complain about passwords, the undeniable fact is that they are very convenient when authentication is required [183] and people will always minimise their cognitive effort if at all possible [82, 194].

In strengthening the appeal and credibility of Tetrad, it needs to be compared and contrasted to competing graphical authentication mechanisms. This will require the finalisation of the prototype, carefully considering and outlining the procedures for authentication secret-creation and identification. Moreover, the services and systems suitable to Tetrad need to be contemplated and defined as Tetrad is *not* a one-size-fits-all authentication solution. The resulting product can be evaluated using traditional metrics, such as login-time and memorability using longitudinal usability assessments.

It is well established that people are the weak link when it comes to security [1, 238] They make clear judgements about costs and benefits. If the cost of authenticating securely is balanced against their risk perception, even if it is inaccurately low, and they might well prefer not to use a mechanism such as Tetrad but rather to accept the risk of traditional mechanisms [288].

However, the idea of buying online content, using a device such as a Wii, iPhone or Apple TV, is relatively novel. Perhaps, as people start using these devices in shared spaces, the issues envisaged will come to the fore and companies will start looking for mechanisms akin to Tetrad to mitigate the threats of shared space authentication.

Tetrad's first evaluation was promising but it is a first step in the journey towards creating an acceptable, secure authentication mechanism for shared spaces, which will be as convenient as possible.

3.6 Conclusion

The chapter introduced an alternative authentication mechanism called Tetrad. The authentication solution was designed to allow consumers to authenticate among others in a shared space. Nevertheless, while the initial prototype of Tetrad showed promise, the evaluation had weak ecological validity.

- The evaluation had weak ecological validity. Tetrad was assessed in a vacuum and was not assessed as part of task or deployed to actual users.
- An envisioned context was outlined for the design and evaluation of Tetrad, i.e. the living room. However, tasks cross context, meaning an authentication mechanism may need to cross contexts as well. An individual may complete a task one day in the living room using a television, but the next on a train using a tablet. Consequently,

The authentication solution was assessed in a vacuum. Tetrad was not part of a task, it did not have an actual registration process. Furthermore, it was not clear how the process would be positioned and deployed alongside other authentication solutions. The reality is that a task is not restricted to a single context but can occur in many. An individual may complete a task one day in the living room using a television, but the next on a train using a tablet. The scenario is different, so too is the device individuals use to perform the task.

Consequently, an authentication mechanism may be expected to cross contexts as a task cross contexts. Therefore, the next step was to determine the correct task for Tetrad. The task not only had to be a realistic use of Tetrad, as to ensure strong ecological validity, but also had to be a task that could be realistically deployed to individuals. A task involving precious and private medical data coupled with an experimental authentication solution would be unacceptable, for obvious reasons.

Therefore, the next chapter outlines and details several potential tasks before determining one for use in evaluations. Furthermore, the next chapter also outlines the research questions that will be explored in future evaluations, using the selected task.

Chapter 4

Evaluation Task

Tetrad is a graphical authentication mechanism that has the potential to replace passwords in the living room [224]. Nevertheless, while Tetrad showed promise, the evaluation, outlined and discussed in §3.3 to §3.5, clearly had weak ecological validity. A situation that is not unusual for experimental authentication mechanisms. Table 2.7 illustrates the reality that few alternative authentication mechanisms are evaluated beyond the laboratory, under realistic conditions. Brunswik emphasised the need to focus on tasks and scenarios to ensure evaluations and experiments were applicable to the world around us [42, 54]. Similarly, Sasse argues systems employing security mechanisms, such as an authentication, should be perceived as socio-technical systems [238].

Consequently, the following chapter outlines the design of an evaluation task for Tetrad that has strong ecological validity. The ensuing section details the requirements of the envisioned task, §4.1, as well detailing the basic structure of it, §4.2. The conclusion is that any task involving authentication comprises of a primary task and secondary authentication task. Therefore, several potential primary tasks are considered, §4.2.1, before outlining the fundamental aspects of the secondary authentication task, §4.2.2. The potential tasks are then assessed for risk, §4.4, before selecting one. Lastly, an application is proposed and research questions outlined for subsequent evaluations of Tetrad, §4.5.

4.1 Requirements

An evaluation with ecological validity required Tetrad to be incorporated into an actual task. The task could be anything, ranging from something simple such as browsing articles in an RSS reader to adding an authentication secret to a password manager. There were two requirements for the task, as follows:

1. *The task had to represent a realistic use of Tetrad*
   The task had to illustrate a real-world use of Tetrad, authentication as a secondary task rather than the primary one. If the authentication mechanism was bundled with an unrealistic application, one never envisioned for use with Tetrad, then any insights extracted would have little validity.

2. *The task had to be suitable for evaluation purposes.*

Individuals using the application should not be placed in jeopardy. The application or experimental mechanism could fail at any minute. Therefore, failure should not result in the user being put at risk.

These aims were set to ensure the resulting package was a realistic implementation of Tetrad. There was no merit in coupling Tetrad with an application that either did not normally use an authentication mechanism *or* would never use a graphical one. Furthermore, although the application had to be realistic, we had to be mindful that it was still experimental. The application or mechanism could contain any number of flaws or errors. Therefore, users of the application could not be placed at risk, if failure occurred. These were the only two requirements in the design of the application.

4.2 Structure

The inclusion of an authentication mechanism within a task must be carefully considered. The task created at the end of the process would be actualised as an application. Consequently, if Tetrad was added without proper consideration of the impact on the workflow, it could cause friction. A conflict that will result in users either abandoning the application or bypassing the authentication mechanism [24].

The importance of such a conflict is often not considered in the evaluations of experimental authentication mechanisms. The reality is that many evaluations often frame an experimental authentication mechanism as the *primary* task rather than a *secondary* task [23, 239]. Authentication is not a primary task, authentication is always a secondary task. The user has a specific workflow or goal and an authentication mechanism is one step in completing that workflow or achieving that goal. Users may act calm and collected in a sterile lab when nothing is at stake and are rewarded with money or course credit. However, when users are distracted, stressed and frustrated, when they are attempting to access important course work, an experimental authentication mechanism may elicit different reactions.

Consequently, a task involving authentication, comprises essentially of two tasks: namely a primary task and a secondary authentication task.

4.2.1 Primary Task

There are several tasks that incorporate authentication, e.g. purchasing a film, reading email etc. The primary task is the focus or goal of the user. The authentication step within a task essentially represents a secondary authentication task, i.e. individuals are required to authenticate in order to purchase a film.

Consequently, Tetrad could not simply be used with any task. The evaluation task had to require an authentication step otherwise it would undermine the ecological validity of any assessment, as the task would be unrealistic. Moreover, the authentication step had to be realistically fulfilled by a graphical authentication solution. The latter point is salient as there are several potential tasks that may require authentication but are not best served by a graphical authentication

solution. Consider the scenario where an individual needs to purchase additional mapping information while driving around a foreign city. The task may require authentication but graphical authentication would not be optimal in such a context. An individual have to stop the vehicle as well as view and select images. An optimal solution would be an audio or voice password, an authentication solution that an individual could perform hands-free.

Therefore, context aspects of a task need to be considered when coupling an authentication mechanism with it. There are several aspects to context but three important aspects are:

- *User*
  The actors within a task include the user and any individuals supporting them, such as an operator or technician. While there is no single type of computer user, they can be broadly divided into two groups: professionals and consumers. The distinction serves to emphasis the support that surrounds the user, rather than any technical expertise they may posses. A professional will receive training and access to real-time technical support. This is not typically true of consumers, who are more self-supporting.

- *Device*
  Tasks may or may not be tailored for specific devices or platforms. A task be more optimal on a dedicated portable device such as smartphone or it may be better suited to a device with a very large screen, such as desktop or laptop computer.

- *Environment*
  The environment the task is performed in is another important consideration. If the environment is heavily populated then any information used within the task may be observed by others. Furthermore, a bustling environment may make it difficult for an individual to focus all concentration on the task at hand.

The aforementioned aspects of task context closely mirror many of the aspects of authentication context, outlined in Table 2.4. Consequently, if both contexts do not align along these aspects, it would suggest the authentication mechanism is not optimal for the specific task.

Nevertheless, even if the contexts of both the task and authentication mechanism do align, the authentication solution must still be integrated into the task. Therefore, aspects of the secondary authentication mechanism must also be considered to ensure an authentication mechanism is optimal.

4.2.2 Secondary Authentication Task

Section 2.1.2 outlines the two main purposes for including an authentication step within a task: *access control* and *accountability* [26]. However, while these may be valid reasons for including an authentication mechanism within a task, there will also be specific motives for considering the use of authentication in the first instance. Understanding the motives for including authentication, affords insight

into specific characteristics an authentication mechanism should exhibit to fulfil the requirements of the authentication step.

Furthermore, asides for the motivates for including an secondary authentication task there is also the consideration of different contexts. The reality is that many tasks cross contexts, individuals no longer perform all computer tasks on a single device. An individual, for example, may purchase a film on their laptop, phone and television. Consequently, as tasks cross contexts so does authentication. However, an authentication mechanism may not easily cross context. The use of passwords on televisions in one such example. Apple has been used to selling individual content on computers but when transitioning the task to a television, passwords persist — despite the obvious concerns of other observing password entry.

Therefore, when determining the optimal evaluation task to couple with Tetrad, the following aspects had to be considered, namely:

- *Motivation*
  The reasons for including an authentication mechanism within the task.

- *Cross Context*
  The strategy used to transition the authentication mechanism across contexts.

The aforementioned aspects of the secondary authentication mechanism had to be considered to ensure Tetrad was properly evaluated.

Motivation

There are various different motives for including an authentication mechanism within a task. If the motivation is to comply with laws of the land, an authentication may need to have an extensive deployment history, experimental authentication mechanisms are rarely deployed, making them unsuitable for such motivations. Nevertheless, there are many motivations for including an authentication mechanism within a task, three potential motivations are:

- *Rules and Regulations*
  Designed to benefit most people within a region, rules and regulations ensure individuals and organisations do not put citizens at risk. The United Kingdom, for example, proposes the use of authentication in regulations regarding the management of personal information [122] and the protection of minors from adult certified content [193].

- *Content Control*
  Rampant piracy of digital music, early in the century revealed the market potential for digital content [17]. Consequently, content creators made digital content available to retailers, under the condition they can control the flow and use of it [111]. There are several organisations that rely on authentication, among other tools, to secure and control the flow of digital content.

- *Regulating Access*
  There are several scenarios where authentication can be used to regulate access. An authentication mechanism can regulate access to sensitive personal information on a device, e.g. smartphone lock, or within an application, e.g. location tracker. Moreover, an authentication solution, especially alphanumeric authentication, is a cheap and effective way to manage users, even if authentication is not necessary for any specific security reason.

The aformentioned motives are among some the most common reasons for including authentication within a task. The motivation of including an authentication mechanism within a task must be considered to ensure the optimal solution is selected.

Crossing Contexts

The reality is that tasks cross contexts, meaning any included authentication step must also cross contexts. Consequently, there are several strategies to expanding an authentication mechanism across contexts, they are:

1. *Distinct authentication mechanism for each device.*
   The user may use an alphanumeric authentication approach when purchasing a movie on a personal computer and use Tetrad when purchasing a movie on a television. However, while such a strategy may be acceptable for a select few authentication tasks it seems unrealistic for most tasks. The user would be required to manage several authentication secrets for a few simple tasks.

2. *Distinct interaction for each device.*
   The task could employ the same authentication secret for all devices but require individuals to perform a different authentication process for each device. Therefore, while an individual may be required to perform a cognitive trapdoor process on the television, see §2.6, they may be required to simply enter the authentication secret on a mobile phone. Banks are yet another example of an organisation using such a strategy. The bank may require an individual to enter elements on an authentication secret on touch-type phone but enter the entire authentication secret to enter a desktop application.

3. *Same Mechanism and Interaction for all devices.*
   The remaining strategy is to utilise the same authentication mechanism and authentication secret across all devices. This is the strategy adopted for most consumer applications. Apple has used the alphanumeric authentication secret used on the desktop version of iTunes and pushed across all versions of iTunes on the television, tablets and smartphones.

The last strategy is the one adopted by industry when constructing consumer applications. However, while an authentication may need to cross contexts it does not need to reach for the lowest common denominator or fit every task. The

authentication mechanism should compliment the optimal context and be functional in a range of others. Otherwise an experimental authentication mechanism will not be adopted.

The optimal experience for watching a film is on a large screen, with friends and family, in the living-room. The authentication mechanism should address the key concerns in that context. Users wanting to purchase movies on a smartphone or tablet can expect the same authentication mechanism to be functional but not necessarily optimal. The assumption is that users will be authenticating more on the optimal device and rarely on suboptimal devices.

Consequently, when considering candidate tasks the *optimal device* had to be identified as it would inform the optimal authentication solution.

4.3 Candidate Tasks

There was a total of seven potential primary tasks considered as the foundation of an application. The primary tasks considered were:

1. *Edit personal information.*
   An individual accesses an application and updates personal information associated with them. The optimal device for such a task would be one centred around a good display and keyboard, namely a desktop or laptop.

2. *Access content from a digital store.*
   An individual accesses an application containing a library of digital content and selects items for download. The optimal device for such a task would depend on the content itself, e.g. accessing a movie on a television.

3. *Make a phone-call.*
   An individual accesses an application to make a voice call to another individual. The optimal device for such a task would be a device able to place calls, namely a smartphone.

4. *View a friend's location.*
   An individual accesses an application or service to view an associate's last known location. The optimal device for a such task would be one with a good display, meaning a range of devices.

5. *Read an article.*
   An individual accesses an application or service to peruse articles they have collated from various sources. The optimal device for such a task would be one with a very high-quality display with enough space to display pages of text, such as a tablet.

6. *Personalise a service.*
   An individual accesses a service that has been personalised to their taste based on information collected about them. The optimal device would depend on the service and could be practically any device.

7. *Access a password manager.*
   An individual access an application that contains all of their passwords, when they need to access another proceed system or service. The optimal device would be one to hand, most likely a smartwatch or smartphone.

The next step in selecting a task was determining the risk associated with each task. Therefore, a risk evaluation was performed on each task.

4.4 Risk Evaluation

The candidate tasks, discussed and outlined in §4.3, all carry risk when deployed to an actual user-base. However, the risk posed by each workflow is not necessarily the same. The workflow deployed to the user-base must be realistic but not place users at risk. Therefore, the risk-level of each task is considered. The risk level associated with each workflow will be a deciding factor in what workflow to use. Consequently, a risk evaluation was performed for each of the potential workflows.

4.4.1 Risk Evaluation Process

The level of risk associated with each workflow is determined using steps outlined by the National Institute of Standards and Technology (NIST) [259]. NIST states risk evaluation is part of the Systems Development Life Cycle (SLDC). The risk evaluation process used seven of the steps recommended by NIST, they are as follows:

1. *System Characterisation*
   Outline the nature of the system, i.e. hardware and software, as well as the interconnected systems.

2. *Threat Identification*
   Determine the threats associated with the system, i.e. natural, human and environmental.

3. *Vulnerability Identification*
   Determine any potential vulnerabilities associated with the system, e.g. known software exploits.

4. *Control Analysis*
   The measures in place to reduce any potential threats, as well as the measures in place if any vulnerabilities are exploited.

5. *Likelihood Determination*
   Generate an overall likelihood rating of vulnerabilities being exploited, considering control measures in place.

6. *Impact Analysis*
   The overall impact from vulnerabilities being exploited, along three dimensions: integrity, availability and confidentiality.

7. *Risk Determination*
   Determine risk level through consideration of *impact analysis*, *likelihood determination* and *control analysis*.

Likelihood	**Impact magnitude**				
	Insignificant	**Minor**	**Moderate**	**Major**	**Catastrophic**
Certain	Medium	High	High	Extreme	Extreme
Likely	Low	Medium	High	High	Extreme
Possible	Low	Low	Medium	High	High
Unlikely	Negligible	Low	Low	Medium	High
Rare	Negligible	Negligible	Low	Low	Medium

Table 4.1: Josang et al. NIST-based Risk Matrix [141, p.271]

The risk associated with each workflow will be the outcome of step seven. The risk level will be determined using the *likelihood* against the *impact*. Table 4.1 will be used to determine the risk level. The outcomes for the first three steps for each workflow are the same. They are, as follows:

1. *System Characterisation*
   The system characterisations for each authentication task are the same. The task will need to operate across an ecosystem of devices but will be optimal on at least one. The envisioned ecosystem comprises of a desktop, laptop, tablet, television and smartphone.

2. *Threat Identification*
   The primary threat is an individual, who is not the user, is able to complete the authentication task.

3. *Vulnerability Identification*
   Flechais *et al.* [83] argues there two types of vulnerabilities, technical vulnerabilities and social vulnerabilities. The primary vulnerability in all the authentication tasks is a technical one, i.e. the experimental authentication mechanism. The mechanism is a vulnerability as it could contain any number of unknown usability errors and flaws, that could be exploited.

The remaining steps will be evaluated for each of the potential workflows. The potential workflows and risk evaluation are detailed as follows.

4.4.2 Risk Evaluation of Potential Tasks

Edit personal information

The primary task is editing personal information. The secondary authentication task will positioned before an individual is able to edit information. The motivation for including a secondary authentication task is rules and regulations. The steps 4 to 7 of the risk evaluation are detailed, as follows:

4. *Control Analysis*
   A *technical control method* can be put in place to ensure the application can be remotely deactivated if a breach occurs. This would rely on users informing us of a data breach. Therefore, a *nontechnical control method* will also be used where any changes to personal information will result in a confirmation email. The email will be sent to the user's previously specified email address and current email address. The email will instruct users to contact us, if they have not authorised any changes to personal information.

5. *Likelihood Determination*
   Personal information is valuable, therefore attackers are well incentivised to compromise the system to gain access to it. Furthermore, there is no real method of detecting when an unauthorised individual actually views information but does not edit it. The information can still be captured and used. The attacker is well incensed and is unlikely to be detected. Therefore, it is deemed *likely* that attackers may try and comprise the system to gain access to personal information.

6. *Impact Analysis*
   If an attacker was able to compromise the system there is a *loss of integrity* and *loss of confidentiality*. The *loss of integrity* occurs when an attacker is able to edit personal information. The attacker could potentially poison data with gibberish and nonsense. This could be counterbalanced be retaining previous versions of data. However, this would require storage of information users may not wish to be kept. If users delete their phone number, for example, they may be unhappy to learn the number is retained for precautionary reasons.

   The *loss of confidentiality* occurs because once an attacker compromises the system they are able to view an individual's personal information, even if they choose not to edit it. The loss of confidentiality has serious implications as it puts users at risk. If a user's home address was captured, for example, there is little the user could do, other than move house. Therefore, the impact magnitude is *catastrophic*.

7. *Risk Determination*
   Therefore, with a likelihood rating of *likely* and impact magnitude of *catastrophic*. The risk level of the authentication task, according to Table 4.1, is **extreme**.

Purchase content from a digital store.

The primary task is purchasing digital from a store. The secondary authentication task is to confirm purchase. The purpose for including a secondary authentication task is accountability, while the motivation is content control. The steps 4 to 7 of the risk evaluation are detailed, as follows:

4. *Control Analysis*
   A *technical control method* can be put in place to ensure the application can

be remotely deactivated if a breach occurs. Furthermore, purchased digital content can be restricted to the application itself. Therefore, if an individual is able to purchase digital content and subsequently download it, it would reside with the application. If the application is remotely deactivate, an attacker will not be able to consume content.

5. *Likelihood Determination*
   The motivation would depend on the content available from the digital store. If the store offered premium digital content, e.g. Hollywood productions, attackers would be sufficiently motivated. The attackers could transfer the content on to other mediums and sell it at a much lower cost, as they have bore no real cost in obtaining it. However, if the digital content is not premium then there may be little motivation to attack the system. The digital content our mechanism would be protecting would obviously not be Hollywood productions. However, at the same time the content would be above free-distribution. Therefore, there would be some motivation to compromise the system. Nevertheless, if digital content remains within the application, access to the content can be controlled by deactivating the application itself. Therefore, it is deemed *possible* that the system will be compromised.

6. *Impact Analysis*
   There is no real concern if an attacker is able to purchase digital content. The user is not placed at risk, it is only the content supplier that is placed at risk. The content supplier can make the decision if the risk is worth access to the user-base. Furthermore, if content is exposed, the application can be remotely deactivated and an attacker would no longer be able to consume it. Therefore, the impact magnitude is *moderate*.

7. *Risk Determination*
   Therefore, with a likelihood rating of *possible* and impact magnitude of *moderate*. The risk level of the authentication task, according to Table 4.1, is **medium**.

Make a phone-call

The primary task is making a telephone call. The secondary authentication task is to ensure only authorised individuals are able to access the phone. The purpose for including a secondary authentication task is regulate access, while the motivation is piece of mind. The steps 4 to 7 of the risk evaluation are detailed, as follows:

4. *Control Analysis*
   A *technical control method* can be put in place to ensure the application can be remotely deactivated if a breach occurs.

5. *Likelihood Determination*
   The motivation would depend if an attacker wanted to make a call to a specific person or gain access to a user's phone. An individual's personal mobile

phone can potentially hold a lot of information about a user, including conversations with other individuals, precious photographs as well logs of call made to others. There is little that can be done if an attacker is able to make a phone-call or compromise a user's phone. The only option would be to remotely deactivate the application or phone, a potentially great inconvenience to the user. Therefore, given a user's phone presents a respectable treasure to an attacker and that little be can down to thwart them, once compromise. It is deemed *likely* that the system will be compromised.

6. *Impact Analysis*
   The impact on the user is disastrous if an attacker is able to access their mobile phone to make a phone call. A user can recover from the incident, in terms of explaining to others there any calls where not from them and the phone can be restored from a back-up. However, information on the phone or call made by an attacker have the potential to not only harm users but others as well. The other individuals affected in a potential breach are not using the experimental application or even partaking in it. However, this arguably depends if the final implementation of the mechanism is an application that makes voice-calls between users or is a screen-lock for a mobile phone. Nevertheless, the impact would be great on the user. Therefore, the impact magnitude is *major*.

7. *Risk Determination*
   Therefore, with a likelihood rating of *likely* and impact magnitude of *major*. The risk level of the authentication task, according to Table 4.1, is **high**.

View a friend's location.

The primary task is viewing a friend's location. The secondary authentication task is to ensure only an authorised individual is able to access the application. The purpose for including a secondary authentication task is to regulate access, while the motivation is user management. The steps 4 to 7 of the risk evaluation are detailed, as follows:

4. *Control Analysis*
   A *technical control method* can be put in place to ensure the application can be remotely deactivated if a breach occurs. Furthermore, all individuals using the application can be contacted via email or text message, informing them of the breach and to destroy the application.

5. *Likelihood Determination*
   The motivation of the attacker would depend if they wanted to see the location of a specific individual or collection of individuals. The location of an individual is very sensitive information as it can reveal personal appointments and habits. If an individual was located in a clinic, bank or a local fast-food chain, for example. An individual may want to share with a friend but not attacker. The information extracted could be used to blackmail or compromise another. An attacker could contact a specific bank and detail

that they conducted a transaction at a branch, at such a time. Furthermore, if the location of another individual is compromised there is little action that can be taken, other than informing the individual immediately to move location. Nonetheless, the location could be an individual's home or workplace. Therefore, given the sensitivity of the information and value to an attacker, it is deemed *likely* that the system will be compromised.

6. *Impact Analysis*
   The impact on users is cataclysmic if an attacker is able to compromise the location information. An attacker could potentially access location information several times before the breach was recognised. Furthermore, the breach of the system puts potentially many users at risk, not just a single user of the system. Therefore, given the grave risk users are placed in and the fact little can be done once the information is leaked, the impact magnitude is *catastrophic*.

7. *Risk Determination*
   Therefore, with a likelihood rating of *likely* and impact magnitude of *catastrophic*. The risk level of the authentication task, according to Table 4.1, is **extreme**.

Read an article

The primary task is reading an article from a reading list. The secondary authentication task is to ensure only an authorised individual can access the application. The purpose for including a secondary authentication task is to regulate access, while the motivation is user management. The steps 4 to 7 of the risk evaluation are detailed, as follows:

4. *Control Analysis*
   A *technical control method* can be put in place to ensure the application can be remotely deactivated if a breach occurs. Furthermore, all individuals using the application can be contacted via email or text message, informing them of the breach. Moreover, all articles could be stored for a specific period before being actually deleted from the service. Therefore, if an attacker does compromise the system and delete articles, the application could be 'rolled-back' to an agreed state.

5. *Likelihood Determination*
   The motivation of an attacker would be very low, as there seems to be few if any discernible reasons as to why an attacker would compromise a reading-list application using an experimental authentication mechanism. Therefore, given the lack of motivation, it is deemed *rare* that the system would be compromised.

6. *Impact Analysis*
   The impact on the user from an attacker compromising the system would relatively minor. The attacker could delete the articles collected by an individual but this concern can be overcome with technical control mechanism.

Therefore, given the fact users would not be dramatically inconvenienced, the impact magnitude is *minor*.

7. *Risk Determination*
   Therefore, with a likelihood rating of *rare* and impact magnitude of *minor*. The risk level of the authentication task, according to Table 4.1, is **negligible**.

Personalise a service

The primary task is personalising a service. The secondary authentication task is to ensure only an authorised individual can access the personalised service. The purpose for including a secondary authentication task is to regulate access, while the motivation is user management. The steps 4 to 7 of the risk evaluation are detailed, as follows:

4. *Control Analysis*
   A *technical control method* can be put in place to ensure the application can be remotely deactivated if a breach occurs. Furthermore, all individuals using the application can be contacted via email or text message, informing them of the breach. Moreover, all data used to personalise the service would be purged.

5. *Likelihood Determination*
   The motivation of an attacker would be relatively low, as there is little benefit for an attacker to compromise a personalised service. An attacker would only be able to receive tailored responses to specific requests or searches. This suggests an attacker would have no real motivation to compromise the system. Therefore, given the lack of motive, it is deemed *unlikely* that the system would be compromised.

6. *Impact Analysis*
   The impact on the user from an attacker accessing a personalised service appears negligible. The tailored responses sent to attacker would not have any real any impact on the user. Therefore, given the fact they stand to experience no real impact, the impact magnitude is *insignificant*.

7. *Risk Determination*
   Therefore, with a likelihood rating of *unlikely* and impact magnitude of *insignificant*. The risk level of the authentication task, according to Table 4.1, is **negligible**.

Add something to password manager

The primary task is adding an authentication secret to a password manager. The secondary authentication task is to ensure only an authorised individual can access the password manager. The purpose and motivation for including a secondary authentication task is to regulate access. The steps 4 to 7 of the risk evaluation are detailed, as follows:

4. *Control Analysis*
   The manager will store various objects, not just passwords. Therefore, the application can not be remotely deactivated as this would inconvenience the user. A *technical control method* can be put in place where users are contacted by email to inform them of breach.

5. *Likelihood Determination*
   An attacker will be highly motivated to compromise a password manager as it contains secrets for a range of application and services. Users may choose to save credit card and debit card information, as well as information access online services. Furthermore, given the fact an experimental mechanism is being used, attackers may feel they could easily find design flaws and bugs. Therefore, given the motivation, it is deemed *certain* that the system would be compromised.

6. *Impact Analysis*
   The impact on the user would be dramatic if not catastrophic. If an attacker is able to access a user's personal password manager they may be able to access any number of services. An attacker could conduct simple to nefarious attacks. These attacks could range from changing passwords to initiating remote wipes on personal devices. If an attacker changes the passwords to services, then users will be unable to access the service. If an attacker is able to remotely wipe devices by accessing such services through the web, the could lose precious data, e.g. photographs. Therefore, given the fact stand to experience no real impact, the impact magnitude is *catastrophic*.

7. *Risk Determination*
   Therefore, with a likelihood rating of *certain* and impact magnitude of *catastrophic*. The risk level of the authentication task, according to Table 4.1, is **extreme**.

Summary

Table 4.2 outlines a summary of the various potential tasks. The tasks with a risk level of *Extreme* or *High* are probably best avoided for use in the field investigations. Table 4.3 outlines some of the concerns surrounding the potential tasks. There are various concerns with the tasks. However, the second task appears to have few concerns as a client or distributor is exposed not the user. These aspects will influence the proposed workflow.

Task	People	Place	Platform	Purpose	Motivation	Risk
1	Consumer	Secluded	Agnostic	Regulate Access	Rules & Regulation	Extreme
2		Shared	Television	Accountability	Content Control	Medium
3		Shared	Smartphone	Regulate Access	Regulate Access	High
4		Shared	Smartphone	Regulate Access	Regulate Access	Extreme
5		Secluded	Tablet	Regulate Access	User Management	Negligible
6		Shared	Agonistic	Regulate Access	User Management	Negligible
7		Secluded	Agonistic	Regulate Access	Regulate Access	Extreme

Table 4.2: Summary of Potential Workflows, including associated risk level.

Task	**User unable to complete authentication**	**Evaluation and Implementation concerns**
1	No real risk	Serious implications if an unauthorised individual is able to view personal information. Server-side an attractive target for attacker.
2	No real risk	Content provider may expect reimbursement if content is accessible to unauthorised individuals.
3	Users unable to use smartphone could place them in danger	The mechanism would need to be bundled as part of an smartphone OS, requiring complex install.
4	No real risk.	Location information of individuals exposed to unauthorised individuals. Storage of location information could be compromised on server side.
5	No real risk	Users may use application store sensitive documents or copyrighted articles
6	No real risk	Non-trivial to construct a service worth personalising that users would use regularly.
7	No real risk	Serious implications if an unauthorised individual is able to view authentication secrets for a range of services and systems.

Table 4.3: Concerns surrounding the potential tasks.

4.5 Proposed Solution

4.5.1 Task

The candidate tasks, outlined and discussed in §4.3, were considered and assessed for risk, in §4.4. The second task was selected as the evaluation task, namely:

- *Access content from a digital store.*
  An individual accesses an application containing a library of digital content and selects items for download.

The primary task context was similar to that of the envisioned context of Tetrad, aspects outlined and discussed in §4.2.1, namely:

- *User*
  The envisioned users of both the task and Tetrad are consumers, not professionals.

- *Device*
  The envisioned optimal device for both task and Tetrad is the television. The assumption is that the television is the biggest and best display for most consumers. Moreover, it likely possess the most sophisticated sound system for most consumers. Consequently, it will represent the best device to access several different types of digital content, such as music, films, television programmes and games.

- *Environment*
  The envisioned environment for both the task and Tetrad is a shared space or the living room.

Furthermore, Tetrad was deemed to meet the requirements of the secondary authentication task, aspects outlined and discussed in §4.2.2, namely:

- *Motivation*
  The motivation for including an authentication step within the task is to ensure only authorised individuals are able to access content. The motivation aligns well with Tetrad as users are not exposed to harm, if the mechanism fails. The only element under threat in such a situation is the content itself.

- *Cross Context*
  The task is expected to cross into different context and be used on different devices, specifically smartphones and tablets. Consequently, it was determined that Tetrad would cross contexts unchanged to these devices.

Lastly, the aforementioned task was determined to have a medium risk level, detailed in §4.4. The user is not exposed is Tetrad does not function properly. The next step was actualise the task in the form of an application.

4.5.2 Application

The initial incarnation of Tetrad was designed and developed for large screen displays, such as those found in televisions. The assumption was that Tetrad would form part of a task, performed on televisions. Nevertheless, the reality is that as tasks cross contexts onto other devices, so does an authentication mechanism. Therefore, the initial assumption was that the application, containing Tetrad, would be initially evaluated on televisions, before expanding onto smartphones and tablets. Nevertheless, a television platform that would allow for the distribution of an application containing an experimental authentication mechanism could not be identified. Consequently, the application was designed, developed and evaluated for smartphones and tablets, as (1) acceptable distribution platforms could be identified for both and (2) the authentication mechanism had to be explored in other contexts.

The first step was sourcing a content distributor for the candidate task. This lead to conversations with School of Psychology at the University of Glasgow. The School of Psychology wanted students to be able to download and access lecture recordings on their mobile devices. Furthermore, they wanted access to the lecture recordings regulated by an authentication approach, ensuring only enrolled students were able to access the lecture recordings.

Therefore, an agreement was reached where we would provide such an application protected with our alternative authentication mechanism. The School of Psychology acknowledged the approach was experimental and could have flaws that may expose their content. They agreed this was acceptable as long as flaws could be tackled swiftly and if not the content was to be removed. Furthermore, the School of Psychology required designs to be approved and discussed with themselves before any applications would be deployed to students. Lastly, the School of Psychology made it clear that application was to be of an acceptable quality as they would not deploy a low quality application to students.

Therefore, we agreed to deliver an application that would allow students to download and listen to lecture recordings. The application would require users to authenticate to access it. There was three iterations of the application created, codenamed 'Tom', 'Dick' and 'Harry'. The labels are inspired by the tunnel names mentioned in the 1963 film 'The Great Escape'. The assumption in the film is that not all tunnels would necessarily be successful. Similarly, the assumption when creating the applications was that not all would not necessarily be successful. The design, implementation and evaluation of the initial application, 'Tom' is outlined in the next chapter, followed by similar chapters for 'Dick' and 'Harry', respectively. The application itself would be used to evaluate Tetrad within an actual task. Therefore, a number of research questions were outlined to explore in subsequent evaluations.

4.5.3 Research Questions

There are two keywords in the thesis statement, outlined in §1.2, namely *viability* and *wild*. A number of research questions were used to determine the viability of the authentication mechanism. These questions are detailed after the steps taken

to ensure the mechanism was evaluated in the wild, below:

1. The first step was to create the foundation for a recognition-based graphical authentication mechanism, as outlined in the thesis statement. The existing research and literature was reviewed. The decision was taken to create a novel graphical authentication mechanism that focused on observation-resistant entry rather than resurrect an existing approach. The approach relied on images of faces and entry of the authentication secret was resilient to observation from onlookers.

2. The recognition-based graphical authentication mechanism had to be evaluated in the field, as outlined in the thesis statement. However, before committing to expensive and complex field investigations, the decision was taken to assess the authentication mechanism in a controlled evaluation. Consequently, a prototype was designed, implemented and evaluated to determine if users could cope with the novel authentication approach. The research was presented at BCS HCI 2009 [224]. The controlled evaluation, as well as the discussions at the conference, concluded that the authentication mechanism showed promise and did address an actual authentication concern.

3. The next step was to develop and extend the authentication mechanism for deployment in the field. The approach was to assess it in the wild as outlined in the thesis statement. Consequently, the authentication mechanism had to be primed for use in a specific authentication context. Therefore, a suitable task that required authentication had to be designed and evaluated for risk. Furthermore, a client and user base had to be sourced to use the application. Once these steps were completed all the elements involved in authentication had to be designed and implemented, namely the application and the registration process. These elements had to be throughly considered as they were being deployed, as outlined in the thesis statement, in the wild, i.e. a field investigation with strong ecological validity. Therefore, the elements had to withstand use from individuals that were self-supporting. Lastly, to strengthen ecological validity and ensure the mechanism was facing realistic conditions, as is the case with other applications reliant on authentication, the application was put against competition. Therefore, alternative access to the same resources was made available to individuals to ensure if they did not want to use the application and by extension the authentication mechanism they could abandon it. All these steps were taken to ensure the authentication was being exposed in the wild as outlined in the thesis statement.

4. The initial field investigation revealed that the image set was not suitable. Furthermore, the design of the registration process impacted on user performance during subsequent authentication attempts. Consequently, a second iteration of the application was designed, implemented and evaluated using a personal image collection sourced from a popular social network.

5. The second field investigation revealed that the revised registration process was suboptimal. The sourcing of images was part of the reason why the registration process was bloated. Consequently, the third iteration of the authentication approach attempted to improve the sourcing of images and to reduce the complexity of the registration process.

The aforementioned steps were taken to ensure the authentication mechanism was evaluated in the **wild**, as outlined the thesis statement. The following research questions were used to determine the **viability** of the authentication mechanism. The viability of the authentication can be expanded in terms of:

- *Convenience*
  The authentication mechanism should not be too expensive in terms of time and effort. The assumption is that time may initially be lengthy due to novelty of the authentication mechanism, it will improve with use and practice. Furthermore, the observation-resilience does not impede performance.

- *Memorability*
  The authentication mechanism relies on images and recognition. Consequently, the images must be memorable to ensure the authentication mechanism can be used practically. Furthermore, efforts made to make the authentication secret more memorable and improve observation-resilience must not impede performance.

Therefore, these areas are tackled in the following research questions. The first research question is:

> Is there a difference in the time taken to complete an authentication attempt based on the number of attempts?
> **Research Question 1**

The time taken to complete an authentication attempt is an important measure of viability and more importantly convenience. While an alternative authentication approach may have lengthy initial authentication times, it is assumed these will improve over time. The second research question is:

> Is there an association between the memorability of an authentication secret and the time taken to create it?
> **Research Question 2**

The memorability of an authentication secret is an important consideration has it is important factor in determining the value of an authentication secret. Moreover, the memorability is important in determining the viability for the given authentication context. The third research question is:

> Is there a difference in the time taken to complete a successful and unsuccessful authentication attempt?
> **Research Question 3**

There should ideally be no real difference in the time taken to complete an authentication attempt as the process is essentially the same only the conclusion is different. Consequently, a difference would indicate there are factors that influence the success or failure of an authentication attempt. The fourth research question is:

> Is there a difference in the time taken to complete an authentication attempt if users have selected distractors as well as targets?
> **Research Question 4**

The image set is an important factor for a graphical authentication mechanism. The image set used in the second iteration of the application required users to select distractors and target images. There is concern that users may struggle to locate target among distractors. The fifth research question:

> Is there a difference in the time taken to complete an authentication attempt if target and distractor images oscillate?
> **Research Question 5**

The image set is an important factor for a graphical authentication mechanism. The image set used in the third iteration of the application required users to select distractors and target images and those images oscillated, if possible. The motivation was to improve resilience to observation and improve retention of the authentication secret. However, there is concern that such changes would be difficult for users to discern and impact on time taken. The final research question is:

> Is there a difference in the time taken to complete an authentication attempt if images are reduced in quality?
> **Research Question 6**

The last question focuses on the concern of completing the registration process over a WiFi connection or cellular connection. A cellular connection is susceptible to traffic-shaping that may result in lower quality images.

Chapter 5

Tom

The process of authentication is always part of a bigger picture, a larger task. Tetrad appeared promising when initially assessed, but the evaluation admittedly had weak ecological validity. The evaluation, like many others, framed the alternative authentication mechanism as a primary task rather than a secondary one [23, 239]. The real test of an authentication mechanism is how it performs as part of the big picture. The reality is that if Tetrad proves difficult, unusable or awkward as part of a larger task, users will simply abandon the entire application or bypass the mechanism itself [24].

Furthermore, another uncomfortable reality for authentication designers is that tasks are no longer restricted to a single device. Tasks involving an authentication mechanism now extend to range of devices. Users may predominantly purchase and watch movies on their television, as it is the optimal device to do so, but they will also expect to do the same task on their smartphone when stuck in an airport lounge. The scenario may be rare but the authentication solution must support several contexts. Therefore, while Tetrad is optimised for the shared-space of the living room and television, it must also function in other contexts. The user can not be expected to manage different authentication secrets and/or solutions for the same task across different devices.

Consequently, the following chapter outlines (1) the actualisation of an authentication task that relies on Tetrad as an authentication mechanism and (2) evaluates the package with actual smartphone users. The ensuing section, §5.1, details the design of the application, focusing on important aspects of Tetrad, such as the registration process. The outlined aspects are used to actualise the application, §5.2. Consequently, a prototype application is produced and evaluated, §5.3, with results reported, §5.4, then discussed, §5.5. Lastly, conclusions are drawn, §5.6, with future steps outlined.

5.1 Design

The application payload is downloading and annotating a lecture recording on a smartphone. The student must be authorised to access the lecture recording, as required by the client, the School of Psychology.

The functionality is achieved with a smartphone application. Authentication

observation resistance is achieved with the proposed authentication mechanism, Tetrad. The design of the shared space prototype of the authentication mechanism was outlined previously. However, the initial design targeted the large display of a television, not the small screen of a smartphone. Therefore, the initial step in designing Tom was to translate the big screen design to the small screen.

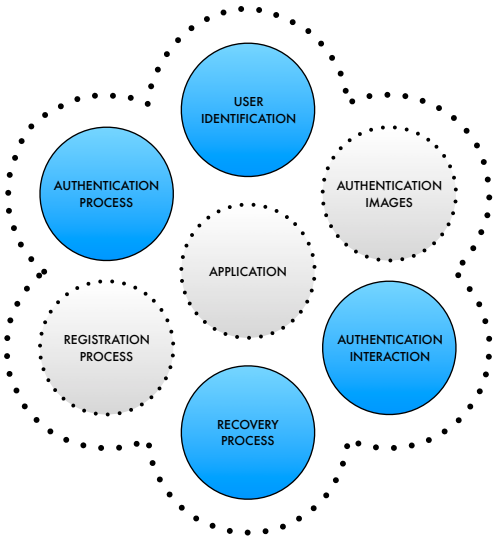

Figure 5.1: The many unspecified elements that had not been considered in the shared space prototype.

Figure 5.1 illustrates the many aspects of authentication that the initial shared space prototype did not consider. The shared space prototype lacks many important elements. There is more to authentication than the process itself [143]. The mundane elements of authentication need to be throughly considered and thoughtfully designed, otherwise it could hamper the entire experience and undermine use of the authentication mechanism or the application itself. A registration process, as well as a recovery process, had to be designed. Therefore, the following elements needed to be designed to ensure an implementation that had ecological validity, as follows:

- *Authentication Interaction*
  The design of the proposed authentication mechanism, specifically interaction, had to be translated to smaller smartphone screens. Section 5.1.1 outlines and discusses the design of the authentication mechanism for a smartphone.

- *Authentication Images*
  The original shared space images relied on an ad-hoc collection of celebrity images. A set of images had to be identified that were suitable for use by Tom. Section 5.1.2 outlines and discusses the image collection used in the design.

- *Registration Process*
  The original shared space prototype lacked a registration process, used to issue or create an authentication secret. Moreover, the registration process determines who is eligible to use the application. Section 5.1.4 outlines and discusses the design of the registration process.

- *User Identification*
  The original shared space prototype lacks a user identification stage. Systems must identify a user to compare an entered authentication secret with one on record. Section 5.1.5 outlines and discusses the design for the user identification stage.

- *Recovery Process*
  The original shared space prototype lacks a recovery process. Users need to be able to recover from failure to remember an authentication secret. Section 5.1.6 outlines and discusses the design of the recovery approach.

- *Application*
  The original shared prototype was not coupled with an application or workflow. Authentication is always a secondary task in reality, not a primary task. Therefore, a suitable application was required to ensure the evaluation had ecological validity. Section 5.1.7 outlines and discusses the design of the application.

The thorough consideration of the aforementioned elements produced a proposed solution, outlined and discussed in §5.1.8, that was used to produce an implementation for use in a ecologically valid evaluation.

5.1.1 Authentication Interaction

The original proposed shared space prototype was designed for a big screen. The initial plan was simply to make the original interface, intact, accessible on the iPhone. Gutwin and Fedak propose panning, zooming or fisheye views for navigating big screen interfaces on small screens [108]. Unfortunately, interacting in such ways is slow and could inadvertently reveal target images, as users linger on specific images of interest.

Therefore, instead of interacting with a big screen interface on the small screen, the decision was taken directly to translate the proposed authentication to the smartphone. However, interaction had to tailored to the iPhone as it relied on touch and not a physical six-button remote control. Three interaction approaches for using Tetrad on a smartphone were considered, as follows:

- *Touch Gestures*
  The user manipulates images using touch gestures, e.g. swiping and tapping the screen of the smartphone.

- *On-screen controls*
  The user manipulates images using on-screen controls, e.g. on-screen direction pad mimicking a physical one.

- *Accelerometer Gestures*
  The user manipulates images using motion gestures, e.g. pitching and yawing the smartphone.

These approaches are all viable for interacting with applications on smartphones. However, not all approaches are necessarily optimal for use with Tetrad. Getting the optimal interaction approach for a smartphone is crucial as the users need to authenticate without help and support. No technician or evaluator will be peering over the shoulder of the user. No one will be around to guide and support the user. They need to be able to authenticate on their own, in a range of environments. Therefore, the advantages and disadvantages of each interaction must be considered.

Touch Gestures

The initial strategy was to take the existing interaction model and simply translate it for a touch-based smartphone. The primary advantage is that a user could transfer knowledge of interaction from the television screen to the smartphone screen. The user is not required to manage multiple interaction approaches for a single authentication mechanism. Figure 5.2 illustrates the interaction mechanism on a touch-based smartphone. The *selection bar* still persists from the original shared space prototype, except that instead of pressing physical buttons, the interface relies on touch gestures. The user performs a vertical swipe to move the selection bar between rows and a horizontal swipe to move the selection bar between columns. The user performs a single tap to reposition images within the selection bar and a double-tap to submit an authentication attempt.

The approach is simple but not without disadvantages. The primary one being that the necessary gestures are not immediately obvious to the user when first encountering the authentication mechanism. Moreover, the interaction model may not be optimal. The user is not directly interacting with images, instead they are manipulating the selection bar. Authentication times could be reduced if the touch gestures directly manipulated the images. However, the use of gestures that are not immediately obvious to a user is still a concern.

Nevertheless, touch gestures could be communicated to the user in advance using video or text instructions. Apple and Microsoft have used both techniques to communicate touch gestures to users [163, 284]. Therefore, touch gestures could not only be used to interact with the authentication mechanism but also to remove a layer of abstraction between the user and the interface. However, the action of the touch gesture must be clear to the user. Pirhonen et al. states that touch gestures must be coupled with feedback and simple single tap gestures should be avoided, as they are easily performed accidentally [208].

Figure 5.2 (d) illustrates an alternative touch gesture interaction approach. There is no selection bar, indeed there are no discernible interface elements. The user is simply presented with images in a grid. The grid comprises manipulatable columns and rows. The user is able to reposition images by directly manipulating a column or row of images within the grid. If users want to move an image to another row, say one beneath, they simply swipe down on the column containing the image. When users are content within their manipulations and ready to

submit an authentication attempt, they simply double-tap. The double-tap interaction gesture is used to avoid accidental submissions from single-taps. The user interface is minimal and focuses attention on the authentication images. Nevertheless, the interface is simplistic and while interactions are not immediately obvious, they can be communicated via text instructions.

On-screen controls

An alternative strategy is to use on-screen controls. A primary concern of using touch gestures is that discrete or non-obvious gestures need to be discovered by users. Therefore, buttons from the Apple remote control, used by the shared space prototype, can be replicated on-screen, much the same way as others have replicated keyboards on-screen. Figure 5.2 (c) illustrates a shrunken authentication canvas with visible dedicated controls. The advantage is the user can see all images. A disadvantage is the images will need to be smaller, making them harder to discern.

Kamba et al. states interface designs for small screens needs to reach a compromise between a physical and functional screen [148]. Interface controls can be displayed on-screen but consume valuable screen space that would otherwise be used to present content. Kamba et al. propose use of opaque on-screen controls [148]. The approach has been deployed in several smartphone applications. Sega have adopted it to port interaction schemes originally designed for physical gamepads to touch-based smartphones. Figure 5.2 (b) illustrates transparent controls overlaying the authentication canvas. The primary advantage of the layout is that the images used in authentication do not need to shrink in size to accommodate the controls. However, a disadvantage is that users may search for images that are underneath their fingertips. If the user lifts fingers from the screen to inspect images underneath, it may be an indicator of target images to onlookers.

Accelerometer Gestures

The use of on-screen controls is desirable as users do not need to interact with images directly. The user does not need to tap anywhere near them or reveal target images. However, shrinking images to use on-screen controls makes them harder to discern and lifting fingers to spot images lurking underneath controls could reveal targets. Therefore, another strategy is to avoid on-screen interaction altogether and use other sensors, such as accelerometers.

Accelerometers determine acceleration in a specific direction: many smartphones use these sensors automatically to rotate the interface between portrait and landscape. Rekimoto states accelerometers can be used to interact with small screen interfaces without sacrificing screen space [218]. Kallio et al. evaluated accelerometer gestures with seven users and reports that none struggle with the interaction approach [147] He et al. states that accelerometer gestures are suitable and practical for many popular smartphone applications [114]. Indeed, the interaction approach has been deployed in popular smartphone applications, such as Instapaper [8]. Therefore, another alternative strategy is to manipulate images

using accelerometers.

Figure 5.2 (a) illustrates the interaction mechanism on a modern smartphone. The selection bar is present but instead of the user tapping an on-screen control or gesture they simply paw and yawn the smartphone to reposition the selection bar. The user single-taps to manipulate images within the selection bar and double-taps to submit an authentication attempt. The advantage of the approach is that the user rarely interacts with the screen and images do not need to be made smaller to accommodate controls. However, a disadvantage is that unintuitive accelerometers gestures may not be immediately obvious to users.

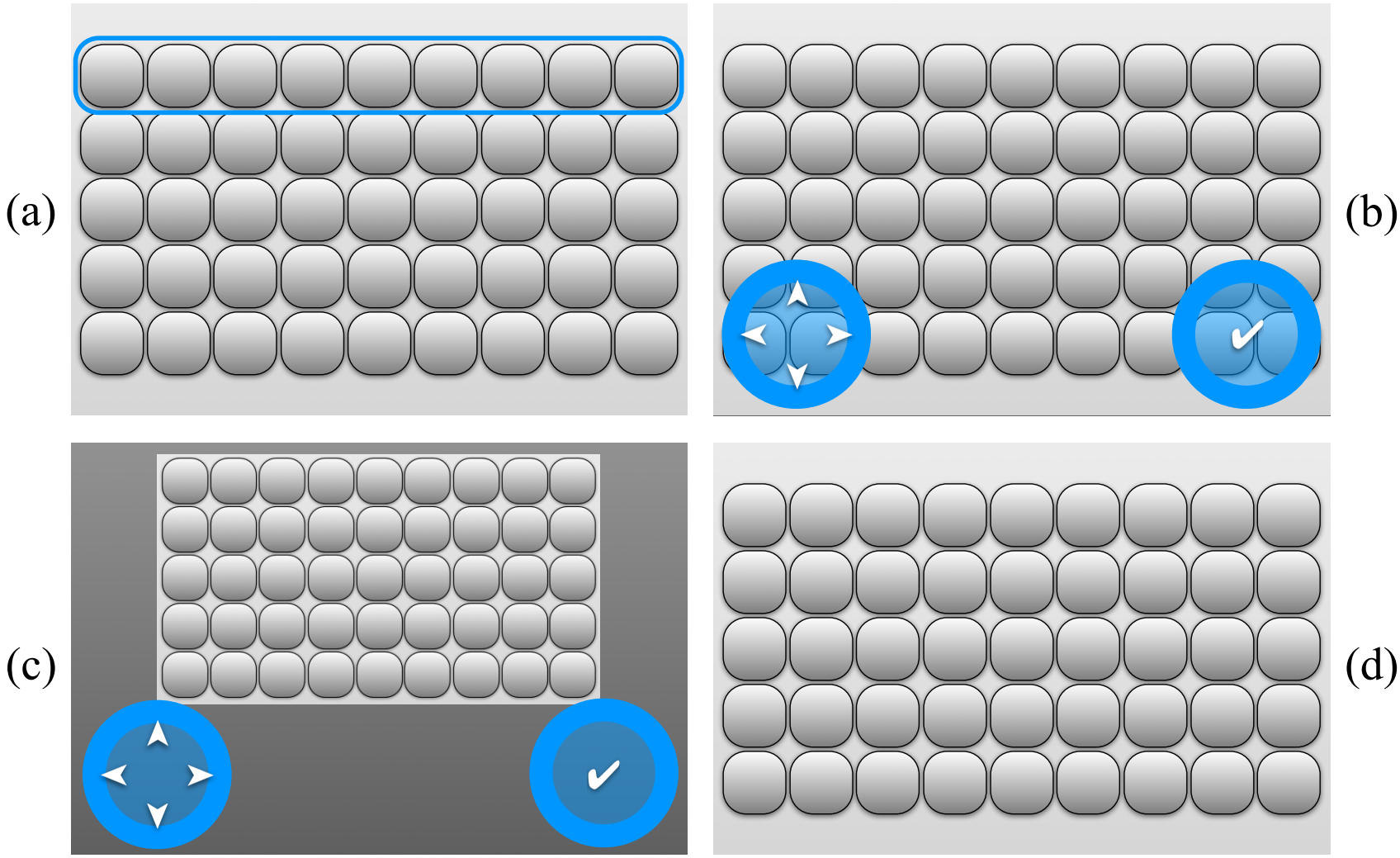

Figure 5.2: The appearance of the authentication mechanism using various interaction approaches

Summary

There are clearly advantages in using any of the aforementioned interaction approaches with Tetrad. However, there are also disadvantages with each approach. Furthermore, it is difficult to assess user experience without developing and exploring prototypes.

Therefore, prototypes were developed and explored before committing to a final implementation choice.

5.1.2 Authentication Images

The alternative authentication mechanism essentially relies on an image set of staff profile pictures, essentially local celebrities. The assumption is that these images will be memorable and recognisable by the user base. All users are presented the same image set during registration and authentication. The image set comprises the following images:

- *Target images*
  The images that comprise the authentication secret.

- *Distractor images*
  The images used to distract attackers away from the important target images.

The user is required to select four images to be their authentication secret: these become the user's target images. The remaining images become distractor images. The image selection process is important as users have to present the same sequence of images to authenticate.

5.1.3 Authentication Process

All users are presented the same image set within a grid. The user is expected to locate the target images they identified during registration among the distractor images. The user manipulates the grid by rearranging columns and rows of images so that the target images align either horizontally, vertically or diagonally. The images must align in the originally selected sequence.

5.1.4 Registration Process

There was no registration process in the shared space prototype. The authentication mechanism was not deployed to an actual user-base. Consequently, there was no need for a registration process that users could complete without guidance. However, when deploying the authentication mechanism to an actual user-base, a registration process is required, enabling users to create authentication secrets without intimate and immediate support.

Registration processes are rarely researched outside the realm of the World Wide Web. Naturally, the ideal solution would be to bypass the registration process entirely [96]. Unfortunately, users typically endure a registration process

before they are able to authenticate. Users should, ideally, only encounter and endure a registration process, once. Nevertheless, the design of the registration process is interesting as it is not itself necessarily bound by the same constraints as an authentication mechanism.

The purpose of the registration process is to associate a user with a system. The registration process determines if the user is eligible to register and then associates an authentication secret with the user, either a user-generated or system allocated one. However, the remaining elements of context — platform, person and place — are dictated by the process itself.

Nevertheless, the registration process is typically the first element of an application users encounter. Therefore, the registration process needs to be treated as a learning stage, that can train and teach interaction approaches to users. Therefore, there are three elements that need to be considered when designing a registration process are:

- *Context control*
  The registration process dictates the context, e.g. users need to complete the process in a secluded space.

- *Collection of personal information*
  The registration process can be used to determine who can register as well as elicit personal information for other purposes.

- *Learning stage*
  The registration process is often the first element users encounter of an application. Therefore, the process itself is an opportunity to train and teach users.

The aforementioned elements all had to be thoroughly considered before outlining a potential design for the registration process.

Context Control

The design does not need to complement the current context the user is operating within: rather the design can dictate the context to the user. The accompanying authentication mechanism is designed to offer resilience to observation. The registration process does not need to offer resilience. There are several reasons observation resilience is not essential, if not potentially damaging, for a registration process. They are, as follows:

- *Potential confusion surrounding the authentication secret*
  The authentication secret could become ambiguous in the mind of the user in a registration process designed to be resilient to observation. The registration process would likely rely on redundancy or indirect entry to achieve observation resilience. The concern is that a user may believe they have created an authentication secret, using specific image elements, when in actual fact they have created an authentication secret comprising of different image elements. The user would not be aware of the error until authentication fails. Even then they may assume they have forgotten the authentication secret, rather than misinterpreted the registration process.

- *Interference effect on authentication*
  The interaction adopted in the registration process to create an authentication secret may impact on entry of the authentication secret. The interaction used in the registration process to achieve observation resilience may result in confusion during authentication. Individuals may perform interactions used during registration during authentication, resulting in failure. The user may not even be sure of the error that occurred, they may assume they entered the wrong authentication secret rather than that they entered the correct authentication secret incorrectly.

- *Difficult to design using software alone*
  The design of a registration process, contained entirely within software, resilient to observation, is difficult to envision. The primary problem is how to communicate with the user in secret. The registration process, for example, could present the user with a grid of images and ask them to reposition images to specific *target cells*. Images located within target cells would become the authentication secret. The only problem is how to communicate the target cells to the user, without others observing. The target cells could be fixed but then any other user who has registered with the system previously, would know the location of the target cells. The target cells could be communicated via another channel but there is no guarantee others would not observe such communication. Furthermore, this assumes the user has a secondary channel in the first instance.

The registration process should perhaps rather instruct users to complete the process in a quiet, secluded environment, free from distractions and onlookers.

Personal Information Collection

The registration process is primarily about generating an authentication secret. A registration process should be designed [97] to ensure that it is:

- *Rapid*
  The registration process can be completed efficiently and effectively by the majority of users.

- *Shallow*
  The registration process should not probe for personal information that is not necessary for authentication, i.e. privacy preserving.

However, the registration process could offer the opportunity to collect personal information. Such information is potentially valuable to an organisation in customising a user's experience and targeting advertisements. Nevertheless, users may feel uncomfortable in providing such personal information to a system or service, especially an experimental one. However, the reality is that most users view personal information as worthless [131] and are willing to sell it for as little as 25 cents [107].

Therefore, users may indeed provide personal information, as long as the process is not particularly time consuming. However, merely because they deem

certain personal information worthless, this does not mean they will answer all questions accurately. There are three aspects to consider when probing personal information [177], as follows:

1. *Not all information is disclosed readily*
   Users may readily surrender certain personal information but are less likely to disclose more sensitive information.

2. *Savvy users are less likely to respond*
   Users who strongly value personal privacy are far less likely to respond to probes for personal information.

3. *Explanations do not improve response rate*
   Furnishing users with explanations or reasons for probing for particular personal information does not improve response rates.

There are also concerns beyond collection of personal information, namely storage. The primary concern in collecting personal information is that users may be put at risk if the experimental system is compromised or collapses.

Presentation

Users were presented four panes that they could vertically swipe through. Figure 5.3 illustrates the layout or presentation of the registration process. The purpose of each pane is:

(a) *Information Screen*
   The information screen greets the user and provides instructions as well as an overview of the registration process.

(b) *Image Pool*
   The user can horizontally swipe through the image pool and double-tap images to add them to the authentication secret.

(c) *Authentication Secret*
   The user can view the authentication secret. They can reposition images within the authentication secret by dragging individual images. Moreover, the user can remove an image from the authentication secret by double-tapping it.

(d) *Confirmation*
   The user is presented a confirmation button. The user is required to press the button to confirm the creation of the authentication secret.

Users can vertically swipe between panes at any stage, in case they want to consult a particular pane before progressing to another, e.g. consulting the information screen before confirming the authentication secret.

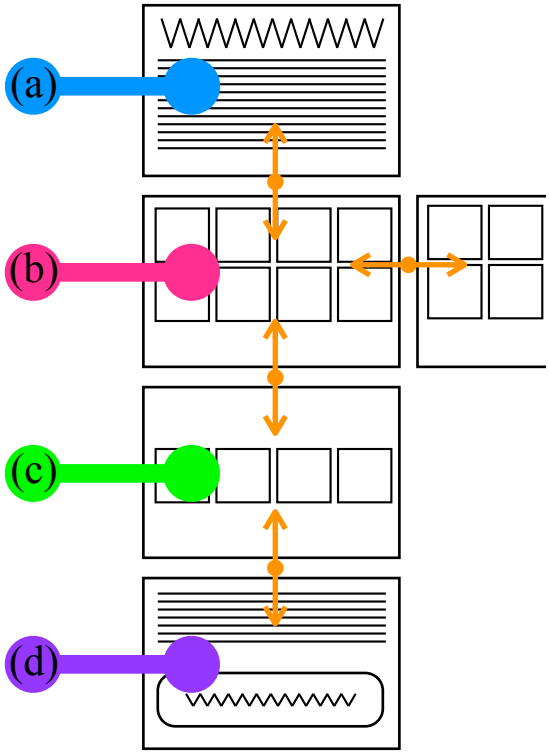

Figure 5.3: The presentation of the registration process.

Summary

The registration process does not collect any personal information, aside from the unique identifier of the device initiating the registration process. The user is not required to provide any personal information or create a username.

5.1.5 User Identification

Just states authentication is not just a single process but comprises of many procedures, user identification being an important one [143]. User identification in most knowledge-based authentication schemes typically relies on usernames, i.e. strings of characters. A username generally acts as an index, used by the system to locate a specific user record. Records can contain a variety of information but at a very minimum will contain an authentication secret for the user.

Therefore, the user is typically required to submit a username and an authentication secret during authentication. There are many different strategies for managing such credentials [142]. Nevertheless, many organisations opt to manage user credentials themselves so as to retain control over valuable user information.

Applications need to ensure each username is unique, if they are to be used to retrieve records. Usernames are generally generated or chosen during the registration process. There are variety approaches to creating usernames, three potential approaches are:

1. *System-generated username*
   Usernames are generated automatically by the system and issued to users. The advantage of the approach is that a username can be created and issued rapidly, streamlining registration. Moreover, using system-generated usernames, makes it far more awkward for attackers and others to track an individuals movements and actions across systems and services. However, a disadvantage of the approach is that users may find it difficult to recall the username when required, especially if they have to manage several such usernames. The solution, at least in modern web browsers, is to allow the web browser to manage credentials or store the username in a small text file on a user's device, i.e. a cookie. Nevertheless, the user may encounter problems if they attempt to authenticate on other web-browsers or devices.

2. *User-generated username*
   The username is generated by the user. The advantages of allowing individuals to create usernames is they should be easier to remember and the user does not need to sacrifice any personal information, such as an email address. However, registration may take longer to complete as users spend time creating a username that is unique. Moreover, users may opt to reuse a username, inadvertently allowing other organisations and individuals to track movements across various systems and services [9, 204].

3. *Email-address as username*
   The username is an individual's email address. Using an individual's email address as a username represents a middle ground between system-generated and user-generated approaches. An email address is assuredly unique, ensuring the registration process will not be lengthy because of username generation. An individual does not need to waste time trying to craft a username, that no other user has registered. An email address is ensuring users will not struggle to recall it [9]. However, using an email address allows other entities and attackers potentially to track and aggregate user movements between systems and services. Moreover, using an email address may be used for purposes other than authentication, such as marketing [9, 204].

The last approach, an email address as username, is likely the best compromise. An email address should be memorable to most users and will not slow down the registration process. Nevertheless, there are still concerns surrounding the collection and storage of email addresses.

However, there is always the option of not using usernames at all, at least not usernames perceptible to users. There are many consumer scenarios where an individual authenticates without entering a username. An example would be paying for shopping at a local supermarket, using Chip and PIN. The user would insert a bank card and enter the relevant PIN. The user is unaware of any identification stage. The bank card itself contains enough information to identify the user.

The user may be accessing the authentication mechanism from a television, desktop or mobile phone. The user may even be accessing the authentication from a public machine, such as a desktop computer in a local library. The device itself

is discounted when constructing a web application. However, if the authentication mechanism is part of an application, the device itself may provide the information necessary to identify the individual without needing to burden the user with providing a username.

- *Recall vs Recognition*
  The user identification stage could require an individual to recall a username or recognise it.

- *Interaction*
  The user identification stage could rely on hardware that is not present on a device, e.g. televisions do not typically come equipped with keyboards.

- *Refacing*
  Usernames are used to *tell* a system about a user. However, they could be refaced or redirected towards a user. The 'username' could tell a user how much a system knows about them.

The aforementioned elements all need to be considered when designing a registration process, suitable for use with the proposed authentication mechanism, Tetrad.

Recall vs Recognition

While an authentication mechanism may incorporate an identification stage, it is not clear what form that stage will take. The stage would have typically relied on usernames, i.e. strings of characters. However, authentication typically takes the form of passwords, i.e. strings of characters. The identification and authentication stages are similar: both rely on string of characters. Furthermore, they are both framed as recall memory tasks. The user is required to recall and type-in an alphanumeric username and an alphanumeric authentication secret.

The same is not true of the proposed graphical authentication mechanism. The proposed authentication approach is reliant on recognition. Therefore, using usernames would result in the coupling of a recall memory task with a recognition memory task. There is nothing inherently improper in creating such a coupling but such a design appears awkward.

Interaction

The proposed graphical authentication approach is not reliant on a keyboard, a great advantage on devices that lack keyboards, such as televisions. The user does not need to struggle with an on-screen keyboard and a four-direction button to enter a lengthy authentication secret. However, expecting users to struggle with an on-screen keyboard to enter a lengthy username seems to negate any such advantage. The authentication approach has not then actually spared the user the inconvenience of using an on-screen keyboard. Moreover, if the user is able to enter an alphanumeric username, they may well be content to enter a password, rather than learn yet another interaction process.

Summary

The user will not be expected to create a username and instead the unique identifier of the device will be used to identify the user.

5.1.6 Recovery Process

There are several possible designs for the recovery approach. Deciding on a recovery approach for an alternative authentication mechanism could take many different directions. The decision was taken to not offer an explicit recovery process since this was outwit the scope of the research. The user will be expected to complete the registration process again to use the application.

5.1.7 Application

The workflow is downloading a lecture recording for listening and annotating on a smartphone. Therefore, the application is an audio player with accompany digital store.

The School of Psychology record all lectures, for all years. The lecture recordings are compressed and distributed through an online resource centre, referred to internally as the 'portal'. The portal is managed by dedicated technicians in the School of Psychology. The portal is accessible to all enrolled students through a typical web browser. The portal is designed and targeted at laptop and desktop computers. Students enter a username and password before they can access it. The lecture recordings are compressed using the MP3 encoding scheme and are not wrapped with any digital rights management. Therefore, students have the option to download lecture recordings from the portal and essentially transfer them to any device that decodes MP3 files, e.g. a digital audio player.

However, the workflow of downloading and transferring a lecture recording to a smartphone is awkward and cumbersome. Other educational institutions have addressed the problem by adopting podcasting. Using podcasts simplifies the process greatly. The user simply subscribes to a podcast using specialised software, available on several devices. When an institution releases a new recording, a user's device is notified and proceeds to download it. Moreover, devices can synchronise to ensure the latest recording is available across device, e.g. a lecture recording is automatically transferred from a personal computer to a digital audio player. A user is almost unaware of the entire process, ensuring the focus is on the lecture recording itself.

The School of Psychology simply sidestepped podcasting, due to concerns of unauthorised access. Nevertheless, the School of Psychology acknowledge that the solution was non-optimal for mobile devices and were interested in a bespoke mobile application that offered authorised access. The primary aims of the application are, as follows:

- *Authorised Access*
  Users need to authenticate to access the application. The student should be able to access any part of the application without authentication.

- *Lecture Store*
  The user is able to peruse a list of lecture recordings and select specific recordings for download.

- *Audio player*
  The majority of audio player applications are design to play music that is 3 minutes in length. A typical lecture is 50-60 minutes in length. Therefore, the audio player needs to be designed to reflect such length.

The initial aim would be addressed by using the proposed authentication mechanism. The second and third aim would be addressed by designing and implementing a store and audio player, respectively.

However, another aspect, from an evaluation perspective, is that the application should incentive users to return to it, promoting recurring use. The incentives must be considered, as inappropriate or unsustainable incentives would dilute the ecological validity of the evaluation. The application should attract users back to it, generating authentication attempts but not in a way that is unrealistic. Furthermore, incentives could not be offered that would undermine or damage the reputation of the School of Psychology. Therefore, after thorough consideration and discussion, the following incentives were agreed with the School of Psychology. They are, as follows:

- *Chapters*
  The smartphone application would allow students to chapter lecture recordings. Therefore, a user could pinpoint specific parts of the lecture recording for later revision. The same way that students use a highlighter to highlight passages of a text.

- *Annotations*
  The mobile application would allow students to attach an image and notes to a chapter. Therefore, a user could attach a specific diagram or image to a specific part of the lecture recording.

- *News*
  The mobile application would provide a news feature that would keep students up-to-date with the latest course information and scheduling.

The design of the authentication mechanism was considered in §5.1.1 and §5.1.2. However, the remaining aforementioned elements need to be considered to determine the design of the application.

Lecture Store

The design of the lecture store was modelled around the several successful digital stores operated by Apple. The digital stores operated by Apple are amongst the most successful distribution platforms in the world [207]. Apple and Google stores have both served other 25 billion downloads to customers [7, 245]. There are many reasons for such successes but the design of such digital stores is likely a major contributor.

Apple launched the iTunes Store on the 28th April 2003 with a catalogue of over 200,000 songs [6]. The iTunes Store now has a music catalogue of more than 28 million songs [247]. The initial catalogue size may seem small in comparison until you consider a larger ASDA supermarket carries 40,000 items on its shelves [12]. Therefore, the digital stores are designed to enable users to navigate vast, complicated collections of content.

The digital stores are standalone smartphone applications. The user peruses the digital store, selects an item to purchase. The content is then downloaded and delivered to another application on the system. An example would be a song purchased from the iTunes Store would be downloaded and delivered to the Music smartphone application.

The solution clearly works well for millions of users. However, the School of Psychology requires that only authorised users can access lecture recordings. There is no reason to make the user authenticate to access the lecture store and then authenticate to access the audio player. Moreover, the lecture store would only need space to navigate a few hundred lecture recordings, not the thousands or millions handled by the iTunes Store. Therefore, the lecture store would be incorporated into the main application and not offered as a standalone application.

Furthermore, the audio player is designed for lecture recordings, not general audio files. There is no way to download lecture recording without using the lecture store. Therefore, the audio player and the lecture store have a symbiotic relationship. Lecture recordings downloaded from the store are only playable on the audio player. The audio player is only able to play lecture recordings from the lecture store. Therefore, they should form a single application rather than two distinct applications that are dependent on each other. Furthermore, the user is able to access the audio player, listening and annotating other lecture recordings, while others download.

Audio Player

Most modern smartphones are preloaded with a music application. The music application can be used to listen to a range of audio recordings, encoded in many different schemes. However, music applications on most modern smartphones are designed to showcase content purchased from digital music stores.

Therefore, the music application on the iPhone is designed for the majority of music sold on the iTunes Store, i.e. commercial pop songs. Consequently, the preloaded music application is designed for a 3 to 5 minute listening experience. Figure 5.4 illustrates the audio player preloaded on the iPhone. Four distinct elements are highlighted: (a) toolbar outlining information about the recording, (b) an overlay with audio controls for the current song, (c) a large view of album art and (d) audio controls for all songs.

There are several interesting aspects to the overlay with audio controls for the current song. The overlay for one can easily be dismissed by the user and will not reappear unless summoned, unlike other elements in the audio player. The assumption is that the controls are not necessarily important or crucial, suggesting designers at Apple do not envisage many users scrubbing through

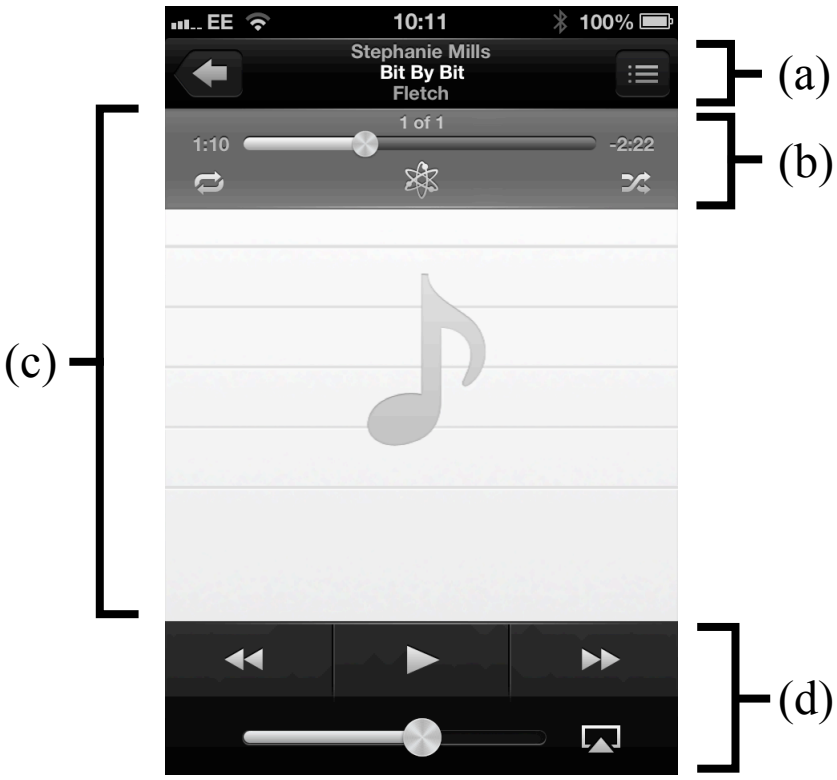

Figure 5.4: Apple 'Music' application interface.

such short audio files. Therefore, the design focuses on users navigating between songs, rather than within a single song. Moreover, the scrubber interaction object is not particularly informative of the differing lengths between audio files. The interface object almost implicitly assumes most audio files are of a similar length, as abnormal, longer lengths are difficult to observe and navigate.

Therefore, the scrubber interface object is not optimal for recordings longer than 5 minutes. There is arguably a need for an audio player designed for lengthier lecture recordings. A design that elegantly communicates a 60 minute length but is itself not confusing to many users. The School of Psychology was adamant that students' focus must be on learning content and not on learning confusing or complicated novel interfaces.

Therefore, the foundation of the interface is a familiar design for representing 60 minutes, recognised by many individuals instantly: a clock face. The clock face is a well established metaphor that enables viewers to observe and understand time quickly. The majority of the interface is allocated to the new scrubber. The appearance of the scrubber is a large circle with a button in the middle. The user navigates the scrubber by touching and gliding a finger along the face of the circle or by tapping a specific point within the circle. The circle represents 60 minutes of audio. The user can tap at the 3rd hand, for example, to jump 15 minutes into the lecture recording, the 6th hand for 30 minutes etc.

The proposed interface has no software-based volume control; instead the user is expected to use the hardware buttons on the device. The bottom section of

the canvas is allocated to navigation while the top is allocated to a toolbar. The toolbar comprises three elements, from left to right, a button to return to a list of lecture recordings, a label outlining the title of the current lecture recording and lastly, a button to access a list of chapters.

Chapters

A chapter is a common and well understood concept in books and text. However, they are rarely seen in digital video and audio. Digital video discs (DVDs) prompted the idea of navigating lengthy video recordings using chapters. The film or television show is segmented, much like a book, into a series of chapters that an individual can use to jump to a specific part of a video recording, avoiding the hassle and time of scrubbing through various scenes.

Chapters in digital audio have rarely been used as digital audio is often relatively short. However, Apple prompted the use of chapters in 'enhanced podcasts'. Apple provide tools to developers and content creators to create chapters for lengthy audio recordings. The required encoding and format was only widely used by Apple iTunes. Therefore, only iTunes users have access to chapters. Nevertheless, the concept is useful to navigate lengthy audio recordings as it allows users immediately jump to specific topics and ideas without needing to scrub through the entire audio track.

However, in both cases the activity of chaptering is restricted to content creators. Users are unable to create chapters and they are unable to segment or emphasise parts of audio. Research and investigation into chaptering of lecturing recordings by lectures or learners is relatively rare to non-existent. However, the activity in many ways can be likened to highlighting text within written and printed notes.

Fowler and Barker state that highlighted text improves retention of information [89]. However, readers need to have confidence. The user has to have confidence that the highlighter is able to discern between relevant and irrelevant content. While a lecturer would likely be loath to class any content within a lecture as irrelevant, clearly students would deem certain aspects more important than others. Moreover, students are likely to have high confidence that a lecturer is able to discern between relevant and irrelevant content in a lecture recording.

Therefore, an initial consideration was to have the digital audio player present chapters from lectures. However, the School of Psychology was concerned that the inferred importance of chapters by students may be that highlighted content signalled examination or assessment topics. Furthermore, the tools themselves would be exclusive to the application. The School of Psychology would not provide any content that required users to own an iPhone or required them to use the application. Moreover, the School of Psychology was not convinced that all lecturers would provide the necessary content. Therefore, the decision was taken not to offer chapters provided by the lecturer.

Instead the application would focus on students chaptering lecture recordings themselves. Lindner et al. argues any advantage in highlighting text is reaped by the highlighter during review [171]. Therefore, the motivation is that students will listen and chapter audio throughout the session rather than before the traditional

revision period. Consequently, the digital audio player supported chaptering, allowing the user to highlight segments of the lecture recording.

An arbitrary limit was placed on the number of chapters a user could create for a specific lecture recording. Students were limited to 12 chapters. The motivation for the design decision was driven by a number of reasons. However, the primary one was the concern that the function may become a procrastination tool for many users. Students may create a vast number of chapters making it difficult to navigate and discern key segments at a glance, negating the purpose of chaptering in the first instance. The limit would require users to consider the placement and length of a chapter, sparingly placing them, rather than simply creating a chapter every few minutes. The limit of 12 chapters was deemed enough, as it represents essentially 6 minutes of audio in a 60 minute lecture. Moreover, the limit could always be increased if the School of Psychology felt the content they supplied would benefit from the ability to create more chapters.

The chapter button is located in the centre of the digital audio player. The user creates a chapter by pressing the button. Each chapter is automatically associated with one of twelve colours, progressing sequentially through them until all chapters are used. A chapter is represented as a coloured segment on the circular scrubber, ranging from the point it was created until the start of the next chapter or end of the recording. The interface looks similar to Goethe's symmetric colour wheel when covered in chapters [101].

The user can view list of chapters by tapping the 'chapter list' button in the top-right corner of the digital audio player interface. The list contains all the chapters associated with a lecture recording; students can add annotations to each chapter.

Annotations

Students can often be seen annotating lecture slides or writing notes in lectures. Hartley and Davies state there are two reasons for students creating notes [112], as follows:

- *Process*
  Students feel the process of producing written notes is conducive to retention of information.

- *Product*
  The faithful reproduction of a lecture will provide a solid foundation for revision.

Similarly, Di Vesta and Gray argue there are two primary purposes for creating notes [69], as follows:

- *External Storage*
  Generate a faithful reproduction of everything communicated as the basis for subsequent revision.

- *Encoding*
  Strengthen knowledge and understanding of lecture material through connection with previously encountered ideas and concepts.

The process of creating notes is advantageous, as is the reviewing thereof [152]. Peper and Mayer suggest the process of creating notes is valuable as it encourages integration of information [201].

Therefore, annotating lectures is a common and beneficial activity. Unfortunately, few students have access to tools that allow them easily to annotate lecture recordings without advanced technical knowledge and experience. Consequently, the feature was added to the application in the hope that it would promote increased use of the application. The ability to add notes was not part of the portal but students could easily create and manage handwritten or typed notes. Therefore, students were not put at a disadvantage for not using the smartphone application or not owning an iPhone.

The application allows students to annotate each chapter with an image, heading and notes. Figure 5.5 illustrates the annotation user interface for adding text notes to each chapter. The user reads and edits text notes by accessing the chap-

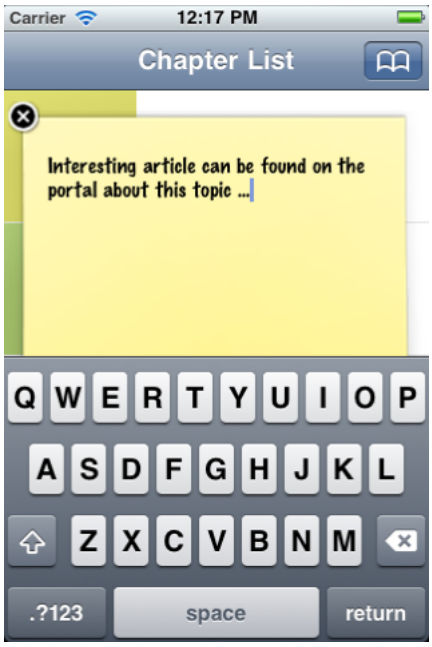

Figure 5.5: Annotation User Interface

ter list. Each chapter element in the list has a note icon on the right: the icon is transparent when there are no notes and yellow if a note has been attached. The user views the notes by tapping the icon. Furthermore, each chapter element in the list has a colour block on the left. The user taps the colour block to attach an image. The image replaces the colour block and acts as the background for the chapter segment on the scrubber in the audio player interface as well.

Therefore, the user can easily see, at a glance, the chapters they have created. The annotations themselves reside on the device and there is no way to export them.

News

Campus and course news is important to students [71]. The School of Psychology portal keeps students up-to-date with the latest course information, such as deadline changes and scheduling information.

Incorporating news into the application not only keeps all students up-to-date with the latest information, it is also the case that an easily accessible news source may promote further use. Consequently, a news carousel was positioned in the lecture store. The carousel presented five news banners. Users tapped on the banners to view a news story. Upon tapping a webpage with the story slides up from the button covering the lecture store. The user simply taps a small 'close' button in the corner to dismiss the webpage.

5.1.8 Proposed Solution

The proposed solution is an application that allows individuals to download and annotate lecture recordings. The user will be expected to authenticate to access the application. The implementation of the application is discussed in the next section.

5.2 Implementation

The 'Tom' prototype was implemented for iOS 4.0 devices, specifically Apple iPhone and iPod touch. The implementation of each component is detailed, as follows:

5.2.1 Authentication Images

The image set used was staff profile pictures. The expectation was that the image set would be memorable and familiar to the user base, minimising need for recovery.

5.2.2 Authentication Interaction

There were a number of possibilities for authentication interaction on mobile devices. The possible interaction solutions for implementation were discussed in §5.1.1. The aim was to provide an interaction approach that was elegant and efficient but did not waste precious screen space, an important factor for mobile devices. There were essentially three interaction approaches to consider, specifically accelerometers, touch-gestures and on-screen controls.

The use of on-screen controls was quickly discounted as they were inelegant and inefficient. The proposed designs either wasted or consumed too much screen-space. The user experience of the remaining interaction approaches, i.e. accelerometers and touch-gestures, was difficult to gauge without actual use. Therefore, preliminarily prototypes of both interaction approaches with the authentication mechanism were created, so as to determine the best suited design.

The interaction approaches were implemented on an Apple iPhone. The implementation, for each case, presented the authentication mechanism in landscape orientation. Accelerometers were initially favoured as a minimum of screen space would be spent on interface objects associated with interaction. The implementation presented a selection bar on-screen that a user positioned left and right as well as up and down by pitching and rolling, respectively. An on-screen tap was required to reposition images within the selection bar.

Unfortunately, the accelerometer approach was far from elegant and was difficult to use. The primary problem with the prototype was that the entire device had to be positioned to move the selection bar. Consequently, the screen was often positioned away from the direct line of sight. Moreover, the screen would often reflect light, increasing the difficulty of viewing the screen. The problem was addressed in part by attuning the approach to be more sensitive to movement, reducing the need to manipulate the device dramatically.

However, such sensitivity was impractical, as the device would respond frequently to movements that were essentially indirect interactions. Furthermore, there was no guarantee that all users would interact with devices in the same way, certain users may make bigger movements. Therefore, the accelerometer approach was discarded. The approach would either not move the selection bar when expected or it would move the selection bar too much when not expected.

The remaining interaction approach was touch gestures. The initial implementation of the touch gesture approach presented the authentication mechanism with a selection bar. The user was expected to swipe up, down, left and right to move the selection bar and single-tap to manipulate images. An advantage of the implementation was that it was essentially the same as the Tetrad approach. Therefore, there is a more consistent approach to interaction across devices. However, consistent interaction is not necessarily an advantage if one device suffers with inefficient interaction.

Unfortunately, the initial implementation of touch gestures felt slow and awkward. The process felt cumbersome as interaction essentially consisted of manipulating a selection bar rather than images. Users wanting to manipulate a column of images needed to swipe the screen several times to position the selection bar over the desired column. Once the selection bar was positioned, a single tap was required to manipulate the images. Therefore, there was the possibility that users would have to perform several touch gestures to make small manipulations. The design of the interaction, although consistent with other prototypes, simply felt slow and awkward.

Therefore, a second implementation of touch gestures focused on manipulating images directly. The interaction was a break with early prototypes, as the selection bar was abandoned. The implementation presented the authentication mechanism and expected direct swipes on the columns and rows of images themselves. The second implementation relied on five gestures: swipe-up, swipe-down, swipe-left, swipe-right and double-tap. Therefore, if an individual wanted to manipulate images within a specific column, they simply swiped up or down on that column. Similarly, if an individual wanted to manipulate images within a specific row, they simply swiped left or right on that row. Lastly, an authentication attempt was submitted with a double-tap. A single-tap was avoided, in case the

user made an accidental tap. The approach was simple and clear.

The implementation relied on an internal iOS framework that essentially determined when one of the gestures had occurred. The specific details of the gesture, such as on-screen position, was used to determine what action should occur. Therefore, if a right swipe occurred, for example, the coordinates associated with the gesture was used to determine the row to manipulate.

The user experience was far superior, in terms of effort, than the original touch-gesture implementation, as users did not need to make several gestures to make a single manipulation. Moreover, no screen-space was wasted on additional interface objects. The entire screen of the smartphone was dedicated to grid display. Therefore, the second implementation of touch gestures was selected for use in the Tom prototype.

5.2.3 Registration Process

The implementation of the registration process comprised two screens or parts. The first screen was registration code entry and the second supported the creation of the graphical authentication secret.

The first screen was presented when the application was initiated for the first time. The view comprised a text field and a block of text. The block of text explained that a registration code was required to use the application and could be obtained from the School of Psychology. The text field was for entry of the actual registration code. When the user entered a registration code the client device confirmed the registration code was valid with the server. Therefore, the user had to have an active data connection to begin the registration process. If the code was invalid, the user was informed and asked to ensure it was entered correctly. If the registration code was valid then the registration code was deactivated on the server and the second screen presented.

The second screen focused on the creation of the graphical authentication secret. The user was presented with a vertical scroll view containing four pages, that they navigated with a vertical swipe. The user was expected to create a sequence of four images from an image set of 45. Each of the 45 images had an associated text label. The image labels combined to represent the user-generated image sequence as a string. An example would be if the user had selected images with labels: ‘A’, ‘K’, ‘X’ and ‘B’, in that order, then the resulting string would be AKXB. The string was then used to generate a hash value. The hash value was stored on the device for subsequent use in the authentication process. The user was expected to recreate the sequence of four images, generated in the registration process, during authentication.

5.2.4 Authentication Process

Three attempts to recreate the sequence were allowed, otherwise the account would be deactivated and the user would be required to re-register.

Tom presented the authentication mechanism directly after the registration process. The user was given instructions on how to use the mechanism; users tapped instructions to dismiss them. The instructions were only presented on the

first encounter and were not presented again, once dismissed. The user submitted an authentication attempt with a double-tap.

Upon entry of an authentication attempt, the system determined the coordinates of the first image in the image sequence. The system then determined a horizontal, vertical and diagonal image sequence using the first image as the start point. The system generated a string, using the labels associated with each image, for each image sequence. The strings were then used to generate a hash value. The hash value was compared with the one generated at registration. If the values matched, the user was authenticated and the authentication mechanism was dismissed. Otherwise the user could make another authentication attempt.

Therefore, the authentication process itself did not include a server-side component. The implementation was designed to allow the authentication mechanism to function without an active data connection. The design was intended to complement the portability of a smartphone in contrast to a desktop personal computer permanently connected to the Internet. Students may need to authenticate in environments where there is no connection to the Internet. Furthermore, students may encounter the situation where they are unable to afford to pay for an active data connection. An individual may have a pre-pay handset rather than post-pay; as funds dwindle students may prioritise and not pay for expensive data connections. Moreover, students may not be using an Apple iPhone but rather an Apple iPod touch. Consequently, while students may have connectivity in a coffee shop or on campus, they will not as they walk to the bus stop.

However, the details associated with the authentication attempt, such as the time taken were logged on the server side. The client device logged the information on the server after the user completed an authentication attempt. Therefore, an active data connection was required during authentication to log details of the attempt itself. Consequently, users were required to have an active data connection to use the authentication mechanism, otherwise a message was displayed that instructed users to check their Internet connection.

Unfortunately, there was still the concern that users may not always have a data connection. Therefore, the decision was taken to afford eight hours of use between authentication attempts. Therefore, once a user had successfully authenticated and the application was active, they would not be required to authenticate again for another eight hours.

5.2.5 User Identification

An alphanumeric string was used for user identification. The Universal Device Identifier (UDID) of the user's iOS device acted essentially as an individual's username.

5.2.6 Recovery Process

The authentication mechanism will not have a dedicated recovery process. The user will instead be expected to delete the application and subsequently install it again.

5.2.7 Application

Figure 5.6 illustrates an open lecture document within the application. The main elements of the interface are detailed below:

1. *Lecture Document Browser*
   The user taps the button to return to a view of the downloaded lecture recordings from the store. If audio is playing it will stop.

2. *Chapter List*
   The user taps the button to view the chapters associated with the currently opened lecture recording.

3. *Action Button*
   The user taps the button to perform an action. If the lecture recording is playing then a chapter is created at the specific time point. If the lecture audio is not playing, it will play the audio.

4. *Chapter*
   These coloured segments represent chapters. The user can change the colour to an image by attaching a specific image to the chapter, in the chapter list.

5. *Lecture Document View*
   The user can tap the tab to view the currently opened lecture document or a list of documents if none are open.

6. *Lecture Store View*
   The user can tap the tab to view a list of available lecture recordings for download. The user can view and download lectures will listening to a lecture.

The application was made available for distribution to students in the evaluation.

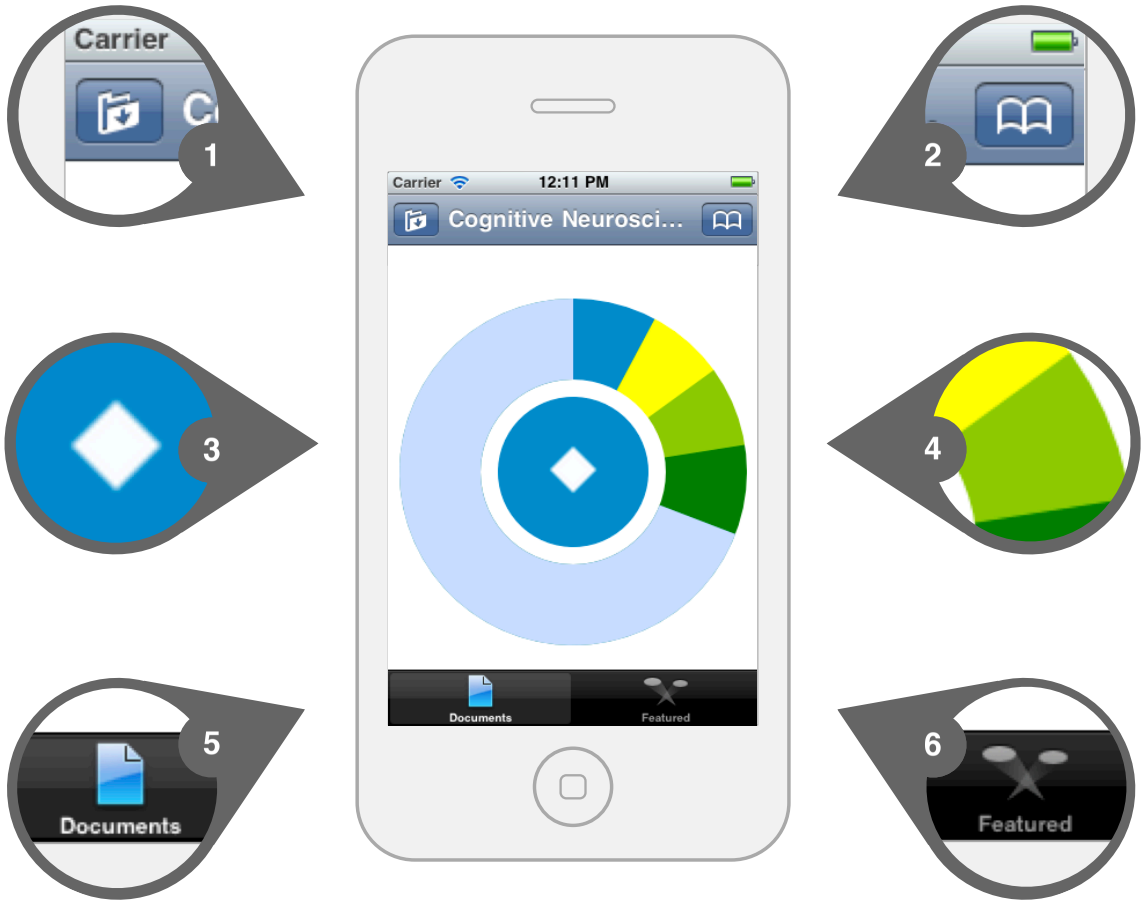

Figure 5.6: An overview of the main functions of the application.

5.3 Evaluation

The details of the evaluations are outlined over the following sections.

5.3.1 Subjects

The participants were enrolled students at the School of Psychology at the University of Glasgow. The application was distributed for free and any enrolled student was allowed to use it. The application itself did not collect any personal information. Therefore, it is difficult to characterise the user-base. However, estimations can be made about the cohort, based on the results of the survey.

5.3.2 Apparatus & Materials

Three technical elements were required for the application to function, namely: (1) a device to execute the application, (2) a server providing content and storing usage logs and (3) a data connection between the two.

Students were responsible for providing two of these elements, specifically a device to execute the application and a data connection. The application binary was designed for execution on iOS 4.0 devices, specifically Apple iPhone and iPod touch. The application could also execute on other iOS 4.0 devices, such an Apple iPad. Students were expected to own or have access to such devices. The application binary itself was distributed for free via the Apple App Store. The Apple App Store was available on all iOS 4.0 devices, as it is packaged with the operating system. Consequently, students were expected to download the application from the Apple App Store, requiring them to have a registered iTunes account.

Furthermore, students were required to have an active data connection to not only download the application but to use it. Moreover, users were completely responsible for managing all costs associated with data connections. Therefore, students choosing expensive cellular connections over inexpensive alternatives, such as on-campus wireless connections, did so at their own discretion. Students could not claim expenses for costs associated with the data connection. The data connection was necessary as the primarily link between the application and the server.

The server solution for the application was a mixture of Amazon Web Services. Amazon Simple Database was used to store usage information collected from the user's device. While Amazon Simple Storage Service was used to store and serve lecture recordings, as well as a listing of available lecture recordings. The listing was an XML file containing specific details about the lecture recordings, e.g. lecture title and lecturer.

The listing file was updated and uploaded to the server solution using a traditional personal computer. Lecture recordings were generated by the School of Psychology and delivered for upload. The School of Psychology provided high-quality audio recordings, MP3 encoded. The audio recordings were transcoded into Advanced Audio Coding (AAC) format before upload, using Apple iTunes on a traditional personal computer.

5.3.3 Procedure

The School of Psychology sent an email to all enrolled students, advising them about the immediate availability of the application. The main features of the application were communicated, namely users could download lecture recordings direct to their portable device as well as chapter and annotate them. Students were advised the application targeted iOS 4.0 devices, specifically Apple iPhone and iPod touch. Lastly, students were advised that a registration code was necessary to use the application.

Students were automatically issued registration codes via email. Students seeking a registration code either had the option of completing the survey or requesting a code using a web-based form. The student was expected to enter the registration code when the application requested it.

The application requested the registration code, upon first launch. The application binary was distributed, for free, via the Apple App Store. The email sent to students contained a hyperlink to the application binary in the store. The application could be downloaded by any registered iTunes user in the United Kingdom. However, without a registration code users could not actually use the application.

Users were presented the registration process, once they entered a valid registration code. Each registration code was limited to one-time use. The user was requested to complete the registration process to generate a graphical authentication secret. Once students had completed the registration process, they were requested to authenticate.

The graphical authentication mechanism was then presented. Moreover, an information panel overlaid the graphical authentication mechanism, detailing how to use it. The panel of information was presented to user the first time they used the authentication mechanism and not subsequently. Users dismissed the information panel with a single tap. Once users authenticated they could access the application for eight hours. Therefore, if they authenticated one day and accessed the application the following day they would be asked to enter their graphical authentication secret. Users were allowed three attempts to authenticate.

There was no reset or recovery procedure. If users were unable to authenticate, then they were required to delete the application and request another registration code. There was no limit on the number of times a user could register.

5.4 Results

The results are discussed in terms of registration process and the authentication process.

5.4.1 Registration

There were 63 registrations, over 168 days. The majority of these (76%) occurred in the first 30 days and few registrations (13%) occurred after 60 days.

The application did not collect personal information, such as an email address, or request individuals to create a username. The application instead relied on the

device UDID as the individual's username. The device UDID was associated with each registration. Consequently, the number of unique users can be estimated from the number of distinct UDIDs. The total number of distinct devices that executed the registration process was 45. Therefore, 45 individuals registered and generated an authentication secret.

However, because there was no recovery process, users unable to authenticate were expected to delete the application, download it and conduct the registration process again. Therefore, some users completed the registration several times. The majority of users though (64%) completed the registration process only once. Nevertheless, several users completed the registration process twice (31%) and some (5%) completed it three times. The registration process was not completed more than three times by any user.

The majority of registrations (60%) were initiated between midday and six o'clock in the evening. Indeed most registrations (68%) were initiated between midday and midnight. Nevertheless, some registrations (16%) occurred during the night between midnight and six o'clock in the morning. The remaining registrations (16%) occurred between six o'clock in the morning and midday. Furthermore, the majority of registrations (30%) were initiated on a Thursday. Indeed most registrations (90%) occurred on a weekday rather than at the weekend.

Image Selection

Students were expected to select four distinct images during registration, duplications were not allowed. Therefore, an individual could not select the same image twice in a single registration. The mean number of selections for each image was approximately six, over all registration attempts, with a median of 2 and mode of 1. However, the range in image selections was 35 with a minimum of 0 selections and maximum of 35 selections per image. The standard deviation for image selections, over all registration attempts was 9.16.

However, specific registration attempts were re-registrations, i.e. users conducted the registration process a second and third time as they were unable to complete authentication. Excluding subsequent registrations by the same device, the mean number of selections per image was 4 with a median and mode of 1. However, the range in image selections was 24 with minimum of 0 selections and maximum of 24 selections per image. The standard deviation for image selections, for initial registration attempts was 6.65.

Therefore, there was variance between images in terms of selections. Figure 5.7 illustrates the frequency distribution of image selections made during registration. Figure 5.7 depicts image selections from all registration attempts contrasted with image selections from initial registration attempts. A total of 8 images were never selected over all registration attempts, 10 when considering only initial registration attempts.

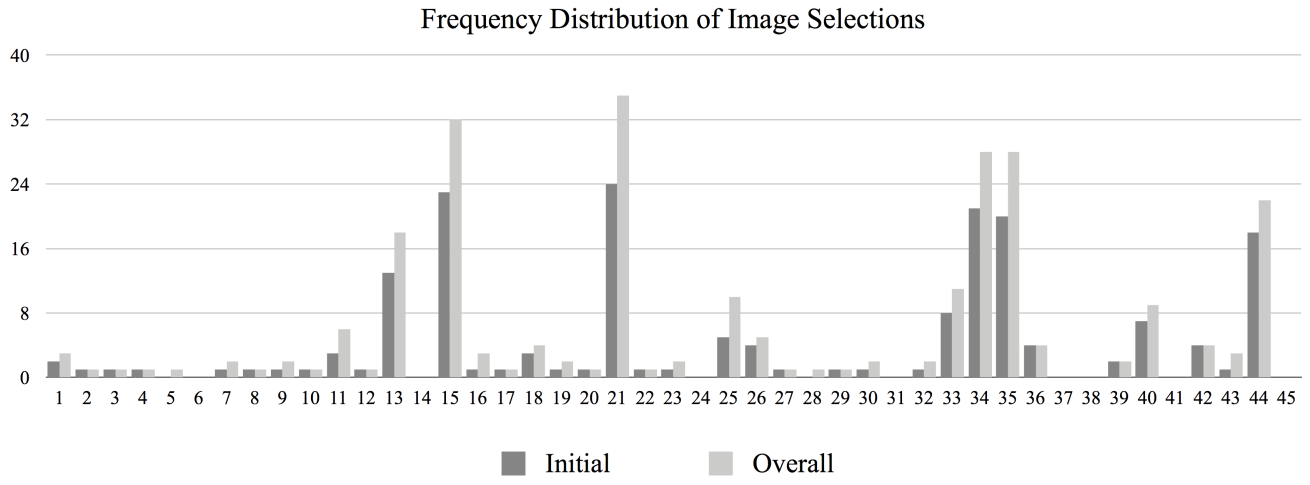

Figure 5.7: Frequency distribution of image selection for all registration attempts and initial registration attempts.

Rank	#	Staff Title	Times Selected		Position							
			Overall	Initial	Overall				Initial			
					1	2	3	4	1	2	3	4
1	21	Senior University Teacher	35	24	12	6	6	11	8	3	5	8
2	15	Teaching Assistant	32	23	10	10	4	8	8	5	4	6
3	34	Head of Department & Professor	28	21	9	4	7	8	8	3	5	5
4	35	Reader	28	20	7	10	7	4	6	8	3	3
5	44	University Teacher	22	18	5	6	6	5	4	5	4	5

Table 5.1: Five most popular image selections made during registration.

Time

The average time taken to complete the registration process overall was 312 seconds. However, there was wide variation in the time taken to complete registration (minimum = 25.17, maximum = 5041.94, range = 5016.78). The majority of users (93%) completed the registration process in less than 1000 seconds. The average time taken to complete the registration process for those below 1000 seconds was 163.23 seconds. However, there was still wide variation in the time taken to complete the registration process (minimum = 25.17, maximum = 763.62, range = 738.45). It is important consider that the registration process would be novel to most users and take time for some to adjust to the process.

The time taken to create an authentication secret is incredibly important as it represents a component of the overall cost of an authentication. The time spent on registration needs to be coupled with the time spent on authentication attempts to give a real reflection of the time associated with an authentication secret. If an individual spends 100 seconds on registration and then makes only two authentication attempts, then each authentication attempt is weighed down with an additional 50 seconds. The majority of laboratory investigations do not discuss registration cost; even if such costs are discussed it is hard to couple them with authentication attempts. The reason is that laboratory investigations typically rely on a controlled number of attempts at fixed intervals. The user is not given room to authenticate as much, and as often, as desired. Therefore, laboratory investigations can sidestep the actual cost of registration and ignore the fact that this cost has to be added to each authentication attempt.

A key concern is that users should be able to complete registration swiftly. Moreover, it is expected users may take time with novel registration. The expectation is that users can become better at registration and reduce the overall time taken to complete registration, as they become more familiar with the process. Consequently, a key research question is whether there is any difference in the time it takes a user to complete registration, specifically between the initial encounter and subsequent encounters with the registration process.

Therefore, a paired-samples t-test was used to determine if a statistically significant difference in mean time existed between participants re-registrations and initial registrations. The data exhibited no outliers, as determined using a box plot. The data was normally distributed for both re-registrations and initial assessments, as determined by a Shapiro-Wilks test (p=.741). The reality is that participants completed re-registrations (M = 116.95, SD = 44.32) quicker than initial registrations (M = 153.83, SD = 62.48). The mean difference was a statistically significant ($p < .05$) decrease in time of 36.87s, 95% CI [-70.9417,-2.8127], $t(13)$= -2.339, p=0.036, d= -0.63.

5.4.2 Authentication

Users made a total of 387 authentication attempts over a period of 339 days. The majority of authentication attempts (58%) ended in success. The remaining (42%) were unsuccessful.

A total of 46 users made at least one authentication attempt. The average number of authentication attempts was 8.41 (SD = 6.235). However, many users made far fewer authentication attempts (mode = 3). The number of authentication attempts generated by users was mixed (minimum = 1, maximum = 30, range = 29), suggesting that the distribution of authentication attempts among users was more varied than indicated by the mean. Figure 5.8 illustrates users grouped by the number of attempts they generated. The majority of users (33%) made between 1 and 3 authentication attempts.

Furthermore, although all users made at least one authentication attempt (N = 46), not all attempts were successful. The average number of successful authentication attempts was 4.91 (SD = 5.349). However, the mean may not be indicative of actual performance for most users (mode = 0). There was a large span in the number of successful authentication attempts (minimum = 0, maximum = 22, range = 22). Figure 5.9 illustrates the number of successful and unsuccessful authentication attempts made by individuals. Figure 5.9 serves to emphasise that several users (20%) did not make a single successful authentication attempt.

The average number of unsuccessful authentication attempts was 3.50 (SD = 2.438). The mean may be indicative of overall performance (median = 3, mode = 3) as comparatively there was not wide variation in the number of unsuccessful attempts for most users (minimum = 0, maximum = 10, range = 10). However, a small number of users (11%) made no unsuccessful authentication attempts. Nevertheless, the majority of such users (80%) only made a single authentication attempt. The reality is that most users (89%) made at least a single unsuccessful authentication attempt. Moreover, the vast majority of users (74%) made 3 or

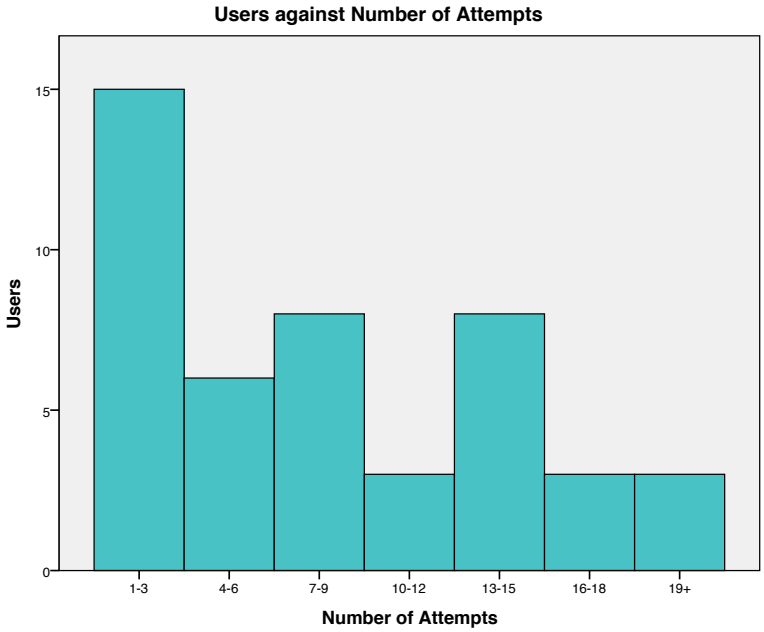

Figure 5.8: Users grouped by the number of attempts they generated.

more unsuccessful authentication attempts.

Users were constrained in that each registration or account was limited to 3 sequential unsuccessful authentication attempts. The account was deactivated once the limit was reached. Users had to re-register to use the application. Figure 5.10 illustrates authentication attempts in terms of registrations. Users completed a total of 63 registrations ($N = 45$), all individuals were accounted for in authentication attempts, asides one. Therefore, authentication attempts are associated with an individual that is not accounted for in registration logs. There are several possible explanations for the situation, although the most likely is that the registration information was simply not logged on the server. The individual generated 4 authentication attempts. Consequently, a total of 383 authentication attempts were generated across all registrations.

The average number of authentication attempts for each registration or account was 6.06 ($SD = 5.217$). However, many accounts generated far fewer authentication attempts (mode = 3). The average number of successful authentication attempts for each registration or account was 3.52 ($SD = 5.060$). Similarly, many accounts generated far fewer successful authentication attempts (mode = 0). The average number of unsuccessful authentication attempts for each account or registration was 2.54 ($SD = 1.803$). The mean may be indicative of overall performance as several accounts generated unsuccessful attempts (mode = 3).

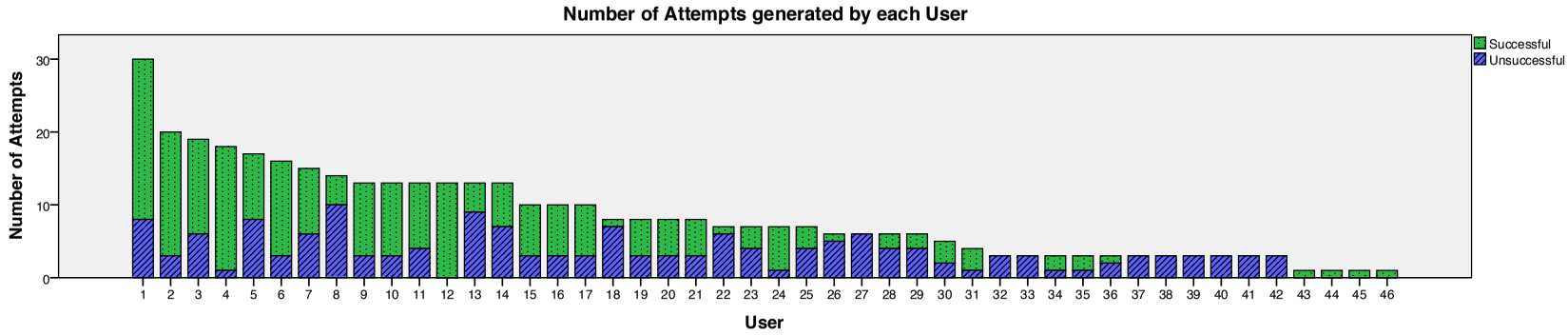

Figure 5.9: The number of successful and unsuccessful authentication attempts for each user.

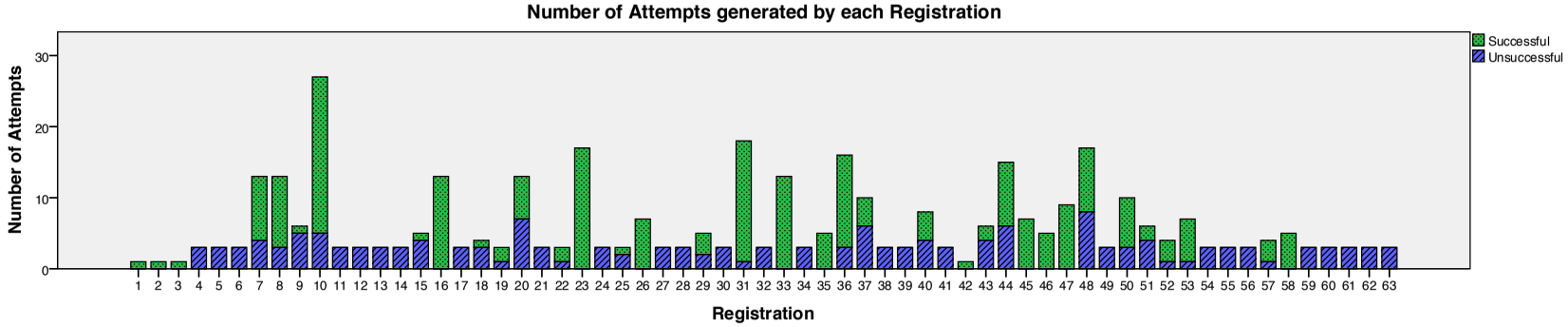

Figure 5.10: The number of successful and unsuccessful authentication attempts for each registration.

The position and occurrence of unsuccessful authentication attempts is important as 3 sequential authentication attempts resulted in an account being deactivated. Figure 5.11 illustrates the occurrence of successful and unsuccessful authentication attempts with each registration or account. The majority of accounts (56%) were deactivated while the surviving accounts (44%) remained active.

The deactivated accounts can be characterised in terms of position of the unsuccessful authentication attempts. The majority of deactivated accounts (49%) represented *false starts* by users, i.e. the user's initial account was deactivated but the user created another account. Therefore, and the deactivated account did not deter the individual, instead they tried again, re-registered. However, this was not always the case: there were several deactivated accounts (23%) that represented *casualties of authentication*. Such users could not make a successful authentication attempt and simply abandoned the application entirely. The application lost users due to authentication. There were also several deactivated accounts (23%) that represented more traditional authentication failure, i.e. users generated at least one successful authentication attempt but subsequently could not authenticate and the account was deactivated. These deactivated accounts can be characterised with a *cap* of unsuccessful authentication attempts, e.g. Figure 5.11 account 36. The 'cap' should be no longer than 3 unsuccessful authentication attempts but this is not case for some deactivated accounts, e.g. Figure 5.11 account 9 and 37, have lengthier sequences of unsuccessful authentication attempts. The presence of these lengthier sequences suggests users were able to circumvent the limit or the application contained flaws. The remainder of the deactivated accounts represented a user who had re-registered but did not submit a single successful authentication attempt. Consequently, several deactivated accounts (29%) represented users who did not gain access to the application.

Therefore, deactivated accounts essentially represent users who failed to learn the novel authentication approach or were simply frustrated by it. Consequently, the majority of unsuccessful authentication attempts (75%) were associated with deactivated accounts. The remaining unsuccessful authentication attempts (25%) were associated with active accounts.

Furthermore, the majority of successful authentication attempts (80%) were associated with active accounts. There are several active accounts (46%) that have no unsuccessful attempts associated with them. Moreover, many of these unblemished accounts (69%) consist of several successful authentication attempts, as illustrated in Figure 5.11. The vast majority (89%) of these unblemished accounts that contain several successful authentication attempts are from users who initially made a false start. Therefore, the majority of these exemplar accounts are from users who made disastrous false starts.

The remaining aspect to consider about successful authentication attempts is the direction. The target images could have been aligned either horizontally, vertically or diagonally. The majority of successful authentication attempts (80%) were horizontal. The average number of horizontal authentication attempts for each registration or account was 2.81 ($SD = 4.782$). The remaining authentication attempts (20%) were vertical, no diagonal authentication attempts were submitted by users. The average number of vertical authentication attempts for

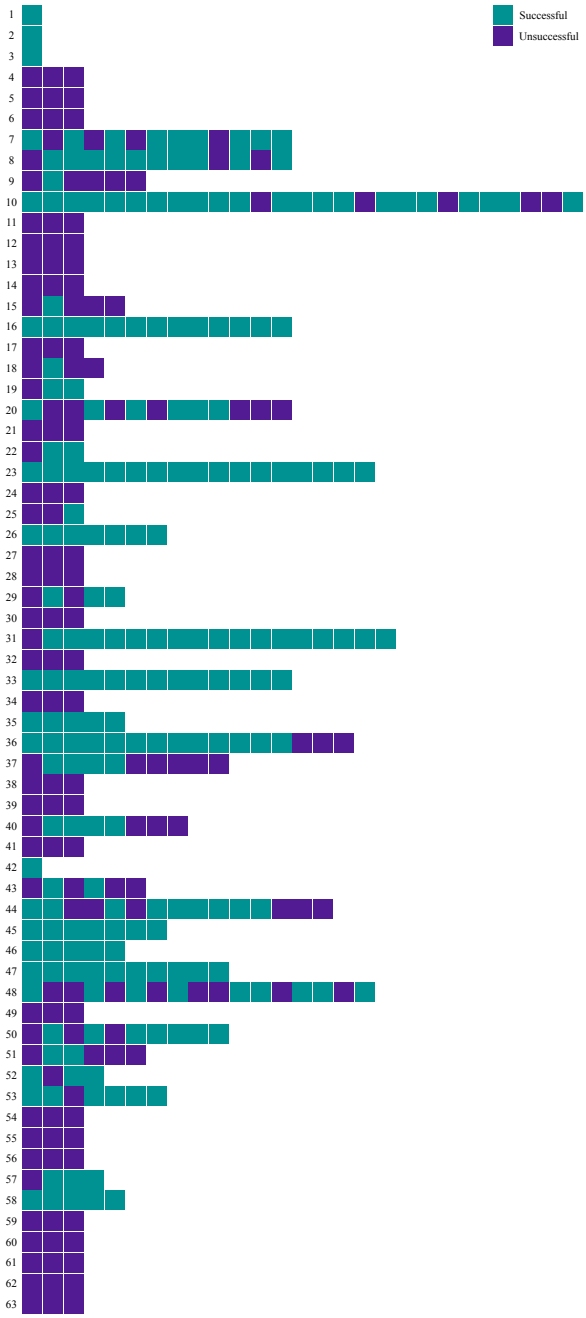

Figure 5.11: The successful and unsuccessful authentication attempts associated with each registration or account.

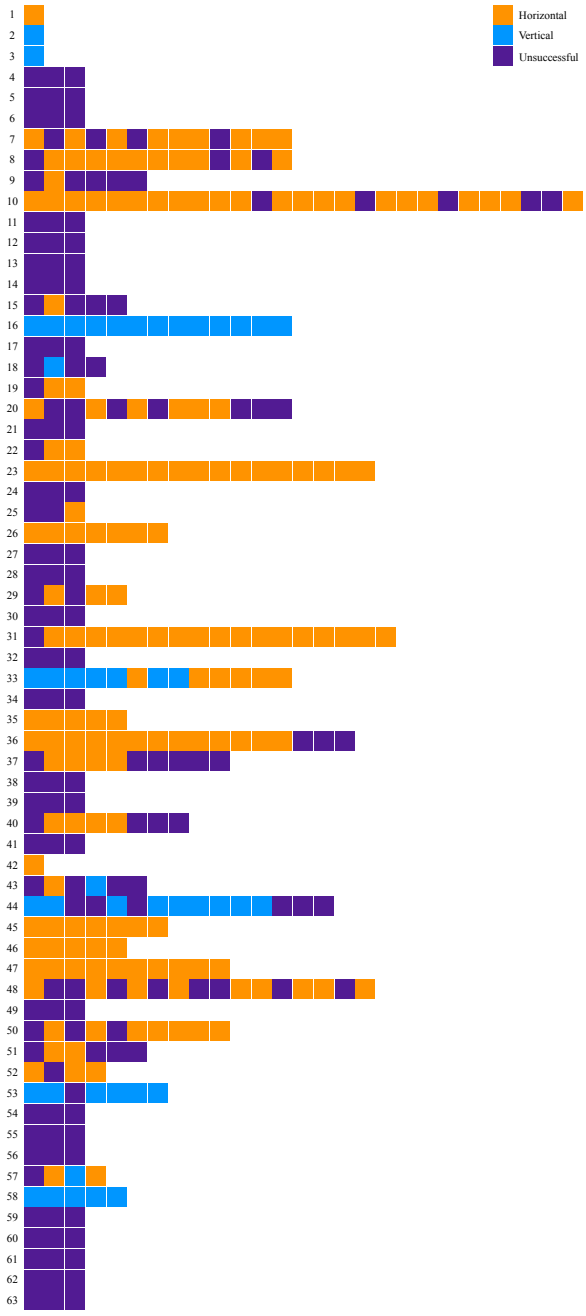

Figure 5.12: The horizontal and vertical authentication attempts associated with each registration or account.

each registration or account was 0.71 ($SD = 2.317$). The reality is that users overwhelmingly submitted horizontal authentication attempts.

Time

Users spent a total of 423671.55 seconds on 387 authentication attempts or 4.9 days. The average time for an authentication attempt was 1094.76 seconds ($SD = 14005.335$). However, there was a wide variation in the time taken on each authentication attempt (minimum = 1.43, maximum = 266919.88, range = 266918.44). There was several authentication attempts that took considerable time.

Therefore, if we limit authentication time to 200 seconds, then only a small number of authentication attempts (5%) are removed from consideration. The majority of the discarded authentication attempts (72%) were successful authentication attempts. The average time for the remaining authentication attempts was 36.66 seconds ($SD = 28.24$). However, there is still wide variation in the time taken for each authentication attempt (minimum = 1.43, maximum = 194.31, range = 192.87). Table 5.4.2 outlines descriptive statistics for the time taken to complete an authentication attempt from various different perspectives.

The time taken to complete an authentication attempt is reflective of the use of the authentication approach in a real world task. The wide range in authentication times is potentially explained by users gaining familiarity with a novel authentication approach and then becoming increasingly better at the authentication approach. The concern with laboratory experiments is that users typically complete an authentication attempt at fixed intervals disconnected from the stress of completing an actual task. The controlled experiments do not necessarily reflect the cost of an authentication attempt in actual use with an actual application. The user is arguably motivated in a real world setting to become superior at the authentication approach to complete the task more efficiently.

Therefore, another key research question is whether the individual is able to improve or reduce the time taken to complete an authentication attempt as they become more adept. The expectation would be that initial authentication attempts would be lengthy due to the novelty of the authentication approach but then rapidly improve as users become more familiar with the authentication approach.

Therefore, a repeated measures analysis of variance (ANOVA) was performed to determine if there was any statistical significant difference in the time taken for users to complete their first, second and third authentication attempts. Only active accounts with three authentication attempts below 200 seconds were considered. Figure 5.13 illustrates box plots for each of the three groups of authentication attempts.

The first and third groups appear normally distributed. However, the second group of authentication attempt times appears to have one case (18) that is more than 3 box-lengths from the edge of its box, labelled in Figure 5.13 with an asterisk. A Shapiro-Wilk's test suggests the second group of authentication attempt times is indeed not normally distributed ($p < 0.05$). There are many different approaches for dealing with outliers. The decision was taken to simply remove the offending case from consideration. A subsequent Shapiro-Wilk's test suggested

	Attempts	Mean	SD	Minimum	Maximum	Range
Successful	226	671.68	4632.666	4.82	52258.76	52253.94
Unsuccessful	161	1688.64	21033.839	1.43	266919.88	266918.44
Successful (<200s)	213	45.02	27.975	4.82	194.31	189.49
Unsuccessful (<200s)	156	25.25	24.406	1.43	168.56	167.12
Horizontal	180	678.10	356.972	4.82	52258.76	52253.94
Vertical	46	646.59	4008.495	11.28	27240.56	27229.28
Horizontal (<200s)	170	44.18	25.619	4.82	154.75	149.93
Vertical (<200s)	43	48.34	35.990	11.28	194.31	183.03

Table 5.2: Descriptive statistics for length of time for authentication attempts.

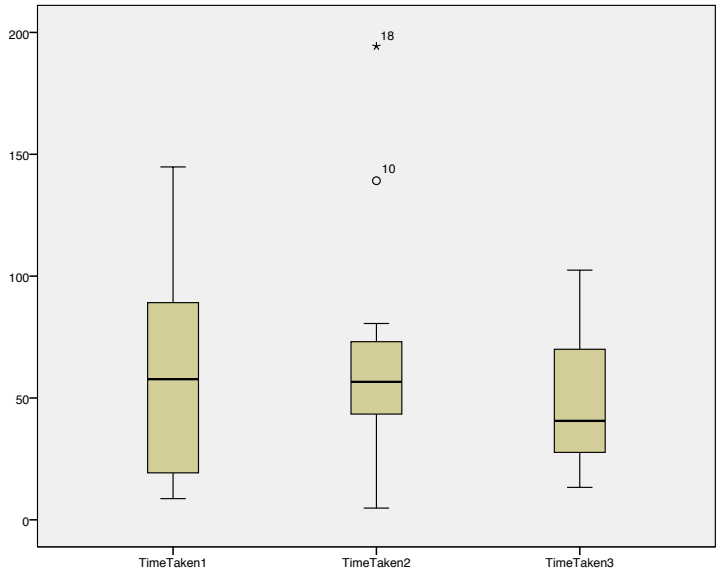

Figure 5.13: Box plots of time taken for first, second and third authentication attempts by users.

that second group of authentication attempts as well as the first and third group were normally distributed ($p > 0.05$). The length of time for each authentication attempt increased from the first attempt ($M = 56.77$, $SD = 37.718$ seconds) to the second attempt ($M = 57.35$, $SD = 27.684$ seconds) but then decreased by the third attempt ($M = 50.02$, $SD = 27.764$ seconds). The assumption of sphericity was not violated as confirmed with a Mauchly's Test of Sphericity, $\chi^2(2) = 2.184$,$p = 0.335$. However, there was no statistically significant difference in the time taken to complete each authentication attempt, $F(2,36) = 0.236$, $p = 0.791$.

Therefore, the time taken to complete an authentication attempt did not appear to improve significantly as they completed more attempts. However, such improvement may be difficult to gauge over the initial three authentication attempts, considering the initial six authentication attempts may reveal improvements. Therefore, a repeated measures ANOVA was performed to determine if there was any statistically significant difference in the time taken to complete the first, second, third, fourth, fifth and sixth authentication attempts. Only active accounts with six authentication attempts below 200 seconds were considered.

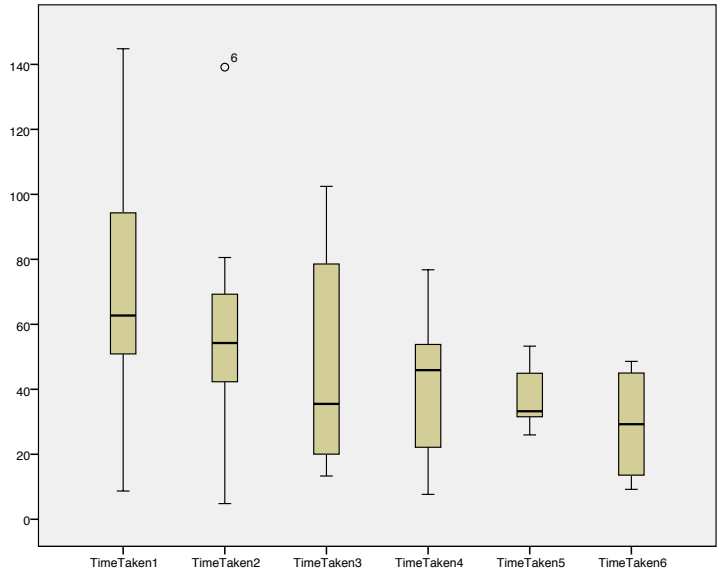

Figure 5.14: Box plots of time taken for first, second, third, fourth, fifth and sixth authentication attempts by users.

Figure 5.14 illustrates box plots for each of the groups of authentication attempt times. The groups all appear normally distributed asides from a case (6) that is more than 1.5 box-lengths from the edge of its box, as labelled in Figure 5.14 with a circle. The case does not represent an extreme outlier, consequently will remain in the group for consideration. A Shapiro-Wilk's test suggests all six groups are normally distributed ($p > 0.05$). Figure 5.14 illustrates a trend that authentication attempts time appear to decrease with each successive attempt. The length of time decreased from the first authentication attempt ($M = 68.90$, $SD = 39.208$ seconds) to the second attempt ($M = 58.97$, $SD = 34.899$ seconds) and continued to decrease by the third attempt ($M = 46.72$, $SD = 31.599$

seconds). The trend persisted as the time decreased further for the fourth attempt($M = 41.64$, $SD = 23.056$) and then for the fifth attempt ($M = 37.31$, $SD = 9.466$) and still decreased by the sixth authentication attempt ($M = 28.82$, $SD = 14.543$). The assumption of sphericity was violated as indicated by Mauchly's Test of Sphericity, $\chi^2(14) = 31.590, p < 0.05$. Consequently, a Greenhouse-Geisser correction was applied ($\epsilon = 0.520$). While the trend may have appeared that authentication attempt times were improving with each successive attempt there were no statistically significant differences in the times taken to complete successive authentication attempts, $F(2.600,23.401) = 2.522$, $p = 0.090$.

However, there may be improvement over time with successful authentication attempts as users become increasingly familiar and successful with the authentication approach. Therefore, a repeated measures ANOVA was conducted to determine if there was a statistically significant difference in the time taken for users to complete their first, second and third successful authentication attempts. Only active accounts with three initial, sequential successful authentication attempts below 200 seconds were considered.

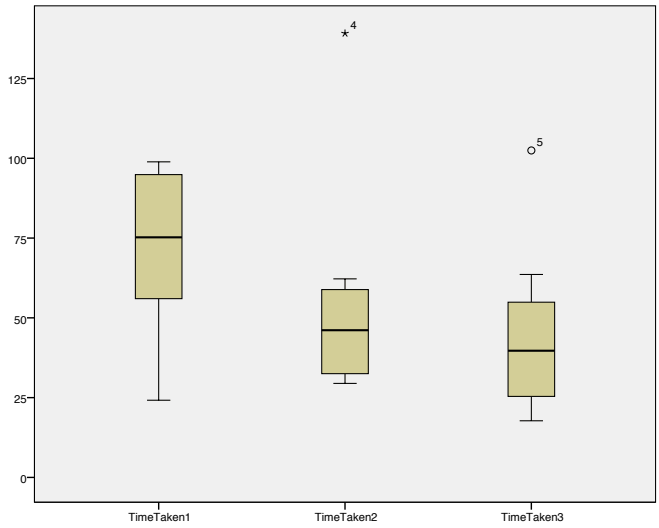

Figure 5.15: Box plots of time taken for first, second and third successful authentication attempts by users.

Figure 5.15 illustrates box plots for each group of authentication attempt times. The first and third group appear normally distributed. The second group contains a case (4) that is 3 box-lengths from the edge of its box, as labelled in Figure 5.15 with an asterisk. Consequently, the case was classed as an extreme outlier and removed from consideration. A subsequent Shapiro-Wilk's test confirmed that all three groups were normally distributed ($p > 0.05$). Figure 5.15 illustrates a trend that the time taken to complete a successful authentication attempt appears to decrease from the initial attempt but not continue. The time taken to complete a successful authentication attempt decreased from the first attempt ($M = 78.74$, $SD = 18.89$ seconds) to the second attempt ($M = 43.49$, $SD = 12.688$ seconds) but then increased by the third attempt ($M = 46.98$, SD

= 29.223 seconds). The assumption of sphericity had not been contravened as suggested by Mauchly's Test of Sphericity, $\chi^2(2) = 3.713$, $p = 0.156$. The time taken for successful authentication attempt was statistically significant depending on the attempt, $F(2,12) = 4.143$, $p < 0.05$, partial $\eta^2 = 0.408$. Post-hoc analysis with Bonferroni adjustment displayed a statistically significant decrease from the first attempt to the second attempt (35.24 (95% CI, 10.30 to 60.18) seconds, $p < 0.05$).

However, the difference is not particularly surprising as initial authentication attempts are likely to be higher than others as it will be the first encounter of the authentication approach for many users. Furthermore, the sample used for the statistical analysis was particularly small ($N = 7$). Therefore, the weight or relevancy of results are debatable. It may be more prudent to consider the time taken on unsuccessful attempts for that there is a larger sample ($N = 24$).

Consequently, a repeated measures ANOVA was conducted to determine if there was a statistically significant difference in the time taken for users to complete their first, second and third unsuccessful authentication attempts. Only inactive accounts with three initial, sequential unsuccessful authentication attempts were considered.

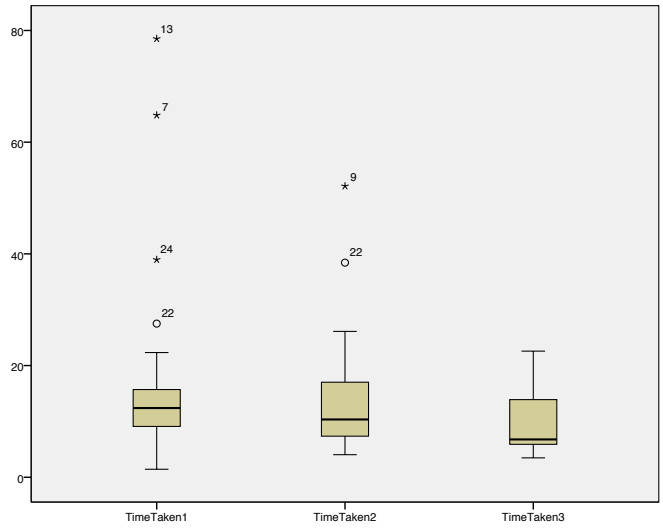

Figure 5.16: Box plots of time taken for first, second and third unsuccessful authentication attempts by users.

Figure 5.16 illustrates box plots for each group of authentication attempt times. The first and second group appear to contain several cases that are more than 3 box lengths away from the edge of their respective box. A Shapiro-Wilk's test confirms that all three groups were not normally distributed ($p > 0.05$). Therefore, the respective outliers were removed and subsequent Shapiro-Wilk's test confirmed that the first and second groups were normally distributed ($p >0.05$) but not the third group ($p < 0.05$). However, given the third group contained no extreme outliers and the impact is debatable the analysis continued with all three groups. Figure 5.16 illustrates a trend of the time taken on unsuc-

cessful decreasing by the third attempt. The time taken for an authentication attempt increased from the first attempt ($M = 10.34$, $SD = 4.980$ seconds) to the second attempt ($M = 13.42$, $SD = 6.639$ seconds) but then decreased by the third attempt ($M = 7.46$, $SD = 3.136$ seconds). The assumption of sphericity had not been violated as suggested by Mauchly's Test of Sphericity, $\chi^2(2) = 2.749$, $p = 0.253$. The length of time for an unsuccessful authentication attempt was statistically significant depending on the attempt, $F(2,30) = 4.506$, $p < 0.05$, partial $\eta^2 = 0.231$. Post-hoc analysis with Bonferroni adjustment displayed there a statistically significant decrease from the second attempt to the third attempt (5.97 (95% CI, 0.61 to 11.32) seconds, $p < 0.05$).

Therefore, it would appear that in terms of unsuccessful authentication attempts, it is possible for users to improve authentication times. However, several cases were removed from consideration and the last group was not normally distributed, arguably undermining the strength of the statistical analysis. Nevertheless, users do appear to reduce the time taken to authenticate to around 7.46 seconds, on average. However, the key aspect is that users did not actually authenticate, they did indeed fail to authenticate. The user was unsuccessful in gaining access to the application. Furthermore, 7.46 seconds for an unsuccessful authentication attempt is considerably lower than 43.49 seconds for a successful authentication attempt.

Therefore, it would appear that successful and unsuccessful authentication attempts are fundamentally different. Naturally, these authentication attempts are fundamentally different in conclusion but the process and the time taken by users should be similar as the user is striving to succeed and not fail. Ideally, there should be no difference between successful and unsuccessful authentication attempts in terms of time: only the conclusion should be different. Therefore, another key research question is whether an authentication attempt exhibits a difference in time depending on whether it is successful and whether it is the first, second or third attempt.

Therefore, a within-within ANOVA was conducted to determine if there was any interaction between the success of an authentication attempt and whether it was the first, second or third attempt. A Shapiro-Wilk's test on residuals indicated that groups of times for the first, second, fourth and sixth authentication attempts were normally distributed ($p > 0.05$) but not for the third and fourth authentication attempts ($p < 0.05$). However, considering the majority of groups were normally distributed and the impact was considered minimal the analysis continued with no cases transformed or discarded. The time taken to complete a successful authentication attempt increased from the first attempt ($M = 61.45$, $SD = 23.742$) to the second attempt ($M = 65.01$,$SD = 37.053$) but decreased by the third attempt ($M = 45.29$, $SD = 27.353$). The time taken to complete a unsuccessful authentication attempt increased from the first attempt ($M = 9.53$,$SD = 3.966$) to the second attempt ($M = 14.08$,$SD = 5.531$) but decreased by the third attempt ($M = 6.07$,$SD = 2.150$). The assumption of sphericity was not violated as suggested by Mauchly's Test of Sphericity, $\chi^2(2) = 0.882$, $p = 0.644$. There was no statistically significant interaction in the success of authentication attempt and attempt number on the time taken, $F(2,12) = 0.306$, $p = 0.742$. However, the main effect of the success of an authentication attempt

suggested a statistically significant difference on time taken, $F(1,6) = 440.952$, $p < 0.05$, $\eta^2 = 0.987$.

Therefore, there was a difference on time taken between successful and unsuccessful authentication attempts. However, the sample used in the analysis was small ($N = 7$). Therefore, the weight of any results are debatable. Nevertheless, the distinction in time taken between such authentication attempts is curious as the process should be similar only the outcomes different. The difference suggests that unsuccessful authentication attempt represent something other than simply not being able to remember the authentication secret.

However, there is another aspect to consider of successful authentication attempts: the alignment selected by users. The expectation is that users would favour the most convenient alignment, i.e. the alignment that affordd rapid entry of an authentication secret. The reality is that one alignment, horizontal, was used considerably more than either vertical or even diagonal. The diagonal alignment was not used in a single authentication attempt and the vertical alignment was used rarely in comparison to the horizontal alignment. Therefore, the expectation is that users make authentication decisions that favour convenience. Therefore, another key research question is whether there is any difference in the time taken to authenticate with different alignments.

An independent-sample t-test was conducted to determine any mean difference between horizontal and vertical attempts. Diagonal attempts were not considered, as not one single authentication attempt used the diagonal alignment. Few vertical vertical attempts were made, making a comparison with horizontal authentication attempts difficult. Consequently, only initial vertical authentication attempts were compared with initial horizontal authentication attempts ($N = 14$). A Shapiro-Wilk's test suggested that both groups were normally distributed ($p > 0.05$). The time taken for horizontal authentication attempts ($M = 73.86$, $SD = 43.146$) was shorter than vertical authentication attempts ($M = 97.741$, $SD = 59.400$). There was homogeneity of variances for authentication times for both horizontal and vertical groups, as determined by Levene's Test for Equality of Variances ($p = 0.972$). The reality is that vertical authentication attempts took 23.88 ($SE = 27.749$) seconds longer than horizontal authentication attempts. However, the difference was not statistically significant, $t(12) = 0.860$, $p = 0.406$.

Therefore, while vertical authentication attempts exhibit lengthier authentication times, the difference is not significant. However, few users even experimented with the different alignments. Therefore, even if the difference was statically significant it would be debatable if that is why users favoured horizontal alignment. The majority of users authenticated using a horizontally alignment with very little experimentation. It is unclear why users favoured one alignment over all others.

Memorability

The memorability of the graphical authentication secret can be uncovered by determining the number of days between the creation and last submission of an authentication secret. Therefore, if we consider only accounts or registrations with no duplicates and have at least one successful attempt we have 36 accounts.

The average number of days for an authentication secret was 39.59 days (SD = 61.583). However, the average may not be indicative of overall performance as many authentication secrets were not remembered as long (mode = 0). There was much variation in the memorability of authentication secrets (minimum = 0, maximum = 333, range = 333). The was several users could not remember (33%) of authentication were not memorability beyond a day.

However, this can be contrasted with several users (39%) that could remember the authentication secret beyond 30 days. Indeed specific users (22%) were able to recall the authentication secret beyond 60 days.

There may be a correlation between the memorability in days and the time taken to register. The accounts with registration time below 100 second and memorability above 0 days but below 200 days was considered. A Shapiro-Wilk's test suggested the data was not normally distributed ($p < 0.05$). Consequently, registration time below 200 second and memorability above 0 days but below 100 were considered. This left 15 accounts to consider, a Shapiro-Wilk's test confirms the data is normally distributed ($p > 0.05$). While there was a small positive correlation between the time taken and the number of days in terms of memorability, $r = 0.106$, the correlation was not significant, $p = 0.708$.

However, there may be a correlation between the memorability in days and the time it taken to register for active accounts. The registration time was limited to 1000 seconds for registration time and memorability above 0 days but less than 300 days. A Shapiro-Wilk's test suggests the data is normally distributed. While there was a slight negative correlation between the time taken and the number of days in terms of memorability, r = -0.101, the correlation was not significant.

There may be a correlation between the number of successful authentication attempts and registration time. A Pearson correlation was performed on 28 active accounts. However, registration times greater than 1000 seconds were not considered, leaving 26 cases. A Shapiro-Wilk test reveals data not normally distributed for successful authentication attempts ($p > 0.05$). Therefore, those cases removed where there was only one successful authentication attempt had been made and all those who made more than 15 successful authentication attempts. A Shapiro-Wilk test suggests the data is normally distributed for both registration time and successful authentication attempts ($p < 0.05$). There was not statistically significant correlation between the time taken at registration and the number of successful authentication attempts, $r(15) = 0.183$, $p = 0.483$.

If we only consider active accounts. The average number of days in terms of memorability for an authentication secret was 43.22 days (SD = 66.241). However, many authentication secrets were not remembered beyond a day (mode = 0). Indeed there was variation between memorability of authentication secrets (minimum = 0, maximum = 333, range = 333).

However, the last area of interest, specifically in the wild was the user expectations. Specifically, those users that have made a successful authentication attempt and have subsequently returned and been unable to authenticate. The user expects to be able to authenticate, i.e. assumes they remember their authentication secret. The average number of days between the successful and subsequent last entry was 41.07 days (SD = 49.842). The mean may not be indicative of the typical performance though (minimum = 0, maximum = 142,

range = 142). However, the majority of users (60%) attempted to gain access after 30 days. There were also attempts at 116 days and 142 days. The reality is that users expected these great gaps in memorability.

5.5 Discussion

The purpose of the application was to evaluate an alternative authentication mechanism within its expected context, i.e. as part of a realistic workflow, executable on an actual device, targeted at intended users. The application was successfully accepted into the Apple iTunes Store and was available to all students with an iOS device. Consequently, the application went through the same process of scrutiny, assessment and deployment as any other application available for distribution to consumers.

The application laid the foundation for the evaluation of the authentication mechanism in the field. There were no set intervals, no set requirements; students could easily abandon the application if they were not impressed. The application was initially well received as several students downloaded the application. The entire pool of students were never expected to download and use the application as sophisticated mobile computers, e.g. smartphones, were not particularly mainstream at the time. The expectation was that a small but sizeable base of students would use the application. The application was deployed with expectations but no guarantees.

However, much like a real-world product it was open to criticism, comment and discussion. Unfortunately, the reality is that while the application was relatively well received the authentication mechanism was not. The anecdotal evidence was that students generally appreciated the application and the ability to download lectures to their portable computer. Moreover, some students felt the application would be a useful addition to many other courses and felt the application should be deployed elsewhere. However, there were no positive comments surrounding the authentication mechanism.

The criticism directed at the authentication mechanism took two forms. The first was that some students felt the authentication mechanism was unnecessary. It was felt that an application that delivered lecture recordings to an individual's mobile computer should be free from the shackles of authentication. Authentication seemed gratuitous and some users of the application obviously felt the authentication mechanism was an unrealistic addition.

However, such an argument can be contrasted with the requirements laid down by the content owners. The School of Psychology was adamant that the application contained an authentication mechanism, during and after the evaluation period. Therefore, determining a realistic use of an authentication mechanism is difficult to define. There were users who felt the authentication mechanism was unnecessary and detracted from the application as a whole, while the client felt the application could not exist without an authentication mechanism. The second form of criticism was that the authentication mechanism was simply awful. There were some anecdotal comments that suggested the graphical authentication was not well received.

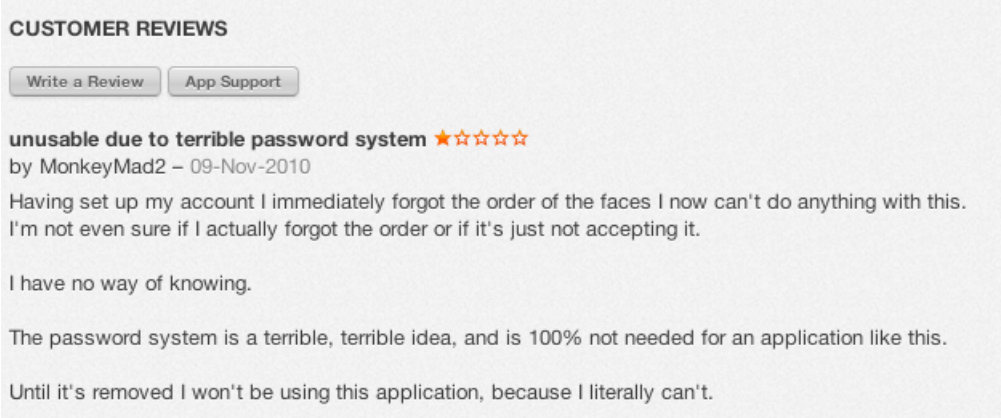

Figure 5.17: Review of the application from an individual in the iTunes Store.

Figure 5.17 illustrates a review from the iTunes Store that a student posted after they had downloaded and used the application. Figure 5.17 encapsulates both forms of criticism. The comments are visceral and emphasise that the authentication mechanism essentially drove the reviewer away. However, the comments about the alternative authentication mechanism are not surprising. The authentication approach was novel to user. The novelty of an authentication mechanism has many connotations, not all positive.

The aforementioned review provides some insight into these negative connotations. The reviewer is arguably distrusting of the authentication mechanism. The reviewer is not convinced that they entered the authentication secret incorrectly or whether the authentication mechanism itself processed it incorrectly. Alternatively, the comment does make a valid user interface remark that the user is unable to make corrections as they are not sure what they have entered. Therefore, the user is not sure what the system actually assessed. In an environment where exploration is limited by the number of unsuccessful authentication attempts, there is little room for learning.

Furthermore, the last remark in the review is that the reviewer would not use the application until the authentication mechanism was removed. The remark makes another essential point that not only did the reviewer think the authentication mechanism was expendable but that it was experimental. The reviewer was not going to waste time dealing with an experimental authentication mechanism, as something untested would not be used to protect something valuable.

The user base was intelligent and arguably did not appreciate being used as guinea pigs to evaluate an authentication mechanism to access valuable learning resources. However, in fairness, students were not forced to use the application and could use the portal to access lecture recordings: a portal that also required a password. Furthermore, students were not being treated differently from those in other schools, as no other school within the university had such an application, indeed few even made lecture recordings available to students.

Nevertheless, the authentication mechanism clearly alienated some students. The reality is that several users (23%) abandoned the application after authen-

tication and did not use it. Therefore, it many respects the alternative authentication mechanism was a failure from the outset. It frustrated too many, drove too many away. The argument could be made that it was not the concept but the implementation of the authentication mechanism that alienated users.

However, the application itself was well received, anecdotal evidence suggested that those students who accessed the application appreciated the quality. Furthermore, the School of Psychology were content with the application and keen to continue to use it. Moreover, the Apple engineer who assessed the application, approved it for wide-spread use. Therefore, it is unlikely those who abandoned the application did so because of the implementation. The reality is that the alternative authentication mechanism produced many casualties. However, the remaining students that soldiered on can be used to assess user performance.

The time taken to complete an authentication attempt is a key area of performance assessment. The expectation was that as the authentication mechanism was novel, initial authentication attempts may take time but the user would be able to reduce time with successive attempts. The time taken to complete an authentication attempt is an important consideration in determining the convenience of an authentication mechanism.

The statistical analysis did reveal a statistically significant difference between the initial and second authentication attempts submitted by users. However, this, in itself, is not particularly surprising as it was expected that users would initially generate lengthy authentication attempts as the authentication approach was novel. Users would have taken time to learn how to use the authentication approach. Therefore, it is not surprising that a subsequent second authentication attempt reduced the time taken. The assumption was that the time taken would improve with subsequent attempts. Unfortunately, while there was certainly a trend in that users appeared to take less time with successive attempts but there was nothing conclusive.

Furthermore, much of the analysis required the pruning of data to ensure it was normally distributed. There was wild oscillation in the time taken to complete an authentication attempt, some authentication attempts took only a few seconds while others took considerably longer to complete.

These wild oscillations can be explained by the design of the authentication mechanism, specifically from two aspects. The first aspect was the design of the authentication mechanism itself. The authentication mechanism was a *search-metric*, i.e. users were required to locate target images among distractor images. Furthermore, the position of target and distractor images was random. Therefore, unlike a keyboard or keypad, users did not know in advance the position of the target elements. The user was required to locate them, each and every time. Moreover, the user not only had to locate the target images, they were required to align them. Therefore, users had to strategise to determine the best approach for alignment. Consequently, if images were scattered close together, surrounded by forgettable faces, the users would find it relatively straight-forward to locate and align them. If, however, the images were dispersed across the grid and surrounded by distinctive images, users would find it harder to locate and align them, increasing the overall time of the authentication attempt.

Therefore, as target images were not fixed and locating them is potentially

hindered by distinct distractors surrounding them, it is not unexpected that authentication attempt times would essentially oscillate. The argument could be made that the images be placed in fixed positions. Moreover, the images could be assessed and analysed to ensure they are suitably distinct from those that surround them. This design is plausible and not beyond the realms of implementation. However, such design decisions come at the cost of reduced resilience to attackers.

If target and distractor images are fixed in positions and not at random when loaded, attackers do not need to know images, they merely need to monitor interaction and perform the same actions subsequently. The same interactions performed on the same canvas will produce the same successful authentication attempt. The tailoring of distractor images that orbit target images to reduce confusion for the user, introduces an element that could be hijacked by an attacker. Therefore, ultimately the design decisions were taken to deliver an authentication approach that was resilient to attackers but at the cost of authentication time.

The second aspect of the design, that resulted in wide variation of authentication time, was the interaction language laid out during the registration process. The registration process required users to double-tap to enter, i.e. select an image for use in the authentication secret, long-tap and drag to reposition images within the authentication secret and swipes to move between steps in the process. The user was required to authenticate once they had completed the registration process. The first time the user is required to authenticate, they are presented a screen of text that instructs them how to use the authentication mechanism.

Unfortunately, when users were presented the authentication mechanism the interaction language used during registration was transferred to the authentication mechanism. Therefore, the short authentication attempt times suggest users located target images and double-tapped them. However, during authentication, a double-tap submitted an authentication attempt. Therefore, users likely double-tapped and were informed the authentication attempt was unsuccessful. Then, unsure of what occurred, users likely double-tapped on a target image again. If the user performed the action one more time, their account would have been deactivated.

The authentication mechanism was novel to users and efforts were made to communicate the idea that a double-tap acted as entry and text instructions outlined how to use the authentication mechanism. Unfortunately, some users clearly ignored instructions and simply remembered that a double-tap on an image resulted in action, during the registration process. The statistical analysis revealed that there was a difference in the time taken between successful and unsuccessful authentication attempts. This was unexpected as the process for a successful or unsuccessful authentication attempt should be same: only the conclusion is different.

Nevertheless, as there was a statistically significant difference between the time taken to complete a successful and unsuccessful authentication attempt suggests there was something distinct between the two processes that results in the respective conclusions. The answer is clearly that the interaction language laid out during the registration process was transferred to the authentication process. Therefore, the registration process was more important than first thought.

The prevalence of horizontal alignment among successful authentication attempts may also be explained by the registration process. The registration process presented the authentication secret in a horizontal alignment, it did not advertise or communicate to the user the fact they could make vertical or diagonal authentication attempts. Therefore, arguably the user transferred experience from the registration process to the authentication mechanism.

The importance of the registration process was not fully realised during the creation of the application and authentication mechanism. The reality is that most research on alternative authentication mechanisms simply does not discuss the registration process. The primary reason for this is that alternative authentication mechanisms are rarely deployed, negating the need for the consideration of the registration. However, if anything, evaluation in the field with students emphasises the importance of the registration process and how it is the first aspect of the alternative authentication mechanism that users encounter.

However, one aspect of the registration process that was considered important was the time spent completing it. The expectation was that the longer an individual spent creating an authentication secret, the more memorable it would become. The assumption was that a lengthy registration process signalled thorough consideration of the images within the authentication secret. However, statistical analysis did not reveal any significant correlation between the time an individual spent completing the registration process and the memorability of the authentication secret.

Nevertheless, other aspects could be connected to the memorability of the authentication secret, such as the number of authentication attempts made by an individual. The assumption was that authentication attempts, it would reinforce the memory of sequence of images. However, statistical analysis revealed no significant correlation between the number of authentication attempts an individual made and the memorability of the authentication secret.

Nonetheless, an authentication secret comprising of faces appeared memorable for at least some users. Unfortunately, users were not particularly diverse in the creation of their authentication secret. The reality is that users selected similar images and some images were more popular than others. The expectation is that users selected images that were familiar and memorable so as to ensure that they could authenticate without inconvenience. Consequently, many of the popular images were of core staff from the first year programme. These members of staff had been in place for several years. Therefore, every student would have encountered the members of staff and would likely have enduring memories of them, as they were some of the first teaching staff they encountered at the university. The phenomenon of individuals favouring specific images of faces is not unknown [62].

The explanation of the popularity of specific images is that individuals select images based on qualities such as race, attraction and similarity. However, these qualities were not particularly diverse within the image set. The popular images were of learned and experienced teaching staff being selected by inexperienced and youthful undergraduates. The assumption is that users were selecting familiar faces. The staff that were selected arguably represent celebrities within the department. The reality is that users strove to avoid the inconvenience of forget-

ting the images and favoured images they felt they would be able to remember. The problem for the authentication mechanism is that the predictability of image choice can essentially be utilised by an attacker. The resulting image set is essentially suboptimal.

The image set is the foundation of a recognition-based graphical authentication mechanism. However, even a strong image set is undermined by user selections, i.e. users will make similar choices. Ideally, an optimal solution would be to tailor the image set for an individual so that it is distinct. However, the creation of such an image set would not only require the user to provide images but also some sort of process or algorithm. The process would have to be kept secret otherwise an attacker could simply apply it to an image set to undermine the authentication mechanism. Nevertheless, ensuring specific algorithms and processes are kept private only increases the complexity of the authentication mechanism and potentially impacts on the scaleability as well as the deployability of the authentication mechanism.

However, an alternative authentication approach may be for the individual to create the image set. The user can provide the initial group of images and then prune and pick images until they have an optimal image-set. This would potentially increase the complexity and time taken to create an authentication secret but would tackle the problem of predictability without impacting on the memorability of the authentication secret.

5.6 Conclusion

The chapter introduced the design of the application and the integration of the alternative authentication mechanism. Furthermore, the chapter outlined the evaluation of the entire application in the field with undergraduates psychology students. The results of the evaluation revealed that authentication time was not reduced significantly with successive attempts and that the memorability of an authentication secret did not correlate significantly with the the time taken to create an authentication secret or with the number of times it had been submitted. However, the evaluation did reveal that there was a significant difference in time between successful and unsuccessful authentication attempts.

A potential explanation for the inability of users to reduce authentication times over successive attempts is that the authentication mechanism was a *search-metric* and, as such, required individuals to locate images before they could be manipulated. The explanation for the difference in time between successful and unsuccessful authentication is potentially explained by the design of the registration process that used one interaction language while the authentication mechanism used another. The remaining aspect discussed was the fact that users favoured some images more than others. The popularity of specific images, confirms that user choice is predicable, a quality that could be exploited by an attacker.

These concerns could be addressed in subsequent revision of the authentication and application by tackling the following issues:

- *User specific image collection*

There is no single image set, instead each user defines their own personal image set. The user provides a group of images and then prunes and picks images until a set is formed containing target and distractor images.

- *Re-invigorate registration process*
  The registration process needs to be re-designed and re-considered from the outset as the evaluation revealed that it is an incredibly important component of an authentication mechanism.

The two aforementioned aspects were tackled in the subsequent revision of the application, Dick. The design, implementation and evaluation of Dick are discussed in the next chapter.

Chapter 6

Dick

Tetrad appeared promising when evaluated in a vacuum. Nonetheless, when framed inside the Tom application and evaluated with actual smartphone users, Tetrad appeared less than desirable. The reality is that users made predictable image choices, undermining the image-set used in the initial incarnation of the application. The ecologically strong evaluation of Tetrad uncovered problems with the registration process, an area that is rarely discussed in authentication research. Self-supported registration processes are arguably not needed for controlled investigations. Nonetheless, it was clear there would need to be renewed focus on the registration process in subsequent iterations of the application.

Consequently, the following chapter outlines an iteration on Tom that (1) utilises a personalised image-set and (2) refines the registration process. The ensuing section, §6.1, details the design of the application, focusing on the alternative image set and registration process. The outlined aspects are used to actualise the application, §6.2. Consequently, a prototype application is produced and evaluated, §6.3, with results reported, §6.4, then discussed, §6.5. Lastly, conclusions are drawn, §6.6, with future steps outlined.

6.1 Design

The previous chapter outlined the various elements of authentication. Figure 6.1 illustrates the various elements, all of these were designed, implemented and evaluated in the previous chapter. However, the design decisions made for some elements had undesirable outcomes when evaluated, specifically: authentication images and the registration process as indicated with dotted lines in Figure 6.1. Therefore, these elements were revisited in the second iteration of the authentication mechanism. Furthermore, the application functionality was also revised to refine it and add additional features. The motivation for adding additional features was to provide incentives for users to return to the application.

6.1.1 Authentication Images

The previous incarnation of the application relied on an image set comprised of staff. There was a single image set for all users. Unfortunately, users made pre-

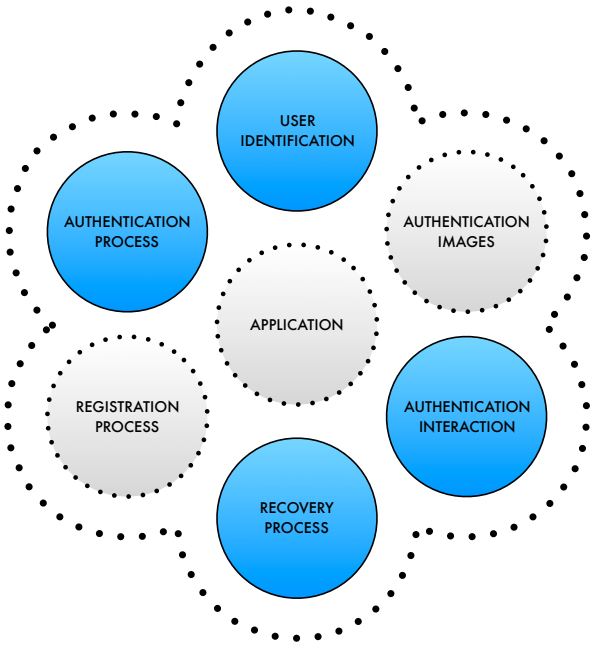

Figure 6.1: The many authentication elements, the elements in blue are settled while the elements in grey with dotted lines have to be refined.

dictable choices and some images were more popular than others. Consequently, a focus of Dick's design was to determine an alternative image-set.

Section 2.5 outlined and compared various image-types, such as scenes, objects and faces. The conclusion, §2.5.5, was that faces were the optimal choice for graphical authentication as they can be recognised rapidly and retained for decades. Therefore, the initial design decision taken was to continue using of faces, but to address the concern of predictable image choices.

The immediate element to address is the use of a shared image-set for all users, as predictable images choices are far more problematic in such a design. Consequently, an alternative direction was taken and the design focused on use of a personalised image-set for each user. The problem with a such solution is the practicalities of an individual sourcing such images. However, social networking sites such as Facebook can be used to generate the necessary image-set. Consequently, the second iteration of the application generated an image-set using images from a user's own social network. The images were generated using the profile pictures of friends. The profile picture was downloaded and processed to extract a single face, for use in authentication.

The solution not only has the potential to reduce the concerns of predictable image choices, but can also potentially improve the usability of the authentication solution. Ida Gobbini et al. state that familiar faces induce neural activity in areas traditionally associated with social knowledge. Moreover, Ida Gobbini et al. argue unfamiliar faces induce activity in areas, hypothesised as a 'social brake', when assessing potential threats [121]. Therefore, a more custom and familiar image-set may improve the authentication solution.

Nevertheless, there are still concerns that close friends and family may be

able to circumvent the authentication solution as they can potentially guess the authentication secret an individual is likely to create. The reality is that many knowledge-based authentication solutions are circumventable by friends and family, as they are highly familiar with an user's knowledge and taste. Nonetheless, an authentication solution should ideally be resistant to the efforts of such individuals. The assumption was that Tetrad already had such resilience as even if close friends and family were able to determine images used in the authentication secret, they would not necessarily be able to determine the sequence required to authenticate.

6.1.2 Registration Process

During registration the user was requested to enter their credentials for the popular social network Facebook. The user was also asked to grant the necessary permissions for the system to access the profile pictures of friends. The profile picture of each friend was downloaded and processed to generate a list of images. The user was asked to select all 45 images for use in authentication and to indicate the four images that would act as the authentication secret.

The registration process was more complex than is typically expected for an authentication mechanism. However, an individual should only need to endure the registration process once. Nevertheless, an unwieldy and awkward registration process could drive users away from an application. Therefore, every effort had to be made to ensure the registration process was as practical and painless as possible.

The processing of profile pictures to extract faces is increased the complexity of the registration process. The system had to present a list of images, containing extracted faces. Therefore, each profile picture had to be downloaded from Facebook's servers and processed to extract a single face. The process had to be performed for every single friend on a user's friend list. It relied on a steady Internet connection and powerful processor, if users were to complete the registration process in a timely fashion. Moreover, users could struggle to understand the complex process, given the novelty of the registration process. Consequently, three registration strategies for conducting the process were considered, as follows:

- *Device*
  The registration process could be conducted on the user's smartphone. The concern is the strength of the data connection and the limited resources of the device. The image processing could take a considerable amount of time to complete on the device. Moreover, the complex process could confuse users and result in many not completing the registration process, as there would be no support at hand to help users.

- *On-site*
  The registration process could be conducted on-site at a specific location. Users could register on specialised systems that have a powerful central processing unit as well as a fast and steady Internet connection. The time

taken to complete the image processing part of the registration process could be negligible using this strategy. Furthermore, someone could be hand to support and guide users through the registration process.

- *Cloud*
  The registration process could be conducted on the user's smartphone but the image processing part of the process could be completed on the server-side. The image processing could occur on a cloud instance rather than on the user's device. This drastically reduces the time taken to process images. There are several concerns connected with such a strategy, including cost and legality. Moreover, concerns still remain that the users may not be able to complete the complex registration process without guidance.

The initial strategy of expecting users to complete the registration process on-site is commonly used in biometric- and token-based authentication. However, it is rarely used in knowledge-based authentication, as these mechanisms can often be used without necessitating a specialised registration process.

Therefore, adopting an on-site registration process for the experimental authentication mechanism essentially undermined a primary advantage of using the knowledge-based authentication approach. Requiring on-site registration removed an advantage of the approach, making it less competitive with other authentication approaches.

However, arguably, for the purposes of evaluation the registration process could be completed on-site, although such a requirement would undermine the ecological validity of the evaluation. If users had to register on-site in a laboratory, then arguably every aspect of the evaluation could be controlled. The user could simply return each week to authenticate. The reality is that a realistic registration process could be so repulsive as to deter actual users from completing the registration process. An on-site registration process could mask users' displeasure with a novel registration process.

The remaining registration strategies retain a strength of the knowledge-based authentication mechanism, as users can complete the registration process without specialised equipment. The user in both strategies conducts the registration process on their own device but the image processing aspect is either handled in the cloud or on the device. The motivation for relying on cloud computing, rather than the device, for the image processing stage is to reduce the time taken to process images.

Cloud computing offers much better performance than modern smartphones. The time taken to complete image processing and thus the registration process could be dramatically reduced.

The inclusion of a cloud computing component within the registration process to improve it may be far more complex than first anticipated. However, the reality is that even if the cloud computing aspect were employed the resources would still cost considerably more than entry of a simple alphanumeric authentication secret. The transmission, storage and processing of alphanumeric authentication secrets would be dramatically less than that of graphics.

Therefore, even if an cloud solution were possible the cost might be too high when contrasted with other solutions. Companies are ultimately answerable to

shareholders. The drive to cloud computing is to avoid high capital expenditure and push for a manageable operating expenditure. However, there is the possibility that organisations could drive to push application power to users themselves.

The organisation could harness the power of user devices and push the costs to them. Furthermore, a key advantage of pushing the processing down to the user's device is that the personal information resides only on the user's device. Therefore, all analysis, processing and storage is on the user's device. Consequently, the approach is privacy preserving. Therefore, the analysis and creation of images will happen on the user's device rather than on an external cloud architecture.

6.1.3 Application

The initial design of the annotation user interface in 'Tom', was focused on *process* rather than *product*. The design aimed to enable students to enter notes effectively and efficiently. Unfortunately, the design was less than optimal when it came to reviewing notes.

The initial design of the chapter list interface only required users to tap on an element to edit it. Therefore, an individual only needed to tap an image to change it or tap the chapter title to change it. The chapter list is in permanent edit mode. The minute the user taps an image or title belonging to a chapter, they are immediately presented with a keyboard or image chooser. The design is perfectly acceptable as long as the user does not accidentally tap objects.

It is unrealistic to expect precise interaction from users is expecting too much. The reality is that users regularly make accidental taps [208]. Furthermore, encountering image choosers and keyboards can be incredibly irritating when users are not expecting them, especially when they are trying to review lecture notes. The user may be in a stressed state, preparing for examinations, and does not need the hassle of dealing with an awkward interface.

Furthermore, aside from interaction in the chapter list view, the annotation view, as illustrated in Figure 5.5, is also in permanent edit mode. The design of the annotation interface, on reflection, is terrible. The motivation for using a faux yellow sticky note and Marker Felt font was to emphasise a sense that the notes were short, simple and unformed. The sticky note was used to communicate the idea that users had limited space to make a note, they they should consider what they are typing. However, users were not actually restricted: they could create notes of almost any length.

However, such motivations clashed with the actual design of the interface. The design was focused on entering and editing information quickly but we expected users to produce considered and insightful notes. Moreover, if users do create a lengthy note they have to swipe up and down through text as the view was restricted between the top of the faux sticky note and the top of the keyboard. Therefore, the design squandered space that was already precious [148]. The surrounding screen real estate was scarified purely to represent a sticky note that had no real perceivable value, on reflection.

Consequently, the design of the annotation interface was refined to be optimal for editing and reviewing. The first step was to implement a drill-down interface. The chapter-list user interface still presented a list of chapters, each with an

associated colour, image and title. The user taps an element in the list to view annotations associated with the chapter.

Both the chapter-list and annotation user interface default to review mode, rather than to edit-mode. The user can tap and swipe anywhere on the screen and no keyboard or image chooser will appear. Instead an edit control in the form of a button, titled 'Edit', is now located in the top right hand corner of the interface. The user simply taps the button to enter editing mode; upon tapping the button is relabelled to 'Done'. The user taps the same button to exit edit mode and automatically save changes.

The annotation user interface simply consists of three simple elements. A colour banner with space for a single image, the chapter title and space and a notes area for annotations. When entering text into the notes area, the canvas slides up to accommodate the keyboard, ensuring that the text does not flow behind the keyboard. Furthermore, when reviewing the text the user can use the whole screen and swipe up and down: no space is wasted.

The aim of these changes is to provide a simple and clean interface to users. The changes should make it easier for users to review and edit content without encountering an unexpected keyboard or image chooser. Moreover, the design enables the introduction of the chapter-sharing feature. Annotations shared with other users are not editable. Therefore, shared chapters appear in the chapter-list alongside the chapters generated by the user. However, users can drill-down on any shared chapter to view annotations associated with it. When the user is presented with associated annotations, the edit button is absent.

Hence, user interaction is consistent across all chapters, except that the additional edit operation is absent for shared chapters.

6.1.4 Proposed Solution

The proposed solution is an application that allows individuals to download and annotate lecture recordings. The annotations are shared across users to provide the sharing of content. The feature addition was meant to promote regular use of the application and consequently the authentication mechanism.

The authentication mechanism relies on a personal image collection. The decision was taken to download and analyse images on the user's device rather than use a cloud component. The motivation was to preserve privacy as the user's personal information would not be transferred and analysed on external cloud components. Furthermore, another motivation was to reduce the costs associated with performing analysis on external cloud components.

The implementation of the application and alternative authentication mechanism are discussed in the next section.

6.2 Implementation

The implementation of the application and authentication mechanism is detailed over the next few sections.

6.2.1 Registration Process

The registration process consisted of four stages. Dick's registration process is more complex than Tom's, as it requires the user to provide the image collection used in the graphical authentication mechanism. The decision was taken to rely on an individual social networking account, specifically Facebook. Consequently, users were required to have a Facebook account. The application would essentially download the profile picture of each friend in an individual's friend and extract a face from the image. The user was then expected to create an image set for the authentication mechanism. They then selected four of the images as their authentication secret. The steps involved in the registration process are detailed below:

1. *Enter Facebook Credentials*
   The implementation relied on the Facebook SDK to connect with the user's Facebook account. The SDK provides a standard user interface for connecting external applications with an individual's Facebook account. Users tap the button to initiate the connection process. Facebook manages the connection process.

   The initial step in the connection process requires the user to enter their Facebook credentials. Once authenticated, Facebook presents additional dialogues informing the user that the iOS application is seeking access to their Facebook account. Furthermore, the dialogues outline the specific permissions the application is seeking, e.g. access to photos. The user is free to allow for the application to be connected to their Facebook account or deny such connection. Once the user grants the connection the registration process can proceed. Facebook generates a session for the application and provides the necessary data to access the user's account. The application must store and present information to gain access to the account later.

   Users are free to subsequently sever the connection at anytime via the Facebook website. If users choose to deny access in the first instance, the registration process does not continue and they are unable to use register and use the application.

2. *Select 45 friends for use in authentication*
   The second step essentially involves downloading and processing the profile picture of every friend on the user's social network. Users are presented an empty UITableView, in the second step.

   The table view header contains a prompt that instructs users to select 45 images for use with the authentication mechanism, as illustrated in Figure 6.2. Underneath the prompt is a progress bar with accompanying label. The label indicates how many friends, or rather profile pictures, have been processed. The table view is populated with selectable entries as each profile picture is processed. An entry comprises of the extracted face and the name of the friend used to generate it.

   The image processing subtask continues until all images have been processed or the user has selected 45 images. The prompt in the table view header

is updated inline with the number of images selected by the user, e.g. if the user has selected 5 images, the prompts states "Please select 40 Friends from the list...". When the user has selected 45 images, a button is enabled for the user to progress to the next step of the registration process.

If the user chooses to progress to the next step, the image processing task is terminated, regardless of whether the profile pictures have been processed.

Figure 6.2: An example of user interface in step two of the registration process.

3. *Select 4 of the 45 as the authentication secret*
   The user merely has to identify 4 of the 45 images that will comprise their authentication secret.

   Users are presented a UITableView containing the 45 images they selected during the previous step. Users are instructed to select 4 of these images to act as their authentication secret. Once an individual has selected 4 images, a button is enabled for the user to progress to the fourth and final step.

4. *Sequence is important, re-order before confirming*
   The final step of the registration process requires the individual to determine the sequence of the images in the authentication secret.

   The user is presented a UITableView containing the four images they selected in the previous step. The interface object allows individuals to re-order entries within a table view. Users tap and hold the 'sequence control' to detach the entry from the table view and reposition it. The user can reposition entries until they are content with the sequence of their authentication images.

   The user taps a button labelled 'Done' in the table view header to complete the registration process.

Image Generation

The second step of the registration process is the most complex step. An overview of the three key steps performed in the second step are outlined, as follows:

1. *Query Social Network*
   The registration submits two queries to Facebook. The first query request the user's friend list, i.e. a list of the all the relationship an individual has on the social network. The second query requests a profile picture for each associate, on the list returned from the first query.

   The implementation submitted queries using the Facebook Query Language (FQL). The social network does not use the *de facto* Structured Query Language (SQL) and instead favours its own variant FQL. The returning results are parsed to produce an array of unified resource locators (URLs) to download the necessary images.

2. *Generate Image*
   The profile picture associated with each friend is processed to locate faces within the image. The image processing subtask returns an array of co-ordinates for each face located within an image. The system is not recognising faces, merely identifying face-like objects within an image. Therefore, the face located within an image is not necessarily that of the profile owner. The system may locate several faces within an image but only one is required. Consquently, the system simply selects the first set of coordinates from the array.

   The general assumption is that a face will be extracted and will likely be that of the profile owner. Nevertheless, the extracted object, may not be that of the owner, it may not even be an actual face.

3. *Display Image*
   The extracted image is presented along with the name of the profile owner, as an entry in a table view. Therefore, users see the actual image before they confirm selection.

The face problem did consider building an interface that allow individuals to see the images located and change locations but this complicates the situation even more. A user interface displaying images and allowing users to edit would also represent a cost.

Performance

The analysis and extraction of faces was especially intensive. However, every effort must be made to ensure the process is as efficient as possible to ensure the user does not need to spend considerable time completing the registration process. The following aspects of iOS development are expensive[138], as follows:

- *NSManagedObject*
- *Saving to disk*
- *Network*
- *Parsing*

The aforementioned aspects are all involved in the creation of the user's personal image set. Therefore, the process must be careful considered and structure to ensure the process is efficient. The steps involved in the image processing are discussed in the next section.

Image Processing

The decision was taken to use a multi-threaded approach to maximise performance on smartphones, such as iPhones. The steps are as follows:

(a) *Generate NSManagedObject*
A NSManageObject, essentially a record, is generated for each profile picture. The NSManageObject is contains information pertaining to the profile picture.

(b) *Download Profile Picture*
The URL for the profile picture was used to download the image onto the user's smartphone. Establishing, maintaining and using a network connection is expensive and is dependent on the speed of the connection. The optimal scenario is for a user to complete the task using a strong Internet connection. The connection would preferably be over a wireless local area network rather than a cellular network. A cellular connection would require considerably more energy to maintain and opens the user to potential expensive data costs.

The image itself was not written to disk as the entire image is not needed, merely a single face from it. Furthermore, the process of writing an image to disk is expensive, even on a separate background thread. Therefore, writing to disk unnecessarily would only increase the overall time taken to complete the registration process. Consequently, the image is passed for image processing once downloaded.

(c) *Locate face(s) within profile picture*
The OpenCV library is packaged with the application to make use of several optimised routines, functions and algorithms for image processing. However, even when using such specialised and refined routines, the image processing subtask still consumes a reasonable amount of resource. The demand on resources only increases with the size and quality of the actual profile picture. There is a concern that large profile pictures could exceed memory allocations and lead to the termination of the entire application by the device's operating system.

Therefore, the size and quality of images are determined before image processing commences. Profile picture exceeding a specific threshold are rejected before image processing commences.

Those images not exceeding the threshold are passed to the image processing subtask. The subtask does not recognise specific faces but merely locates potential faces within the profile picture. When a potential face is located, the coordinates are added to a results array. The processing technique does

not necessarily locate all faces, neither does the process only locate faces, it may well incorrectly identify face-like object as faces.

(d) *Extract face from profile picture*
The coordinates array generated in the previous step is used to extract a face from the profile picture. The array may contain several coordinates: only one face is necessary from each profile picture. Therefore, the first set of coordinates in the array were used to extract a face from the profile picture.

(e) *Image written to disk*
The extracted face is not inserted into the NSManagedObject, as this would increase the resources needed to subsequently load the managed object. Instead, the extracted face, generated in the previous step, is written to disk and the profile picture used to generate it.

(g) *Discard profile picture*
The entire profile picture consumes precious resources and is no longer needed once a face has been extracted. Therefore, the profile picture is discarded once a face is extracted.

(h) *Update NSManagedObject*
The NSManagedObject is updated with the physical location of the image written to disk.

(i) *Merge*
The changes made to the local NSManagedObject are merged with the central NSManagedObjectContext.

(j) *Discard thread*
The thread is discarded and resources released.

The aforementioned steps are used to create the user's personal image set. The process was part of the registration process and was packaged with the application.

6.2.2 Application

Dick's essential functionality was essentially unchanged from the previous incarnation, Tom. The main change was the reinvigoration of the chapter component, specifically the chapter list. The chapter list component needed to accommodate the sharing of chapters with other users.

The annotations created by an individual are shared with all others. The aim is to reduce the complexity of managing privacy controls and settings. The user is advised that if they do not want to share annotations with others, they should simply not create them.

The previous incarnation of the chapter component assumed the user was in edit-mode. The new approach assumed the user was in 'consumption-mode' and users could browse their own annotations and the annotations of others. The

user could enter 'edit-mode' by selecting a chapter and tapping the edit button. The user could not edit annotations by other users. Furthermore, while users could attach an image to their own annotations, these images were not shared with others. The primary motivation for this design decision was to reduce the traffic being exchanged between devices and associated cost.

The application was distributed and evaluated on students, as discussed in the next section.

6.3 Evaluation

The details of the evaluation are outlined over the following sections.

6.3.1 Subjects

The application was distributed and evaluated on undergraduate students enrolled at The School of Psychology at the University of Glasgow. The application required individuals to have a Facebook account. Furthermore, students had to give permission to allow the application access to their Facebook account.

Consequently, the user-base had the potential to be small as users may not have met such requirements. The reality is that many individuals likely have a Facebook account as the social network service is incredibly popular. However, students may be less the willing to connect their personal social networking account with an experimental application. Nevertheless, Nielsen argues that even five users provide enough feedback to evaluate a system [192]. Consequently, the decision was taken to continue with the evaluation, accepting that while a small participation group was undesirable, it still had the potential to be useful.

6.3.2 Apparatus & Material

The apparatus and materials are similar to the evaluation of Tom, see §5.3.2.

6.3.3 Procedure

Similarly, the procedure for the evaluation is also akin to the previous evaluation for Tom. The primary difference was that the application was not distributed via the Apple App Store but via an independent testing platform called TestFlight. The TestFlight system was used to distribute the application to students. The platform required the installation of specific technical profiles on the user's device. Consequently, using the application was more complicated than the previous iteration the added complexity was unavoidable if we used TestFlight.

6.4 Results

The results are discussed in terms of the registration process and authentication process.

6.4.1 Registration

There were 9 registrations over 55 days. The majority of these registrations (67%) occurred within 30 days and few registrations (22%) occurred after 40 days.

Similar to the previous iteration, i.e. 'Tom', the application did not collect any personal information, such as an individual's email address and did not request individuals to create usernames. The application also did not log personal information associated with an individual's Facebook account. The application instead relied on the device UDID as the individual's username. The device UDID was associated with each registration. Consequently, the number of unique users can be estimated from the number of distinct UDIDs. The total number of distinct devices that completed the registration process was 6.

Furthermore, similar to Tom, there was no recovery process. Users unable to authenticate were expected to delete the application, download it again and conduct the registration process again. Similarly, if the application encountered a problem or malfunctioned, the user was expected to remove the application and complete the registration process again. Therefore, while several users (50%) completed the registration process once, several users (50%) completed the registration process twice. The registration process was not completed more than twice by any individual.

The majority of registrations (44%) were initiated between midday and six o'clock in the evening. In fact most registrations (78%) occurred between midday and midnight. Furthermore, no registrations occurred during the night and few (22%) occurred between six o'clock in the morning and midday. Moreover, the majority of registrations (78%) occurred on a weekday rather than at the weekend with most registrations (44%) happening on a Wednesday.

Time

Individuals spent a total of 108885 seconds or approximately 30 hours on the registration process. The average time taken to complete the registration process was 201.64 minutes ($SD = 418.387$). However, the mean may not be indicative of overall performance as there was wide variation in the time taken to complete the registration process (minimum = 18.01, maximum = 1307.93, range = 1289.92). The time taken, on average, is considerable, longer than many films, e.g. individuals could watch Titanic or the finale of Lord of the Rings in the time it takes to complete the registration process.

However, the majority of individuals (89%) completed the registration in under 200 minutes. Inspection of the registration times revealed that the average performance was potentially distorted by an extreme outlier, a single individual that took 1308 minutes or 21.8 hours to complete the registration process. The average time taken to complete the registration process was 63.35 minutes ($SD = 57.97$). However, the mean may not be indicative of overall performance as there was wide variation in the time taken (minimum = 18.01, maximum = 181.43, range = 163.42). Furthermore, the majority of individuals (75%) completed the registration process in less than 60 minutes.

The time taken may be reduced if only initial registration attempts are con-

sidered. The average time taken to complete the registration process for each user was 82.85 minutes ($SD = 67.29$). However, the mean may not be indicative of overall performance as there was wide variation in the time taken to complete the registration process (minimum = 24.81, maximum = 181.43, range = 156.62). The reality is that the registration process took a considerable time to complete.

6.4.2 Authentication

Users made a total of 151 authentication attempts over a period of 77 days. The majority of authentication attempts (53%) ended in success. The remaining (47%) authentication attempts were unsuccessful.

A total of 7 users made at least one authentication attempt. The average number of authentication attempts was 21.57 ($SD = 38.109$). However, many users made far fewer authentication attempts (mode = 2). The number of authentication attempts generated by users was mixed (minimum = 2, maximum = 107, range = 105), suggesting the distribution of authentication attempts among users was more varied than indicated by the mean.

Furthermore, all users made at least one successful authentication attempt ($N = 7$). The average number of successful authentication attempts was 11.43 ($SD = 14.513$). However, the mean may not be indicative of actual performance for most users (mode = 2). There was a large span in the number of successful authentication attempts (minimum = 1, maximum = 42, range = 41). The average number of unsuccessful authentication attempts was 10.14 ($SD = 24.210$). However, the mean may not indicative of overall performance for most users (mode = 1) as there was a wide variation in the number of unsuccessful attempts among users (minimum = 0, maximum = 65, range = 65). Moreover, several users (57%) did not make a single unsuccessful authentication attempt.

Similar to Tom, users were constrained in that each registration was limited to 3 sequential unsuccessful authentication attempts. The account was deactivated once the limit was reached, and users had to re-register to use the application. However, users were also expected to complete the registration process again if the application encountered problems or if they removed it from their device. Therefore, several users (50%) completed the registration process more than once. Users completed a total of 9 registrations ($N = 6$), all individuals bar one, were accounted for in authentication attempts. Similar to Tom there was an individual that was not accounted for in the registration logs. The individual generated a total of 10 authentication attempts.

Furthermore, none of the accounts or registrations appeared to have been deactivated as no individual submitted a sequential number of unsuccessful authentication attempts. Therefore, users appeared to have re-registered as the application malfunctioned or did not behave as expected. Alternatively, users may simply have removed the application and endured the registration process for other reasons.

Moreover, there were several authentication attempts (63%) generated by a individual that could not be associated with a specific registration. Therefore, these attempts were simply not logged on the server. However, the authentication attempts exhibit anomalies that suggest the application or mechanism may have

malfunctioned in these cases. The authentication attempts are discussed in §6.4.2 and were removed from consideration.

Therefore, a total of 46 authentication attempts were generated over 9 registrations ($N = 6$). The vast majority (89%) of these were successful with a few (11%) unsuccessful.

The average number of authentication attempts for each registration or account was 5.11 ($SD = 5.302$). However, many accounts generated far fewer authentication attempts (mode = 1). Furthermore, there was wide variation in the number of authentication attempts associated with each account (minimum = 1, maximum = 16, range = 15). The average number of successful authentication attempts for each registration or account was 4.56 ($SD = 4.953$). Similarly, many accounts had far fewer successful authentication attempts (mode = 1) associated with them. Moreover, the number of successful attempts submitted varied among accounts (minimum = 1, maximum = 15, range = 14). The average number of unsuccessful authentication attempts for each account or registration was .56 ($SD = .527$). In this case, the mean may be indicative of performance (mode = 1) as few unsuccessful authentication attempts were submitted by users (minimum = 0, maximum = 1, range = 1).

Time

Individuals spent a total of 181130.69 seconds on 46 authentication attempts or 2.1 days. The average time for an authentication attempt was 3937.62 seconds ($SD = 20420.913$). However, there was wide variation in the time spent on each authentication attempt (minimum = 11.00, maximum = 131063.97, range = 131052.97). There was several authentication attempts that took considerable time.

There were two extreme outliers lasting 47968.16 and 131063.97 seconds or 13.32 and 36.41 hours, respectively. The two authentication attempts were successful and generated by the same individual. Both of these distort typical performance. Therefore, both outliers removed. Consequently, the average time taken for an authentication attempt was 47.69 seconds ($SD = 27.901$). However, there was wide variation in the time taken for each authentication attempt (minimum = 11.00, maximum = 127.12, range = 116.12) suggesting the overall mean may not be indicative of performance.

Further inspection of authentication times revealed two potential outliers that lasted longer than 100 seconds. These two authentication attempts are unlikely to make a dramatic impact. However, with both were removed, and the average time taken for an authentication attempt reduced to 43.97 seconds ($SD = 22.450$). However, there was still wide variation in the time taken for each authentication attempt (minimum = 11.00, maximum = 97.14, range = 86.14). The time taken for authentication attempts may improve, when considering only successful authentication attempt.

The average time taken for a successful authentication attempt was 44.79 seconds ($SD = 22.711$). Here too, there was wide variation in performance in successful authentication attempts (minimum = 16.47, maximum = 97.14, range = 80.67). The authentication times for unsuccessful authentication attempts may

be shorter. The average time taken for a unsuccessful authentication attempt was 37.88 seconds ($SD = 21.682$). There was wide variation in performance in unsuccessful authentication attempts (minimum = 11.00, maximum = 64.74, range = 53.74). The time taken for unsuccessful authentication attempts appears to be shorter successful authentication attempts but there were far fewer unsuccessful attempts. However, there were far fewer unsuccessful authentication attempts to consider when contrasted with successful authentication. The typical performance for unsuccessful authentication attempts may approach that of successful authentication if more were considered.

The higher number of successful authentication attempts is a relief, as users were required to select target and distractor images. Therefore, it would appear that requiring users to select distractor images may not have impacted on performance. However, while users may be able to successfully authenticate, it may take them longer. An increase in authentication attempt time would be undesirable as authentication times were already considerable. Therefore, a paired samples t-test was conducted to determine if there was in difference in performance between the two authentication mechanisms.

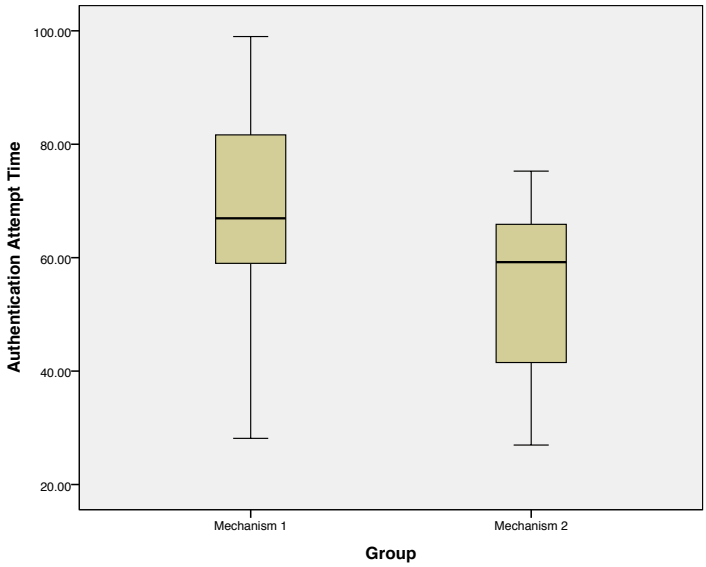

Figure 6.3: Box plots for each of the authentication times for each mechanism.

Figure 6.3 illustrates the authentication time for the first attempt generated by users in both versions of the authentication mechanism. The difference between both authentication attempts generated by users was normally distributed as suggested by a Shapiro-Wilk's test ($p > 0.05$). An authentication attempt took longer to complete using the second iteration of the authentication mechanism ($M = 52.66$, $SD = 19.180$) when compared to the first authentication mechanism ($M = 13.66$, $SD = 5.327$). There was a statistically significant difference of 38.998 seconds, 95% CI [9.8624 to 68.1336], $t(4) = 3.716$, $p < 0.05$, $d = 1.66$. Therefore, users took longer to complete an authentication attempt in the second iteration of the authentication mechanism.

The difference is surprising as the authentication would be initially novel to users. Therefore, it is expected the authentication times using the second iteration will be shorter. However, many of the initial attempts submitted by users with the first iteration of the authentication mechanism were unsuccessful, they were short as users suffered from interaction errors and getting to grips with the mechanism. Therefore, authentication times for the second iteration are more likely to be representative of actual performance with confident users comfortable with the approach. Consequently, it may be better to contrast authentication times for the second iteration with initial attempts from successful users of the first authentication mechanism.

An independent t-test was conducted to determine if there was difference between performance in both authentication mechanisms. Box plots of the authentication times for 'Tom' and 'Dick' were inspected and there was no extreme outliers. The authentication attempt times for each authentication mechanism was normally distributed as suggested by a Shapiro-Wilk's test ($p > 0.05$). An authentication attempt took less time to complete using the second iteration of the authentication mechanism ($M = 54.67$, $SD = 17.850$) than the first authentication mechanism ($M = 54.67$, $SD = 17.850$). There was homogeneity of variances for authentication attempt times for both groups as determined by Levene's Test for Equality of Variances, $p = .632$. An authentication attempt took 12.27 seconds ($SE = 12.201$) longer using the initial iteration of the authentication mechanism. However, the difference in performance between both authentication mechanism was not significant, $t(10) = 1.006$, $p = .338$.

Nevertheless, while the difference may not be significant, users are able achieve similar performance using the second iteration of the authentication. Therefore, expecting users to select distractor images as well as target images does not appear to impact on the performance of users. However, the sample size is small, making the relevance of any analysis questionable. Nonetheless, users appear not only able to authenticate successfully with the second iteration but without an impact in performance. However, users selecting their own distractors may impact on the memorability of the authentication secret.

Memorability

The memorability of the graphical authentication secret was determined by calculating the number of days between the creation and last submission of an authentication secret. The average of memorable of the graphical authentication secret was 17.22 days ($SD = 24.493$). However, there was wide variation in the number of days in the memorability of the authentication secret (minimum = 0, maximum = 61.07, range = 61.07). However, if duplicates are removed, the average of memorability of the graphical authentication secret increases to 24.59 days ($SD = 27.555$). Nevertheless, there was still wide variation in the number of days in the memorability of the authentication secret (minimum = 0, maximum = 61.07, range = 61.07).

Nonetheless, memorability is difficult to assess as users were not assessed at fixed intervals but rather when they decided to authenticate. Therefore, several users did not authenticate beyond a few days. Furthermore, users of the second

authentication mechanism were fairly successful: there were few unsuccessful authentication attempts. There was no sequence of unsuccessful authentication attempts that would suggest failure in memorability, rather many users simply did not use the application.

Nevertheless, there were still many unsuccessful authentication attempts that seem to suggest the authentication malfunctioned rather than that the user could not recall the authentication secret.

Excess Attempts

105 authentication attempts were generated by two individuals and were not associated with any registration record. The majority (63%) of these authentication attempts were unsuccessful. Furthermore, the authentication attempts exhibit anomalies that suggest the application as whole or the authentication mechanism was not functioning properly.

The success of an authentication attempt does not appear to lead to the on-screen dismissal of the authentication mechanism. Therefore, users appear to have submitted several authentication attempts, both successful and unsuccessful, but could still not access the application. The flaw could stem from incomplete registration. The missing registration records suggest that the registration process may have failed in these instances and led to the generation of these unusual authentication attempts.

The average time taken for each of the authentication attempts was 713.87 seconds ($SD = 7099.587$). There was wide variation in the time taken for the authentication attempts (minimum = 3.96, maximum = 72768.11, range = 72764.14). However, there were extreme outliers that likely distorted the typical time taken. Consequently, all five authentication attempts were removed from consideration. The average time taken for each of the authentication attempt was 14.43 seconds ($SD = 8.135$). However, there was still wide variation in the time spent on each authentication attempt (minimum = 3.96, maximum = 44.88, range = 40.91).

The authentication attempts can be parcelled into essentially five misfired attempts with four generated by one user. The initial parcel comprised of 45 attempts, the second comprised of 39 attempts, the third parcel compromised of 4 attempts, the fourth parcel compromised of 7 attempts and the last parcel comprised of 10 attempts. An interesting aspect of these attempt parcels is that users initially submit a successful authentication attempt and then continually submit them, confident in their selection and most likely the system confirms they are correct. However, after submitting several attempts to no avail, user begin to experiment and submit several unsuccessful authentication attempts. In specific parcels users try alternative alignments. The experimentation illustrates the tenacity and determination of users to overcome the authentication mechanism and access the application.

6.5 Discussion

The purpose of the application was to evaluate the authentication mechanism within an actual workflow, on actual target devices with actual users. The previous incarnation of the application, Tom, presented an image set comprising of images of staff faces. Individuals were expected to select target images and use these to authenticate by aligning them within a grid. The concern with the approach was predictability of choice.

There are many potential solutions to the problem, such as tailoring image sets to individuals or requesting individuals to provide target images. However, there are also many potential drawbacks to using such solutions, namely any algorithms used to tailor images could be used nefariously by attackers and user-provided target images could be easily identified from system distractors. Alternatively, individuals could provide their own collection of images and subsequently prune and pick images until they have an image set of target and distractor images for use in graphical authentication.

However, there are two primary concerns with using the aforementioned approach, namely (1) sourcing of a sizeable image collection from each individual and (2) expecting individuals to select distractors as well as target images. The first concern was addressed in Dick by requiring an individual to own and operate a social networking account.

The profile picture of each friend in an individual's friend list was downloaded. The image was analysed for faces and one was extracted and used for use in the image set. The images of faces were presented to users in a vertical list during registration and individuals were requested to generate an entire image set for use authentication. Once users had generated the image set they were required to select four of the images as targets, these four targets would form the user's authentication secret.

The generation of the initial image collection was performed entirely on the individual's device. The images were not stored on servers or residually in any other location. The images that were generated from use in authentication approach resided on the individual's device. Therefore, once the individual deleted the application, the images were destroyed alongside the application. The architecture was meant to preserve privacy and reduce any concerns surrounding the collection and use of personal images and information.

However, the design decision resulted in the registration process being very expensive in terms of time and energy. The limitations of mobile computers also limited the pace of the registration process. Therefore, registration time was considerably longer than previous incarnations of the application, namely Tom, but the expectation was that users would only have to endure the process once.

The application was, not well received. While there was considerable effort to ensure the application did not collect personal information, some users still perceived the authentication mechanism negatively. Indeed, upon presentation of the concept to one individual, they remarked the concept was 'creepy'. Furthermore, on reflection, the registration process expected and required a considerable amount of effort from individuals.

An individual was expected to own a modern and powerful iOS device, have

an active social networking account with at least 100 friends and a strong data connection to the Internet. Users were also expected to devote 10 to 30 minutes to completing the registration process. The energy and effort required to use the authentication mechanism was out of balance with the application. The user was expected to do too much, that became clear in hindsight. The reality is that many students simply refused to endure the registration process and ignored the application. It was used by few individuals because the registration process expected too much.

Therefore analysis can only be performed on the few individuals that did use the application. The key area of interest is whether there is a difference in the time taken to complete an authentication attempt in the alternative version of the authentication mechanism. The requirement of users to select distractors and as well as targets could well increase the time taken to complete an authentication attempt. The statistical analysis revealed that users performed better on the authentication mechanism in Tom than in the second authentication mechanism in Dick. Consequently, the use of distractors could potentially impact on the time taken to complete an authentication attempt.

An independent group comparison between similar performing users suggests that users did not improve their authentication times with the authentication mechanism used in Dick. However, the authentication time does not appear to be impacted from the use of distractors that the individual selected. Therefore, personal image collections could be used in an alternative authentication mechanism. However, the analysis was performed on a small sample. Therefore, the weight of such analysis is debatable. Nevertheless, those users who authenticate appear to be able to authenticate successfully and in a timely fashion.

The impact of distractors is arguably irrelevant as the registration process was incredibly intensive for users. There was also the concern that an attacker could easily identify likely friends within the grid and determine the sequence of images used to authenticate. This may be difficult for an attacker who has little knowledge of the user. However, it may be relatively painless for an acquaintance or associate who observes several authentication attempts.

There was also the concern of 'traffic shaping' that was not fully appreciated during the design phase of Dick. Cellular networks regularly use traffic shaping to reduce the burden and payload on their networks. The operators of the cellular network can interfere with large elements of data such as video and images to reduce the demand on their network. An image or video can be severely reduced in quality. The concern for the authentication mechanism is that high-quality images are reduced in quality to meet the needs of the network. These lower quality images would be stored during registration and presented during authentication. The 'traffic shaped' images could potentially impact on the time taken to complete an authentication attempt, if not make it impossible to use the graphical approach.

These concerns must be overcome to improve the authentication mechanism. However, it is still not clear how the registration process can be improved to reduce the burden on users. Silicon inside smartphones will become more powerful and efficient but the real bottleneck is the data connection to the Internet and that can not be guaranteed. The images still need to be downloaded before being

analysed. The analysis could occur in the cloud but it would require a strong data connection. The reality is that specific users may enjoy incredibly strong and powerful data connections while others have to rely on sporadic and unstable data connections. Even if the concerns of the data connection are resolved, user data would need to reside on external systems and computers. These computers will have backup routines, replication and any number of strategies that will ensure that images and personal information could linger on external services for several months.

The authentication mechanism must preserve privacy otherwise users will distrust the process and may assume it is merely being used to collect personal data. However, there are other ways to generate a personal collection of faces that do not occur on the individual's device. The Facebook service specifically outlines individuals to tag and edit images of individuals online. Therefore, rather than analysing images on device to extract friend's face, the tagged images on Facebook can be used to generate a collection of personal images. The collection could be downloaded to the user's device and once again when the application is deleted the images would be destroyed alongside it. Furthermore, additional images of each friend could be downloaded and oscillated through to reduce the likelihood on an observer being able to identify the target images.

6.6 Conclusion

The chapter outlined the design, implementation and evaluation of an alternative authentication mechanism that relied on a personal image set. The user was required to provide a personal image collection that they subsequently pruned and picked to create an image set comprising of target images and distractors. There was concern that authentication time would be impacted as users were required to discern between distractors and the targets they selected. Users did not appear to struggle with these specific requirements.

However, there were too few individuals to analyse because the registration process was too intense. The authentication mechanism required too much from users and, as a consequence, the registration process was expensive in terms of energy and time. Few users were generous with their time and few registrations were completed. There was also concern that close acquaintances or associates would be able to identify target images. Lastly, there were some concerns that images may be interfered with and compressed by the data connection, resulting in images that were unsatisfactory for use in the authentication process. The authentication mechanism showed promise but these concerns had to be addressed and investigated if the authentication mechanism was to be a practical alternative to passwords.

Therefore, the following two issues had to be investigated and tackled in the subsequent revision of the authentication mechanism:

- *Alternating image collection*
  The target and distractor images would alternate between authentication attempts to ensure the mechanism was resilient to attackers that were close acquaintances or associates.

- *Image quality*
  The data connection used by the individual to complete the registration process would be recorded and subsequent authentication attempts analysed to determine if there is any impact on authentication time.

These concerns and aspects are investigated in the next iteration of application, Harry. The design, implementation and evaluation of Harry is outlined in the next chapter.

Chapter 7

Harry

Tetrad appeared weak after two successive ecologically strong evaluations. The authentication solution, while promising in a controlled evaluation, simply did not appeal to actual smartphone users. The crossover to the smartphone context was arguably not the root of all problems, but rather the lengthy and complicated registration process. The alternative authentication mechanism clearly created a conflict within the application that resulted in users either ignoring the application entirely or abandoning it.

Consequently, the following chapter outlines an iteration on the previous application that (1) attempts to refine the registration process and (2) extends into another context. The ensuing section, §7.1, details the design of the application, focusing on the alternative image set and registration process. The outlined aspects are used to actualise the application, §7.2. Consequently, a prototype application is produced and evaluated, §7.3, with results reported, §7.4, then discussed, §7.5. Lastly, conclusions are drawn, §7.6, with future steps outlined.

7.1 Design

The primary design focus of the third iteration of Harry was (1) iterative improvement and (2) context expansion. The original plan was to extend Tetrad across contexts. The alternative authentication mechanism may have been designed to be optimal on televisions but had to at least be functional on other devices. The motivation was that as tasks cross context so does an authentication mechanism. Consequently, the ground work was laid in transitioning the application to tablets during the second evaluation.

Unfortunately, the second evaluation exposed several problems with the authentication solution that had to be addressed. Therefore, more effort had to be applied to the third application to address the uncovered concerns. Nevertheless, the design decision was taken to continue with expansion to tablets. The motivation was that much of the work had already been completed in transitioning the application to tablets during the second evaluation. Furthermore, the assumption was that considering the design that worked for both smartphones and tablets may produce an acceptable the solution. Moreover, it was felt that offering a tablet application may attract more users. Therefore, the decision was taken to

continue with context expansion and target the iPad.

The primary focus was improving on the second iteration of the application. Figure 7.1 illustrates the various elements of authentication. Many of these elements have been defined in previous iterations of the application. However, some need to be addressed to improve the authentication mechanism. Similarly to the previous iteration, the elements that need improved are the authentication images, registration process and application itself.

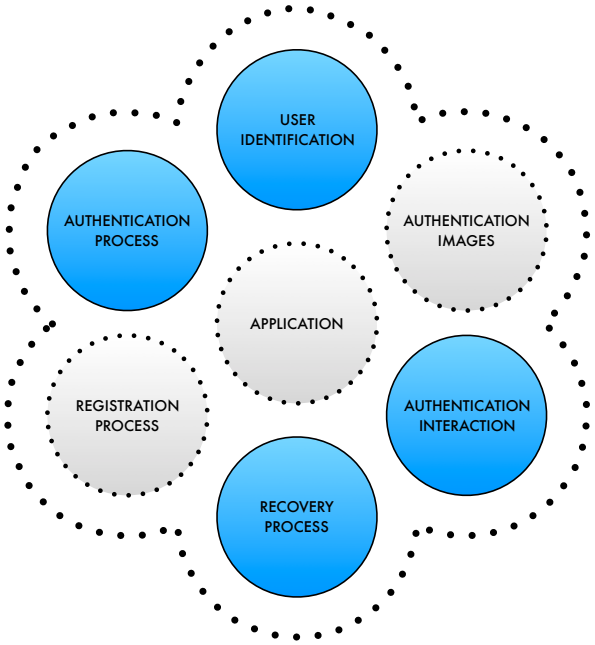

Figure 7.1: The many authentication elements, the elements in blue are settled while the elements in grey with dotted lines have to be refined.

The transition from the television was relatively straight-forward as only the experimental authentication mechanism existed. The other elements of authentication were never envisioned for the television, never mind manifested. The registration process and application were designed for a smartphone. These designs may not be easily transition from the iPhone to the iPad.

The iPhone and iPad has no windowing environment: there is no desktop. The user is presented a screen at a time. Consequently, applications designed for the iPhone target a screen. The design decision makes sense considering how precious screen space is on smartphones. The operating system does not sacrifice screen space to windows and menus. Applications are designed to span screens with clear focus and simple interaction. The approach ensures an individual can continue to use the device while on the move.

The approach was then translated to the iPad. Consumers clearly value the use of screens of content, large buttons and modal interaction as the iPad is popular.

The authentication mechanism is transferred relatively painlessly as it was a single screen of content. However, transferring the registration process and application the design had to consider the fact they were leaping from a small

screen to a large screen. The registration process, for example, consisted of several small screens that could have all been easily presented at the same time, on a large screen.

Therefore, considering the design of the iPad interface forced focus on the user interface for registration. The process previously consisted of steps, users were presented a series of small screens. However, a superior solution may be to present the entire process as a single screen. The focus should be to ensure an individual can register while walking through the high-street on a mobile phone.

The registration process defines the context to the user. The design of the registration process needs to engage the user as they need to memorise the authentication secret they generate. The focus of the design should be to ensure the registration process feels fast and responsive, engages the users but also ensures users memorise authentication secrets.

Therefore, the registration process and application were designed to be optimised for tablets, rather than be tablet-specific. The aim was to create single screen experiences that work well both on smartphones and tablets.

7.1.1 Authentication Images

The previous iteration of the application relied on a personalised image-set created from an individual's social network. The motivation was that a personalised image-set would be relatively distinct for each user. The aim was to reduce the impact of predictable image choices. This threat that is exasperated when a single image-set is shared among all individuals. Furthermore, a personalised image-set of faces may be more desirable as it has the potential to improve the usability of Tetrad. Ida Gobbini et al. state familiar faces, belonging to friends and family, exhibit a stronger response than famous or unfamiliar faces [121]. Consequently, an image-set of familiar faces would seem a wise choice.

Unfortunately, the process used to generate the personalised image-set in the previous incarnation of the application was costly and complex. The process simply took too much time and was awkward and demanding on resources. Nevertheless, a personalised image-set of faces still had potential. Therefore, the design decision taken for the current incarnation of the application was to use an alternative process to create the personalised image-set of faces.

The primary difference, in process between both applications, is that instead of downloading and analysing profile pictures of friends, the authentication approach relied on tagged images from Facebook. Consequently, the application no longer spent resources analysing images: it simply downloaded them directly from Facebook. The authentication mechanism downloaded at least one tagged image and extracted the tagged segment for use in authentication, up to three if available for a given individual. If no tags existed for a specific friend then they were not selectable. The authentication mechanism would randomly alternate between the images downloaded for a given friend.

7.1.2 Registration Process

The registration process requires users to select 45 friends for use in authentication and select 4 of those friends to form an authentication secret. The image tags associated with the 45 friends on the popular social network Facebook are used to generate images for use in the graphical authentication mechanism.

The previous incarnation of the graphical authentication mechanism relied on images from Facebook. However, image tags were not used; instead images were generated using image processing techniques to locate and extract images of faces for use in the graphical authentication mechanism. The images used in authentication were generated during the registration process.

Therefore, the previous registration process was complex and expensive, in terms of time, resource and user-effort. The previous registration process was similar in that users were expected to select 45 friends for use in authentication and then to select 4 of those friends to form an authentication secret. However, users had to wait until images were generated through the processing techniques before they could complete registration.

Under the previous registration process users were guided through several steps to complete the process. The second step took considerable time as users had to wait for the image to be generated. The users were unable to complete any other steps. Users had no incentive to engage with the registration process. The reality is that users would likely perform some other task as they waited for images to be generated, such as watching television or browsing the web.

Users disconnecting or disengaging with the registration process is concerning, as the registration process is crucial. The user not only needs to generate an authentication secret, they need to memorise it. An individual who is disengaged from the registration process may struggle to complete authentication subsequently. Therefore, the registration process of an authentication mechanism must engage users.

However, the registration process for the current incarnation of the authentication mechanism was still fairly complex. Consequently, effort had to be made to ensure the design engaged users and maintained attention until the process was completed. The design of the registration process had to encompass the following elements: a list of friends, an interface to construct and order the authentication secret, an interface for viewing image tags as well as the image themselves.

Therefore, the user interface for the registration process had to encompass many elements, as well as span small and large screens. A design targeting a windowing environment may have opted for separating each element into separate windows. The user could then manage the windows on a smaller screen or have all present on a larger screen. However, managing several windows on a screen could detract focus from the registration process itself. Nevertheless, the design was not targeting a windowing environment but rather the screen of a smartphone and tablet.

The registration process could be presented, using screens, using various approaches. The three potential approaches identified and considered are:

1. *Single full-screen*
   The user interface consists of a single screen that is contained within the

dimensions of the actual physical screen. Therefore, all elements of the interface would need to be accessible within the screen of the smartphone and tablet. The presentation of all the elements within a single screen may be plausible on a table with a large screen but seems unrealistic on a small smartphone device.

2. *Several full-screens*
   The user interface consists of several screens that the user navigates between or is guided through. The approach is well suited to a small screen device but is awkward for a large-screen device, as there may be considerable wasted space on certain screens that have little content.

3. *Single screen*
   The user interface consists of a single screen that is not contained within the dimensions of the actual physical screen. Therefore, the screen could be very large and contain many elements comfortable. The user would navigate the screen using gestures, such as pans and swipe to see different aspects of the screen.

The initial option seems unrealistic as presenting all the elements within a single screen would make them difficult to use, as well as understand. The second and last option are for all intents, the same approach. Users access different elements either through controllers, such as tab views, or gestures, such as pan and swipe. However, while the approaches are certainly similar the latter option translates better between screen sizes than the second option. The second option still relies on segmenting the process into several screens for the user to navigate, as the screens increase in size, screen space is wasted.

Furthermore, the second approach would require permanent on-screen navigation controls, an undesirable requirement as screen space on small devices is far too valuable to waste on such controls [150]. Moreover, any permanent on-screen object, such as headers, panels, prompts or any other interface 'chrome' is undesirable as it consumes precious screen space [88]. Admittedly, such concerns are less prominent on devices with larger screens, such as tablets. Nevertheless, the design needs to span form factors and a consistent design would be desirable across all devices.

The final approach is better suited for translating between screen sizes. The design of the user interface can remain largely intact as users merely see more of it when using a bigger screen. Therefore, the final approach was selected for the design of the registration process. Consequently, all elements of the registration process were encapsulated within a single user interface that users had to pan and swipe to navigate.

However, Jones et al. argues that designers should reduce the amount of scrolling within an user interface, as it interrupts the primary task [139]. Therefore, the design of the user interface must be throughly considered to ensure that users remain engaged with the process. Moreover, Jones et al. argues users are inclined to scroll vertically rather than horizontally [140].

The interface design of Apple's iOS certainly reflects the expectation that users prefer scroll vertically rather than horizontally. The primary applications

bundled with the operating system, expect users to scroll vertically. The *Mail* application presents a vertical list of email messages for users to scroll through. The *Messages* application presents a vertical list of conversations for users to scroll through. The *Music* applications presents a vertical list of songs for users to scroll through. The *Notes* application presents a vertical list of notes for users to scroll through. Indeed, the operating system rarely expects users to scroll horizontally, one of the few occasions is to navigate between pages of applications on the home screen. However, Apple places bundled applications on the front page, negating the need for users to scroll horizontally in the first instance. Moreover, Apple has promoted the use of a larger screen to display more application icons [16], again, deemphasising the need to scroll horizontally.

Therefore, the design of the user interface needed to be considered throughly to ensure the amount of panning, scrolling and swiping was kept to a minimum. Moreover, the position of each element of the registration process needed to be considered to ensure the correct prominence. The four key elements of the registration process are, as follows:

(a) *Friend list*
The list of the user's friend on the popular social network, Facebook.

(b) *Authentication Secret Editor*
The user interface for adding friends to the authentication secret, as well as determining the position of the friend within the authentication secret.

(c) *Tag Viewer*
The generated images that will be used in authentication.

(d) *Image viewer*
The images used to generate image for use in authentication.

Figure 7.2 illustrates the design of the registration process. The large black rectangle represents the physical screen of an iPad or tablet. The registration process presents elements (a) and (b) when first presented to the user. He/she is able to access elements (c) and (d) by scrolling vertically and horizontally, respectively. The last element, the image viewer, is considered the least important user interface in the registration process and consequently requires a horizontal scroll to access. The third element, the tag viewer, is considered more important than element (d) but less important that elements (a) and (b). Therefore, users need to scroll vertically to access the tag viewer.

The registration process is presented slightly differently on a smartphone or iPhone. The interface and interaction are largely intact on a smartphone. However, a primary difference is that users are initially only presented with element (a) and need to scroll horizontally to access element (b). Users need to scroll either horizontally or vertically from element (b) to access elements (c) and (d). The rest of the design is the same across smartphone and tablet screens.

The four elements in the registration process are discussed in the ensuing sections.

Friend List

Figure 7.2 (a) illustrates the list of friends. The friend list is essentially the base of the entire user interface. The remaining elements of the user interface, elements (b), (c) and (d), all react to selections made in the friend list.

The list presents all the friends an individual has on the popular social network, Facebook. Users are instructed to select 45 friends for use in authentication. The individuals in the friend list may have images associated with them on Facebook, indicated by image tags. When users select an individual from the friend list, associated pictures are downloaded. The images used in authentication are generated by extracting picture segments, according to image tag coordinates.

The registration process does not download all images associated with all friends in advance. The strategy would not only be expensive in terms of time and resources but would involve handling numerous images that are not even used in authentication. Moreover, users tend to have many more friends, than the required forty-five. Therefore, pictures used to generate images for use in authentication are downloaded and processed when required.

Users essentially initiate a subtask when they select an individual from the friend list. A query is then submitted to the social network requesting images tags associated with the selected friend. The image tags are then used to download images and extract segments from them. The image segments are associated with the selected friend and used in authentication. The exact steps involved in the subtask are discussed in §7.2. There also the possibility that the query submitted to the social network may return no results.

Therefore, no images associated with the individual can be used in authentication. Consequently, the individual is not used in authentication and can not be selected.

Authentication Secret Editor

Figure 7.2 (b) illustrates the authentication secret editor. The appearance of the user interface can be compared to the felt top of a *crap table* in a casino. Casino craps is a dice game were individuals gamble on the outcome of rolling pairs of dice. Individuals place bets by laying special chips on designated areas, outlined on the top of the craps table.

Similarly, users create and edit the authentication secret within the registration process by placing chips on special designated areas within the user interface. The user interface presents six designated areas to users, as illustrated in Figure 7.2 (b). The six rectangles are effectively used to create and edit the authentication secret. A chip represents a friend of the user, on the popular social network Facebook.

The board or editor is initially empty when presented. The user generates a 'friend chip' by selecting an individual from the friend list. If the friend is approved, i.e. image tags are associated with them, then an active friend chip is added to the board. The friend chip is automatically placed below the line in the second rectangle.

The six rectangles are used to manage images within authentication and each

rectangle has a distinct purpose. The purpose of the second rectangle is to add friends to the authentication image set, i.e. image tags associated with the specific individual. The box below the second rectangle, containing four additional rectangles, is used to add a friend to the actual authentication secret. The four colour rectangles represent one of four positions within the authentication secret. The orange rectangle is first position in the authentication, purple rectangle the second position, cyan rectangle the third position and the red rectangle the final position. All the rectangles below the line are associated with authentication. The single rectangle above the line, the first one, is used to remove a friend from the authentication image set.

Users long press and drag a chip to reposition it on the board. The chip animates and increases in size to communicate to the user that it can be repositioned. If the user attempts to reposition the chip on a non-designated area, the chip merely repositions itself automatically to the closest, suitable designated area. A friend remains part of authentication as long as the chip is positioned in a rectangle below the line. If the chip is below the line and inside the box, then it is part of the authentication secret.

A friend chip below the line and inside the box remains part on the board, at all times. However, when users select another individual from the friend list, the chip becomes inactive. An inactive friend chip can still be dragged and positioned by users. However, inactive friend chips can only be repositioned within the box, i.e. repositioning a friend within the authentication secret. An inactive friend chip is unable to be removed from the authentication secret or authentication. Therefore, active chips can be placed on any rectangle on the board while inactive chips can only be placed on rectangles within the box, below the line. Users seeking to remove an inactive friend chip from the authentication secret, would need to select the individual associated with the inactive chip from the friend list. Users would then reposition the now active friend chip, on the second rectangle or the first rectangle if they want to remove the friend entirely from the authentication secret.

The process for creating an authentication secret and managing the images used in authentication is fairly complex when contrasted with creation of a traditional alphanumeric authentication secret. Therefore, there may be some concern that users of differing abilities may not be able to create an authentication secret. However, through using lessons of how casinos design props to aide users in playing complex games of chance may make the process accessible to a majority of users.

Tag Viewer

Figure 7.2 (c) illustrates the tag view. The tag viewer presents the images used in the authentication mechanism. An image tag is associated with a larger parent image, the tag coordinates are used to extract the segment from that image.

Users access the tag viewer by vertically scrolling downwards from the authentication secret editor, element (b). The interface and interaction is the same across smartphones and tablets. The user can continue scrolling vertically downward to view all images used in the authentication mechanism.

Image Viewer

Figure 7.2 (d) illustrates the image viewer. The image viewer merely presents all the parent pictures used to generate images for the authentication.

Users access the image viewer by scrolling horizontally to the right. The amount of scrolling required, depends on the dimensions of the physical screen. Smartphone and iPhone users initially select an individual from the friend list, element (a), then scroll horizontally to the right to access the authentication secret editor, element (b). Users then continue scrolling horizontally to the right, past the authentication secret editor to the image viewer. Similarly, tablet and iPad users select a friend from the friend list and simply scroll horizontally to the right, once, from the authentication secret editor.

The user then continues scrolling horizontally to the right to continue viewing more pictures. The user can return to the authentication secret editor by simply scrolling horizontally to the left, past all the previously viewed parent pictures. Smartphone and iPhone users need to perform an additional horizontal scroll to the left, from the authentication secret editor to access the friend list.

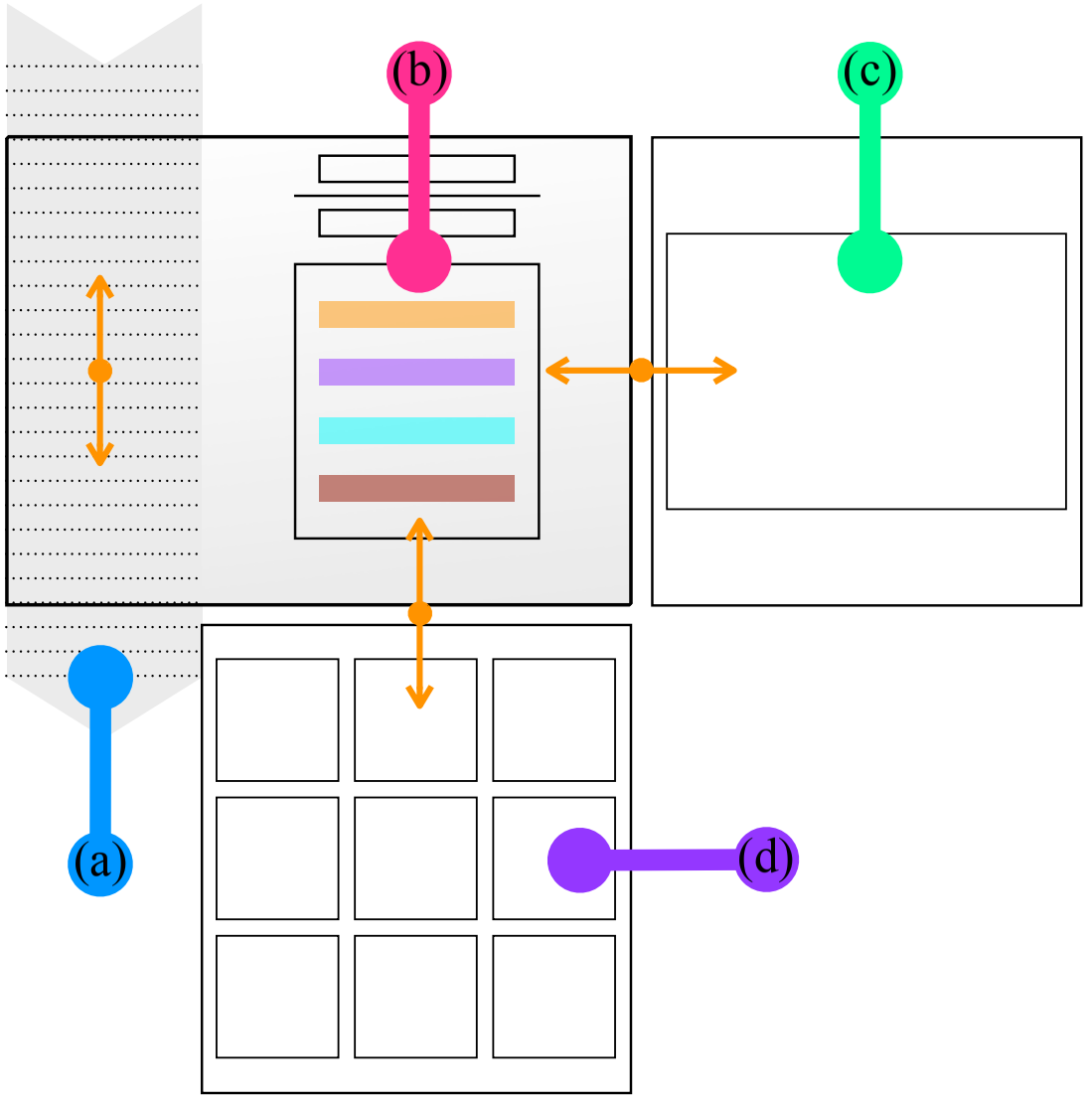

Figure 7.2: An overview of the design of the registration process.

7.1.3 Application

The application was to be revised as an iPad application and the initial design had to be scaled to fit the iPad and iPhone. However, when considering the initial design there were many aspects that seemed less than optimal when revisited and scrutinised. There were four aspects that were less than optimal when refreshing the design for iPads. The four aspects were:

- *Document Browser*
  The document browser contains all the lecture recordings an individual has downloaded to their device.

- *Lecture Store*
  The lecture store is essentially a list of all the available lecture recordings for download.

- *Download Queue*
  The download queue represents the progress of all the lecture recordings currently downloading.

- *News*
  Important course information and news to keep student up-to-date.

Document Browser

The document browser allows users to browse all the lecture recordings they have downloaded. The user swipes right to left to browse through downloaded lecture recordings. Each lecture recording is represented as a piece of paper with an image of an audio player in the centre. The scaled down piece of paper is positioned in the centre of the screen, with the title and author displayed underneath. A button is located underneath each lecture recording, labelled 'Open'. The user taps the button to open the lecture recording. Consequently, the piece of paper animates to fill the screen and the audio player with corresponding lecture recording is loaded. The user taps the button in the centre of the audio player to listen to the lecture recording.

The interaction is acceptable when a user is managing a few lecture recordings but as the list grows it becomes unwieldily. An individual seeking to open the penultimate lecture recording, for example, in a list of thirty, would need to make several swipes. The process is tiresome and inefficient, considering a simple list of lecture recordings could vastly improve navigation as it would reduce the amount of interaction to required to find and listen to a lecture recording.

The initial design, on reflection, appears to be very much style over substance, valuing appearance rather than function. Moreover, the interface is not particularly stylish or even visually informative. The only information that changes as the user swipes left and right through lecture recordings is the title and audio of each recording. Such information may change only slightly: a lecturer will often give several lectures on the same topic, e.g. 'Cognitive Neuroscience 1', 'Cognitive Neuroscience 2', 'Cognitive Neuroscience 3' etc. In these scenarios the only changing aspect as the user swipes left and right is a single number. Therefore,

the process can be monotonous and does not really harness the power of the device, several pixels and screen wasted for no reason at all.

Therefore, the design of the document browser was not stylish, visually informative or particular inefficient. The design was cumbersome, awkward and tiresome and there no real benefit in continue to use it, never mind transition the design to the iPad.

News

The news component of the application was an additional feature, an incentive designed to promote frequent use of the application. However, while other features, such as chaptering and annotation, were inline with the central workflow of downloading and annotating lecture recordings, the news feature seemed misaligned.

The lecture store section of the application featured a news carousel. The feature was relatively inoffensive but it was debatable if the feature offered enough benefit to justify the screen space it consumed. The School of Psychology communicated with the user-base across a variety of channels and was keen to reach students in as many ways as possible. However, inserting news into every space is not desirable, especially at the expense of user experience. Moreover, removing the news carousel is unlikely to have an impact on distribution of news, as the client communicates across so many different channels.

The client was also very aware that the addition of news made the application bloated and made it deviate from the central workflow. The School of Psychology did not want the application to replace the portal, or even constitute a mobile version of it. The application was envisaged as a compliment to existing services and resources, not a replacement. The client simply wanted to provide a more efficient and elegant solution to the problem of accessing lecture recordings on a mobile device.

The news carousel was not simply a feature, additional or otherwise, that was required to achieve the primary aim of the application. Moreover, it consumed valuable space that could be used to list more lecture recordings.

Lecture Store

The aim of the lecture store was distribution of lecture recordings. However, there was some concern about referring to the distribution aspect of the application as the 'lecture store'. The word 'store' has many connotations but the client was concerned about students viewing the interface, as an actual store akin to the iTunes or Amazon Store. Students may interpret use of the word as a signal that The School of Psychology plans to start charging for lecture recordings in the future.

The word 'store' is clearly potent in the world of modern consumer electronics, as Apple and Amazon are battling over the use of it [258]. Moreover, Google recently rebranded the 'Android Market' to'Google Play Store' [104]. The renaming is not surprising as the word 'market' has inappropriate and undesirable connotations, i.e. users should not expect service and support as the stall holder

may not be around the following week. The word 'store' on the other hand has connotations of curated and considered content, where users can receive service and support if they encounter problems. Therefore, the word is important.

However, the School of Psychology was not running a store, not even a lecture store. The client simply wanted to provide students with lecture recordings, incase they missed a lecture or needed to revisit a topic. The store of lecture recordings would not scale beyond the School of Psychology. Therefore, the name is not only undesirable it is inaccurate. Consequently, while the feature was referred to as the 'lecture store', it was decided the name itself would not appear in the application.

Therefore, on the navigation tab the lecture store was accessed through the tab labelled 'Featured', similarly labelled tabs have been used in the iTunes and App Stores, for perusing the latest content. Students were expected to tap the tab, browse the list of lecture recordings and select the lecture they wanted to download. The file was then downloaded to the user's device. The process was similar to that of purchasing a song from the iTunes Store. Therefore, while users may not see the word 'store', they encounter a similar user experience. The word had clearly influenced the initial design of the application. The design treated students, as potential customers when in actual fact they were essentially *subscribers*.

Therefore, when refining the design to cross between the iPhone and iPad, basing the design of the lecture store on a subscription-based application may be a better direction. There are several examples applications where the user is essentially a subscriber, they are:

- *Netflix*
  Netflix is a movie subscription service that requires individuals to pay a monthly fee to access a range of television programmes and films. The Netflix iOS application presents users a single view, containing a list of content. Users tap an element in the list to stream the selection to their device.

- *Podcasts*
  The Podcasts application is a client for listening and viewing podcasts on smartphones and tablets. The application present a single view, containing all the active podcast subscriptions. The application automatically downloads the latest episode of a podcast to the user's device.

- *Metro Newspaper*
  The Metro Newspaper is a daily newspaper that users can view on smartphones and tablets. The Metro Newspaper applications presents users a single view, containing a list of issues. Users download with a single tap to a list element. The same list manages all issues: when one is downloaded, the button label changes from 'Download' to 'Open'.

The subscription-based applications all share the same simplicity of a single view of content but differ when it comes to consuming that content. Streaming and automatic downloads were undesirable options as they may lead to inadvertent

use of expensive mobile data [22]. Therefore, the design used by the digital newspaper seemed a suitable solution to base the refined design on. The only remaining aspect to be revisited was managing downloads.

Download Queue

The user experience of downloads was particularly poor in Dick. Students could download as many lecture recordings from the lecture store as they wanted. The lecture recordings were all added to a single download queue. The progress of the download queue was presented within the lecture store, while the queue was being processed. The progress queue had a cancel button to the right of it.

The management of downloads essentially consisted of the single progress bar and cancel button. Lecture recordings in the download queue were labelled as 'Downloading'. However, users had no insight into state of each download, users did not know the progress of a specific lecture recording or if it had even commenced. Moreover, users were unable to cancel individual downloads, they could only cancel the entire queue of downloads.

Furthermore, asides from actual management of the downloads, presentation of the progress bar only further restricted the space for browsing available lecture recordings. The combination of the progress bar, navigation area and news carousel, essentially leaves little space for browsing or view available lecture recordings. There is a poor use of screen space in the current design.

Therefore, the refreshed design needed to resolve the issues surrounding download management.

Proposed Solution

The application from the outset was separated into two segments that the user navigated between. The first segment belonged to the user. The second second segment belonged to the School of Psychology. Consequently, the first segment contained the document browser, as well as the chaptering and annotation tools. The second segment contained the lecture store and news carousel.

However, reconsideration of the initial design reveals that such separation was needless. The main insight was that students were initially treated as customers, when in actual fact they were more like subscribers. Therefore, a redesign of the application is better based on a digital newspaper application or similar subscription-based service, than a digital store.

Figure 7.3 illustrates the redesigned interface for the iPhone and iPad. The first step was to remove the news carousel and refocus all attention back on content, i.e. the lecture recordings. The new design consolidated all aspects of the application, i.e. document browser, lecture store and download manager, into a single view.

The view presented all available lecture recordings to user. The list was updated every time the application was opened. Each lecture recording is represented in a similar fashion to the initial design with a major difference: icons now offer live previews of all user-generated chapters. The user can now at glance see

the lecture recordings they have chaptered. The lecture recordings not on the device appeared opaque, indicating they were not downloaded.

The application did not download lecture recordings automatically due to the following concerns: (1) expensive mobile data charges and (2) sparse onboard device storage. Instead users were required to initiate all downloads but they could download as many as they wished. Users initiated download of a lecture recording by tapping an opaque icon. A download progress bar would then be overlaid on top of the icon. The user can initiate download of a lecture recording and listen to another recording at the same time. Users opened a downloaded document by pinching on an icon or simply tapping it.

The refreshed design for iPad and iPhone consolidated all aspects into a single view. The design no longer wasted space on navigation controls or chrome to explain views and interfaces. The user no longer needed to swipe between documents: instead a swipe moved between multiple documents. The entire focus and user interaction was now content, on the lecture recordings.

Figure 7.3: The refreshed design of the application for the iPhone and iPad.

7.2 Implementation

The implementation is outlined in the following sections.

7.2.1 Registration Process

The registration process for the authentication mechanism is as follows:

1. *Enter Facebook Credentials*
   The initial step, as was the case with Dick, was for users to allow the registration process access to their Facebook account. The registration process, once connected, submitted a query to Facebook requesting a list of all the user's friends on the social network.

2. *Generate Authentication Secret*
   The second step in the registration process requested the user to create an authentication secret, using images of their friends downloaded from the popular social network, Facebook. The friend list received in the previous step is parsed and presented to the user in a table, each table entry contains the profile image and name of each friend.

   The user is instructed to select 45 friends from the friend list for use in authentication. The user simply taps a table view entry to select a friend. The user selects 45 friends before they can complete the registration process.

   However, unlike the previous version of the registration process the user is not guided through separate stages to generate the authentication secret. The user is able to alter and edit the authentication secret all within the second step. Section 7.1.2 outlines the design of the user interface for the registration process. The specific user interface for managing and editing the authentication secret is to the right of the table view.

   Once the user has selected 45 friends from the friend list and indicated the four friends that comprise the authentication secret, the user is permitted to complete the registration process. The table view automatically scrolls to the bottom and reveals a button to complete the registration process. The user can press the button to complete the registration process or they can make additional alterations and then press the button to complete the authentication process.

Image Processing

The image processing component is different from the previous incarnation of the application. There are several steps that occur when an individual selects a friend for inclusion in the authentication mechanism.

1. *Query Social Network*
   The initial step in the image processing subtask is to query Facebook for all image tags associated with the selected friend. An FQL query is submitted to Facebook, requesting all the image tags associated with a specific user identification string.

2. *Approved for inclusion*
   If the response returned from Facebook in the previous step is not empty, then the friend is approved. The approval is communicated to the user. However, if the response from Facebook is nil, the user is informed the friend was not successfully selected and they will have to make another selection.

3. *Parse response*
   The returned response from Facebook is then parsed to generate an array of objects, each object comprises of an image URL and a set of tag coordinates.

4. *Download Images*
   The number of tags returned for each friend varies. The implementation used a specific threshold value, ensuring only a few images were downloaded for each friend. While not the case in the current implementation, there is no reason why the threshold value could not be based on the current device or available resources, e.g. a powerful device with a strong Internet connection could have a higher threshold value and consequently download more images.

5. *Extract tags*
   The image tags associated with the selected individual are then used to extract a segment of the downloaded image. The tag coordinates identify the segment to be extracted.

6. *Write to disk*
   The downloaded image as well as the segmented extracted images are both written to disk.

7.3 Evaluation

The evaluation is outlined in the following sections.

7.3.1 Subjects

The application was distributed and evaluated with undergraduate students enrolled at The School of Psychology at the University of Glasgow.

7.3.2 Apparatus & Material

The apparatus and materials are similar to the evaluation of Tom, see §5.3.2.

7.3.3 Procedure

The procedure is similar to the evaluation for Tom, see §5.3.3.

7.4 Results

The results are discussed in terms of the registration process and authentication process.

7.4.1 Registration

There were 12 registrations, over 41 days. The majority of these (92%) occurred in the first 5 days and only a single registration occurred after 30 days.

The application did not collect personal information, such as an email address, or request individuals to create a username. The previous incarnations of the application, namely Tom and Dick, relied on the device UDID as the individual's username. The intention was to continue with the same approach in Harry. However, Apple Inc amended guidelines and policies that advocated against the use of such tactics due to privacy concerns and instead favoured generation of a temporary UDID. Therefore, Harry continued to rely on an alphanumeric string for the individual's username but this was a temporary UDID and not the device UDID.

Furthermore, as with Tom and Dick there was no recovery process. Users unable to authenticate were expected to delete the application, download it again and conduct the registration process again. The device UDID was previously used to determine if users had completed the registration process more than once. However, as the device UDID could no longer be used, tracking was curtailed. However, Harry, as unlike previous iterations of the application, collected additional information such an individual's device type, data connection type, the registration code used and the number of friends they had on Facebook. Moreover, Harry also had no limit on unsuccessful authentication attempts. Therefore, users were not forced to re-register because they made too many unsuccessful authentication attempts. The interpretation of the registration records coupled with the fact that the majority of registrations occurred within 5 days, suggested only a single individual re-registered. Consequently, the assumption is that there was 11 distinct users of the application.

50% of registrations were initiated between midday and six o'clock in the evening. Indeed most registrations (83%) were initiated between midday and midnight. Some registrations (8%) occurred during the night between midnight and six o'clock in the morning. The remaining registrations (8%) occurred between six o'clock in the morning and midday.

The assumption was that users would favour inexpensive wireless Internet connections over potentially slower and more expensive cellular data connections. The strength and speed of the data connection was important as images were being downloaded over the data connection. The majority of registration attempts (58%) were completed over a Wireless Local Area Network (WLAN), essentially a 'WiFi' connection. However, many registration attempts (42%) were completed over a Wireless Wide Area Network (WWAN), essentially a cellular connection. Therefore, many users felt comfortable using a cellular connection.

There was interest in whether the user group has been soured by the prior experience of the authentication mechanism. The registration codes used indicated

the year group of the user. The majority of registration attempts (67%) were generated by first-year undergraduates. The remaining registration attempts were generated by second-year (25%) and third-year (8%) undergraduates. There were no registration attempts generated by fourth-year undergraduates as Harry was released at the start of an academic session, the majority of users were new to the application and authentication mechanism.

The Harry application supported the entire ecosystem of Apple products, i.e. iPhone, iPod and iPad. All individuals used iPhones with no individual using an iPod or iPad to complete the registration process. The application, much like Dick, required individuals to have an active Facebook account to populate the authentication mechanism with images. Consequently, there was concern that individuals might not have enough friends to use the authentication mechanism, as 45 slots had to be populated and not all friends would be suitable. The registration process logged the friend count for each individual. The average number of friends for each user was 344 ($SD = 130.987$). However, there was wide variation in the number of friends that each user had on the popular social networking service (minimum = 157, maximum = 548, range = 391). Nevertheless, it seems that most users had more than enough friends.

Time Taken

Individuals spent a total of 15537.06 seconds or approximately 4.31 hours on the registration process. The average time taken to complete the registration process was 21.58 minutes ($SD = 29.472$). However, the mean may not be indicative of overall performance as there was wide variation in the time taken to complete the registration process (minimum = 6.70, maximum = 113.71, range = 107.01). The time taken, on average, although arguably still lengthy is a considerable improvement on the average previous performance using Dick.

The majority of registrations (75%) took less than 15 minutes to complete. Inspection of registration times revealed that average performance was potentially distorted by an extreme outlier, a single individual that took 113.71 minutes or 1.90 hours to complete the registration process. Therefore, the extreme outlier was removed from consideration. The average time taken to complete the registration process was 13.20 minutes ($SD = 5.428$). However, the mean may not be indicative of overall performance as there was wide variation in the time taken (minimum = 6.70, maximum = 24.79, range = 18.09).

The time taken may be reduced if only initial registration attempts are considered. The average time taken to complete the registration process for each user was 13.45 minutes ($SD = 5.674$). The variation in performance was not as dramatic as previous incarnation of the application (minimum = 6.70, maximum = 24.79, range = 18.09). The reality is that the registration process still took time to complete but was vastly improved over the previous incarnation, Dick.

However, another aspect that was difficult to determine over the previous incarnation, Dick, was the data connection used to complete the registration process. The previous application did not record the data connection type. However, the type of data connection could impact on registration times as the images had to be downloaded from servers. Therefore, considering only initial registration

attempts, the typical time taken for each data connection was determined. The average time taken to complete the registration process over a wireless LAN connection was 13.01 minutes (SD = 6.117). There was wide variation in the time taken to complete registration among users (minimum = 6.70, maximum = 24.79, range = 18.09). The average time taken to complete the registration process over a cellular connection was 34.04 minutes (SD = 44.81). However, there was wide variation in the time taken to complete registration among users (minimum = 9.85, maximum = 113.71, range = 103.86).

Users took longer to complete the registration process on a cellular connection than on a wireless LAN connection. However, ultimately in terms of time to complete the specific data connection type is not necessarily as relevant as the speed and strength of the data connection. Nevertheless, registration time was indeed impacted by the use of cellular connections in this instance.

Target Image Selection

The user interface for the registration process allowed individuals to select distractor and target images from a single interface rather than from several steps. The assumption is that individuals would select target images as they selected distractor images. The distractor count was logged when individual confirmed an image as a target image. The reality is that for the majority of registrations (58%), individuals selected all the distractor images and then determined target images. In another registration, an individual selected all the target images before selecting all the distractor images. The majority of individuals simply created steps within the single interface. Consequently, users did not mix interaction by selecting some distractors then a target. Therefore, staged, stepped interaction for creating an authentication secret may be more than enough.

7.4.2 Authentication

Users made a total of 111 authentication attempts over a period of 77 days. The vast majority of authentication attempts (60%) were unsuccessful. However, several authentication attempts (40%) did end in success. The majority of successful authentication attempts used a horizontal alignment (60%), several used a vertical alignment (38%) and one authentication attempt used a diagonal alignment.

The average number of authentication attempts associated with each registration or account was 9.25 (SD = 5.276). Nevertheless, there was a wide spread in the number of attempts associated with each account (minimum = 3, maximum = 19, range = 16). The number of successful authentication attempts associated with each account was 3.75 (SD = 3.415). However, a few accounts had far more successful authentication attempts (minimum = 0, maximum = 12, range = 12) associated with them. The number of unsuccessful authentication attempts associated with each was 5.50 (SD = 3.826). Similarly, a few accounts had far fewer unsuccessful authentication attempts associated with them (minimum = 0, maximum = 12, range = 12). However, authentication attempts can only be considered in terms of users rather than simply accounts.

A total of 11 users made at least a single authentication attempt. If only initial registrations are considered then there are 100 authentication attempts. The average number of authentication attempts was 9.09 (SD = 5.504). The number of authentication attempts generated by users was varied (minimum = 3, maximum = 19, range = 16), suggesting the distribution of authentication attempts among users was more varied than indicated by the mean.

Furthermore, although all users made made an authentication attempt, not all authentication attempts were successful. The average number of successful authentication attempts was 3.91 (SD = 3.534). There was a large span in the number of successful authentication attempts (minimum = 0, maximum = 12, range = 12). However, many users submitted far more unsuccessful authentication attempts.

The average number of unsuccessful authentication attempts was 5.18 (SD = 3.842). However, the mean may not be indicative of overall performance for most users as there was a wide variation in the number of unsuccessful attempts among users (minimum = 0, maximum = 12, range = 12). Therefore, it would appear users submitted more unsuccessful authentication attempts that successful attempts. However, a few users (18%) did not submit any unsuccessful authentication attempts. There were several unsuccessful authentication attempts generated by individuals and the composition of these are discussed in §7.4.2.

The previous incarnations of the application, i.e. Tom and Dick, both had rogue authentication attempts that were not associated with any registration record. No rogue authentication attempts made with Harry: all authentication attempts were made with the application were associated with a specific registration record. This suggests that inconsistencies and flaws in the previous registration processes may have allowed for the generation of rogue authentication attempts.

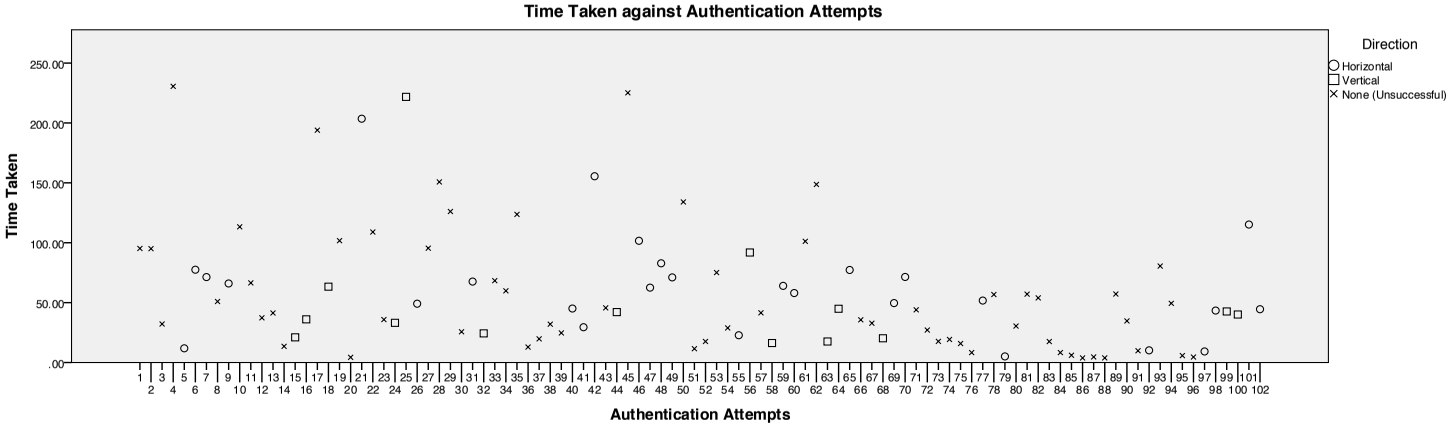

Figure 7.4: Scatter plot illustrating authentication times for successful and unsuccessful authentication attempts.

Time

Individuals spent a total of 145546.83 seconds on 111 authentication attempts or 1.68 days. The average time for an authentication attempt was 1311.23 seconds (SD = 8732.269). However, there was wide variation in the time spent on each authentication attempt (minimum = 3.83, maximum = 70325.84, range = 70322.00). There were several authentication attempts that took considerable time.

There were two extreme outliers lasting 59964.53 and 70325.84 seconds or 16.66 and 19.53 hours, respectively. The two authentication attempts were successful, one vertical the other diagonal. Both of these distort typical performance. Therefore, both extreme outliers were removed. Consequently, the average time taken for an authentication attempt was 139.97 seconds (SD = 569.438). However, there were still several extreme outliers that, while not as severe as the aforementioned, could potentially distort typical performance. Consequently, 7 authentication attempts that were greater than 250 seconds were removed from consideration.

Figure 7.4 illustrates a scatterplot of authentication times, displaying successful horizontal and vertical authentication attempts as well as unsuccessful authentication attempts below 250 seconds. The average time taken for authentication attempts was 57.91 seconds (SD = 51.00). However, there was a wide variation in time taken for authentication attempts (minimum = 3.83, maximum = 230.57, range = 226.74). The authentication time may be improved when considering the success of an authentication attempt.

The average time taken for successful authentication attempt was 59.34 seconds (SD = 46.822). However, there was wide variation in the time taken on successful authentication times (minimum = 5.10, maximum = 221.80, range = 216.70). Furthermore, vertical authentication attempts (M = 51.09, SD = 53.149) appear shorter than horizontal authentication attempts (M = 63.61, SD = 43.644). Overall, unsuccessful authentication attempts (M = 56.96, SD = 53.993) appear to take longer than horizontal and vertical authentication attempts.

Nevertheless, there were far more unsuccessful authentication attempts than successful. There could be several reasons for an individual generating an unsuccessful authentication attempt.

Composition of Unsuccessful Authentication Attempts

An interesting aspect of Harry was that there was no restriction on the number of unsuccessful authentication attempts that users could submit. Furthermore, the application, unlike previous incarnations Tom and Dick, logged the composition of the submitted authentication attempt. Therefore, unsuccessful authentication attempts could be probed and inspected to determine the root of the problem.

The expectation is that an unsuccessful authentication attempt would comprise of a grid with target images scattered across it, unconnected and completely incorrect. However, few unsuccessful authentication attempts (11%) actually approach this scenario. Figure 7.5 illustrates a pie chart displaying the number of

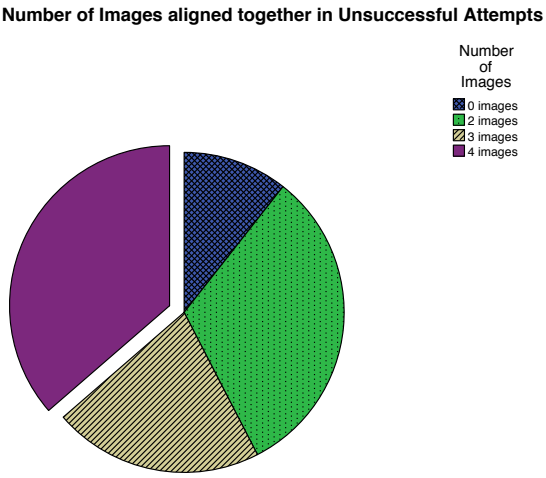

Figure 7.5: Pie chart illustrating the images aligned together in unsuccessful attempts.

images aligned in unsuccessful authentication attempts.

Many unsuccessful authentication attempts (36%) contain 4 images that are aligned but are out of sequence. There were several unsuccessful authentication attempts (21%) that contained 3 images aligned and many unsuccessful authentication attempts (32%) contained 2 images that were aligned.

Therefore, sequence appears to be an incredibly important factor in the failure of an authentication attempt. In many of the unsuccessful authentication attempts the problem is not necessarily memorability with the images but the memorability of a specific sequence. The majority of unsuccessful authentication attempts are the product of an inability to recall a sequence not the images themselves.

Nevertheless, those unsuccessful authentication attempts that contain 2 and 3 images aligned suggest an individual was unable to authenticate due to images. However, many unsuccessful authentication attempts compromise of sessions of several authentication attempts that often conclude with a successful authentication attempt. Users were not limited to a number of unsuccessful authentication attempts. Consequently, in some cases, users made several unsuccessful authentication attempts before concluding the process with a successful authentication attempt. The process suggests that users were refreshing the grid of images, awaiting the arrival of an image they recognised to authenticate.

Time Difference between authentication approaches

An independent-samples t-test was performed to determine if a difference existed in the time taken to complete an authentication attempt between Tom and Harry. Therefore, the time taken for the initial authentication attempt from 10 users of each version of the authentication mechanism was used. Figure 7.6 illustrates box plots of the data and suggested there was no outliers. The authentication

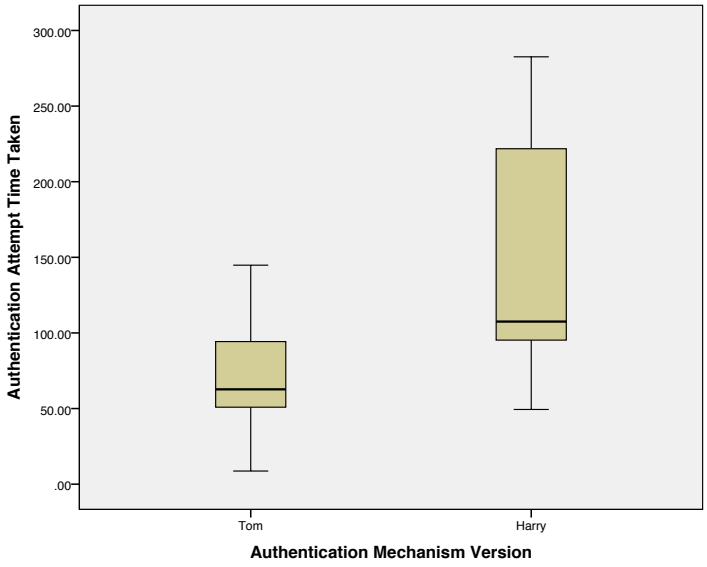

Figure 7.6: Box plot of time taken for initial authentication attempts using Tom and Harry.

times for each group were normally distributed as suggested by a Shapiro-Wilk's test ($p > .05$). The assumption of homogeneity of variances was violated as suggested by Levene's Test for Equality of Variances ($p = .006$). Consequently, a Welch-Satterthwaite correction was used. The time taken to complete an initial authentication mechanism with Tom was ($M = 68.90$, $SD = 39.208$) shorter than Harry ($M = 145.60$, $SD = 78.278$). The initial authentication attempt times for Tom were 76.71 seconds ($SE = 27.685$) shorter than those on Harry. The difference was statistically significant $t(13.249) = -2.771$, $p = .016$.

However, several authentication times are far longer in the Harry group than in the Tom group. Furthermore, there were also much shorter authentication times in the Tom group that may distort typical performance. Therefore, if only authentication times below 120 seconds and above 30 seconds are considered there are 7 authentication times for each authentication mechanism. A box plot of data revealed an extreme outlier in the Harry group was 3 box lengths from the box edge. Consequently, the case was removed from consideration and the longest authentication attempt was removed from Tom to produce 6 cases in each group.

An independent-samples t-test was performed to determine if a difference existed in the time taken to complete an authentication attempt between Tom and Harry. Figure 7.7 illustrates box plots of the data and suggested there was no outliers. The authentication times for each group were normally distributed as suggested by a Shapiro-Wilk's test ($p > .05$). The assumption of homogeneity of variances was not violated as suggested by Levene's Test for Equality of Variances ($p = .689$). The time taken to complete an initial authentication mechanism with Tom was ($M = 68.73$, $SD = 17.604$) shorter than Harry ($M = 88.77$, $SD = 9.189$). The initial authentication attempt times for Tom were 20.03 seconds

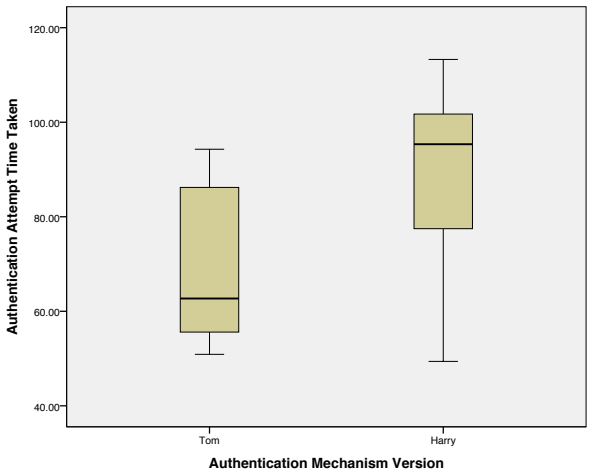

Figure 7.7: Box plot of time taken for initial authentication attempts using Tom and Harry that are under 200 seconds and extreme outliers removed.

(SE = 11.666) shorter than those on Harry. The difference was not statistically significant $t(10)$ = -1.717, p = .117.

Images on registration

7 registrations were completed over WiFi and 5 registrations completed over WWAN. The duplicate registration performed over WiFi were removed from consideration as well as outliers from both groups. Consequently, there was a balanced, albeit small sample of 4 individuals in each group or 8 cases overall.

An independent-samples t-test was performed to determine if a difference existed in the time taken to complete an authentication attempt between Tom and Harry. Figure 7.8 illustrates box plots of the data and suggested there was no outliers. The authentication times for each group were normally distributed as suggested by a Shapiro-Wilk's test ($p > .05$). The assumption of homogeneity of variances was not violated as suggested by Levene's Test for Equality of Variances (p = .898). The time taken to complete an initial authentication mechanism was shorter for those individuals who completed registration over WiFi (M = 117.88, SD = 75.210) than than those who completed registration over WWAN (M = 193.95, SD = 86.012). The initial authentication attempt times for those individuals who completed registration over WWAN was 76.07 seconds (SE = 57.12) higher than those individuals who completed over WiFi. However, the difference in authentication time among the two groups was not statistical significant, $t(6)$ = 1.332, p = .231.

7.5 Discussion

The purpose of the application was to evaluate an alternative authentication mechanism that relied on a personal image collection that alternated during au-

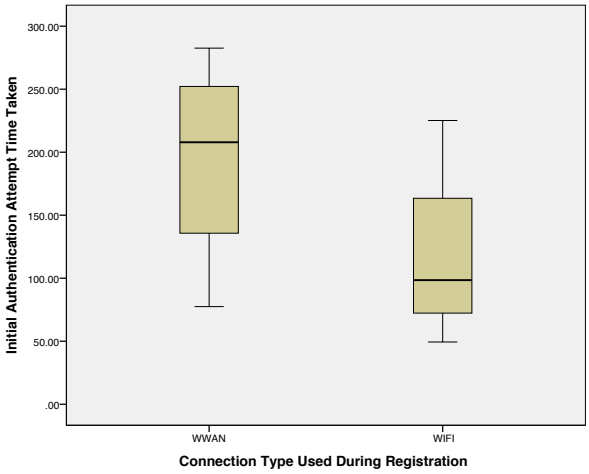

Figure 7.8: Box plot of time taken for initial authentication attempts for individuals who completed registration over WWAN or WiFi.

thentication. The authentication mechanism relied on a Facebook account, much like the previous incarnation, Dick. The user was expected to have an active Facebook connection with at least 100 friends. The user was required to create an image set that comprised of target and distractor images. The user simply selected friends from their friend list. The application would then download at least one of image of the friend tagged on Facebook, up to three if more were available. If there was no tagged images available the friend was not selectable.

The image set comprised tagged images of friends the user selected from the friend list. Therefore, the registration process was more agile and efficient from the previous incarnation of the application as no analysis was performed on images. Instead the registration process merely downloaded tagged images. However, while the application was no longer as demanding as previous incarnations, on reflection the alternative authentication mechanism expected a lot more of users during registration than a traditional authentication approach. The user had to have an active Facebook account with 100 friends as well as active Internet connection and several minutes to complete the registration process.

Therefore, while the application attracted more users than Dick, it was unpopular with the user base. There were clearly many users who felt either the use of Facebook images was 'creepy' or the registration took too long. Alternatively, many users may have simply felt that the registration process and authentication mechanism was too demanding for an application that some felt should not have had an authentication approach in the first instance.

Therefore, analysis could only be performed on individuals that were not alienated by the demands of the authentication approach. The key area of interest was if there was any impact on authentication times from the initial authentication approach used in Tom. The analysis revealed that there was no significant impact on the time taken to complete an authentication attempt between the initial version of the authentication mechanism and the version used in Harry. However, it should be noted that the sample was small and this makes the weight

of such analysis debatable.

Furthermore, while the time taken may not have been impacted, users made several unsuccessful authentication attempts. Therefore, users clearly still struggled with the alternative authentication approach and the alternating images. Many of the logs suggest the impact may result from expecting users to memorise a sequence as well as images. There were many unsuccessful attempts that were the product of an incorrect sequence rather than alignment of incorrect images.

The other concern was that users who completed the registration process over a cellular connection may have been subjected to lower quality images as a result of traffic shaping. The statistical analysis revealed there was no significant difference in the time taken to complete an authentication attempt between those who registered over a cellular connection or wireless connection. Therefore, either traffic shaping was not a problem in this instance or the quality was acceptable enough for users. However, the sample used in the analysis was small and as such makes the weight of any analysis debatable.

Nevertheless, traffic shaping policies depend on the network and operator. Therefore, while traffic shaping does not appear to have impact in the current evaluation, it is still an important issue to investigate. The real world aspects of traffic shaping are something that is rarely discussed, even investigated in the realm of graphical authentication research. The issue would need to be researched and discussed further with possible solutions devised to overcome it. Otherwise, traffic shaping could be another item on the ever growing list of reasons as to why graphical authentication mechanisms are simply not practical, especially with an increasingly mobile user base.

7.6 Conclusion

The chapter outlined the design, implementation and evaluation of an alternative authentication mechanism that relied on an alternating image set. While the authentication mechanism used in Harry was arguably more successful than that used in Dick, the mechanism was still too demanding, on reflection. The requirement to select distractors as well as target images did not appear to have an impact on authentication time. However, users ultimately made many unsuccessful authentication attempts, suggesting that they struggled with the alternating images. Nevertheless, logs revealed that sequence may be more important that first thought, as several unsuccessful attempts were the product of incorrect sequence rather than alignment.

The effect of traffic shaping may not be as important as first thought, as there was significant difference in the time taken in authentication attempts between users who completed registration over a cellular connection and a wireless connection. However, the sample used for analysis was small and the traffic shaping policy very much depends on the network and operator.

The reality is that Tetrad continued to represent a conflict for most users, evident from the low number who actually used the current application. Efforts made to improve Tetrad in the current iteration of the application were sim-

ply not enough. Consequently, Tetrad was unsuccessful in all three application iterations and all field tests. Therefore, the relevancy of the alternative authentication mechanism must be scrutinised as it would appear that any practical implementation of the authentication mechanism is not suitable for most candidate applications. The relevancy of the authentication mechanism as well as all three implementations are outlined and discussed in the next chapter.

Chapter 8

Discussion

The alternative authentication mechanism deployed in all our evaluations was indeed throughly considered. The design of the alternative authentication mechanism was research driven. The approach was, at heart, knowledge-based and merely required an individual to memorise the authentication secret. The knowledge-based approach relies on images as they have superior memorability than the words that name them. The images used were faces, due to the inherent expertise all individuals possess for recognising them. Lastly, the authentication mechanism was a searchmetric, i.e. it was fundamentally recognition-based as recognition is one of the most effective ways to probe memory. Consequently, the authentication mechanism literally presented all the pieces of a user's authentication secret and required users to align them. The mechanism was designed to afford users the ability to align them without fear of onlookers observing the actual authentication secret.

The authentication mechanism was the product of typical design processes. A concept was mapped out, prototypes built, variations developed and explored. The process was organic and all elements were considered. The initial shared-space prototype was assessed in a controlled evaluation. The authentication mechanism was initially well received and showed promise, as users responded positively.

Nevertheless, when the authentication approach was deployed in field evaluations, i.e. in an actual context of use, many problems became apparent. Therefore, while the design of the application may have been driven by research, the result was an authentication mechanism that was poorly received and labelled as 'creepy' and 'terrible'. There were several problems with the alternative authentication mechanism, the three primary problems being:

- *Inconvenience*
  There was wide variation in the time taken to complete authentication attempts. Even worse, authentication times did not appear to improve over time. Authentication attempt time generally remained unacceptably high.

- *User choice*
  Some images were considerably more popular than others, confirming that users make similar choices and thus undesirable predictability.

- *Interaction*
  There were several unsuccessful authentication attempts suggesting poor understanding of the actions used to interact with the authentication mechanism.

There are several potential explanations for the aforementioned problems. However, before such explanations are offered and discussed, many may argue that the alternative authentication mechanism did not need to be deployed into the field to uncover these issues.

Consequently, the merits of controlled and field investigations at uncovering each of the aforementioned problems is discussed in the ensuing section, §8.1. The thesis statement is then reviewed and accepted or rejected, §8.2. Then the potential problem sources are discussed, §8.3, before concluding that observation-resilience is the root of many problems. The discussion continues with the notion that observation-resilience is may not be as important as once thought and alternatives are proposed, §8.4. The discussion concludes by outlining aspects of the field investigations that could have been improved.

8.1 Controlled vs Field Investigations

Controlled evaluations could reasonably be expected to have uncovered many of the problems the field studies uncovered. If this is so, the motivation for several field studies is not necessarily clear. There are some compelling reasons why controlled evaluations would not have delivered these insights. It is not clear that the same problems would have been uncovered in controlled evaluations. Therefore, each problem is reviewed in terms of a controlled evaluation uncovering issues associated with it.

8.1.1 Inconvenience

The time taken to complete an authentication attempt could have been captured in a laboratory setting. There are several research papers that report and discuss authentication times captured through controlled evaluations. However, the ecological validity of such evaluations is questionable and subsequently the weight of such captured metrics debatable. The authentication mechanisms in question, are often evaluated in isolation, devoid of context and executing on non-target devices. Such laboratory evaluations frequently substitute mobile computers for paper and pen or Java-based web browser mockups. It is not clear how mouse interactions or pen movements mimic that of touch on a smartphone. Moreover, the silent and steady laboratory environment may not be representative of the actual expected environment, such as a train-carriage or bus-stop.

Therefore, any captured metrics have to come with the caveat that they were recorded in ideal conditions, not actual or realistic conditions. However, in fairness, many controlled investigations of authentication mechanisms do state that further research in the field is required. Unfortunately, follow-up field investigations are rarely, if ever, reported. Consequently, as authentication times in one

setting are not necessarily representative of performance in another, the value of such controlled evaluations is not clear.

Nonetheless, the primary rebuttal would be that controlled evaluations are necessary to determine the viability of an alternative authentication mechanism before committing to a complex field investigation. Indeed, Tetrad was exposed to a controlled evaluation to determine the viability of the approach before committing to a field investigation. However, the controlled evaluation explored the ability of the user to contend with the novel approach rather than exploring realities such as authentication time.

Nevertheless, arguably if authentication time had been a focus of the initial controlled investigation, the authentication time concern could have been uncovered sooner. Then there would be no need to develop an application or to deploy the mechanism on target devices. However, in many respects this suggests a predefined upper limit on the time an authentication attempt can take. Naturally, many would argue that such limit is the time it takes to enter a PIN on an ATM or a typical password on a personal computer. However, such a preconceived notion of authentication time does not respect the context of authentication.

The pre-defined upper limit on authentication time essentially assumes all tasks involving authentication are similar and all steps within the task are performed sequentially. In fairness, many experimental authentication mechanisms are outlined to solve a similar sequential task: regulating access to a virtual learning environment. Therefore, the process is sequential, e.g. an individual must authenticate before they can see a list of class notes. However, while tasks involving authentication may be performed sequentially from the perspective of the user, the technical architecture does not need to adopt the same structure.

Consider an alternative task, such as purchasing a digital film. The digital film would need to be buffered or downloaded to the individual's device before it could be watched. The digital film will take time to download. There is no reason why the technical implementation does not simply initiate the download immediately. The user is simply prevented from viewing the digital film until they have authenticated. Therefore, the upper limit on authentication time is not that of PIN entry on an ATM but the window of time between the download being initiated and the availability of the film to watch. If the film takes 100 seconds to buffer, for example, as long as the authentication time does not take longer, then the task as a whole has not been delayed by slow authentication.

Therefore, using a controlled evaluation focused on authentication time, with notions of an upper limit, could lead to the premature dismissal of a potentially viable alternative authentication mechanism. Therefore, while authentication time could have been captured in a controlled setting, the *value* and *use* is not necessarily equivalent to that collected from a field experiment with strong ecological validity. The danger is that an authentication mechanism may simply be abandoned purely due to the concern of some preset limit on authentication time. Authentication time was not a focus of the initial controlled evaluation of Tetrad.

Nonetheless, if the initial controlled evaluation had focused on authentication time, arguably it could have been used to improve the overall application. A potential improvement would have been initiating the download of lecture recordings when the application is initiated rather than after the user had successfully

authenticated. The conclusion would be the same as the user would not be able to listen to the lecture recordings until they had authenticated. In this way the overall task time would have been reduced.

The next primary problem to consider was that of *user choice*. There are several research papers that report and discuss user image choice.

8.1.2 User Choice

The controlled evaluations exploring user choice often expect individuals to select images in solitary environment, free from onlookers. Nevertheless, such evaluations have fairly strong ecological validity as most registration processes in the wild could reasonably expect a similar context. However, the registration process would still need to designed *for* and evaluated *on* target devices. The reason is, as was reported earlier: that the design of the registration process could actually influence user choice rather than just the images themselves.

An example would be if the registration process expected users to scroll through several sets of images to select target images. There is the possibility that users simply may not realise they are able to scroll through images and then simply select the first images they encounter. Moreover, some users may select the first images they encounter rather than waste time and energy scrolling through a lengthy list of images. Consequently, the design of the registration process could impact on subsequent user actions. There are several examples of how the design of an interface can impact on user choice. The Microsoft Windows web browser ballot screen is one such prominent example: producers of web browsers argued that position on the ballot screen could impact on user choice [290].

Therefore, the design of the registration process would need to mimic that of the registration process in actual use to ensure the design does not influence the choices made by users. Nevertheless, arguably the registration process could have been designed and evaluated in a controlled setting, avoiding the need for a field investigation. However, in some respects, it is not clear that controlled evaluations would have been possible for versions of Tetrad that relied on personal image collections culled from Facebook. The data downloaded and analysed in the evaluations resided on the individual's personal device and nowhere else. Consequently, if an individual wanted to abandon the application and destroy the downloaded data, all they had to do was the delete the application from their personal device. If a controlled evaluation were used, data would need to be stored and protected on external devices that the user had no control over. Moreover, such external storage would be open to standard processes and policies of back-up and recovery. Consequently, if an individual did decide to leave the evaluation, their data could reside within a network of storage solutions for far longer than desirable. Indeed, there may be no way to concretely confirm their data was destroyed. Furthermore, many may argue as long as the individual is informed of such caveats and consents, then there is no real concern.

However, others may object, such as the organisation, i.e. Facebook. The external company has spent money marketing and developing services and could be unhappy with the prospect of another entity simply amassing data out-with their control. Moreover, Facebook will have terms of service and conditions that

apply to the use of their service. These can change frequently and while it may be acceptable for an external organisation to amass data one day, it may not be acceptable the following day. Lastly, even if an external organisation, such as Facebook, allowed an external organisation to amass data it collected initially, the laws and regulations of the region or country may not allow it. Therefore, data generated and collected in America may not be exported to servers in the United Kingdom. These concerns make a controlled evaluation challenging.

Furthermore, the ecological validity of a controlled evaluation that requires individuals to provide an image collection from their social network is questionable. If an individual attends a controlled session, they may feel coerced to provide the personal image collection and create an authentication secret as there will be an incentive, such as payment or course credit. However, when the incentive is merely access to an application, users may not feel pressured to provide an image collection. Moreover, they may simply not trust the application as much as an evaluator in a white coat. Alternatively, users may feel the process is too intrusive for the purposes of authentication. Therefore, a registration process relying on a personal image collection assessed in a controlled evaluation would have relatively weak ecological validity. Consequently, a field evaluation that presents the registration process to the individual is probably a better approximation to actual user behaviour.

8.1.3 Interaction

The remaining issue was interaction. There was much confusion about how to use the novel authentication approach and arguably much of this stemmed from the design of the registration process. The ad-hoc exploration of different interaction approaches for the authentication mechanism did not reveal any specific problems. However, when deployed in the field, several users struggled with the necessary actions to complete authentication. The problem arguably stemmed from the fact that specific instructions on how to use the mechanism were presented only once and then dismissed; there was no way to recall them. Predictably, users did not read the instructions and were unsure about what action to perform. They merely resorted to the actions they had previously used during the registration process.

The assumption is that a controlled evaluation could have rooted out all the interaction errors and that a field evaluation was not necessary. However, if the authentication mechanism was merely evaluated in isolation, separate from the registration process, it is not clear that users would have made any interaction errors. The errors themselves are arguably a product of users relying on the language of action presented during registration rather than interacting with the authentication mechanism without any baggage. The registration process required users to double-tap an image to select it as a target image, i.e. part of the user's authentication secret. Unsurprisingly, when users were presented a grid of images they probably simply double-tapped the images. However, during authentication, double-tap was used to submit an attempt. Consequently, the authentication attempt would be unsuccessful as users had not yet aligned their images.

The interaction errors also further emphasised that an authentication mechanism is more than simply the process itself and that is why the entire product needs to be evaluated, in context, on target devices with actual users. Nevertheless, arguably a controlled evaluation would have sufficed in extracting the aforementioned interaction errors.

However, arguably the only reason a registration process was developed was because the application was destined for deployment in the field with actual users. The process had to be thoroughly considered as users would be expected to self-support, as is the case with many application and services. A controlled evaluation would rely on a prototype registration process handled by an evaluator. If this was the case, arguably no interaction errors would have been uncovered.

Nevertheless, a controlled evaluation could have sufficed for uncovering interaction errors. However, this is the case only where the product was developed and designed for actual deployment for self-supporting individuals. Crafting an experience for users to operate on their own takes considerably more time and careful consideration than a prototype, partially controlled by an evaluator in a lab. Therefore, once a product is developed for actual deployment, there seems little motivation to perform a controlled evaluation, other than to catch errors with a small group before exposing it to a much larger group in field investigation.

Considering all three field investigations, it is not clear that controlled evaluations would necessarily have unearthed the aforementioned issues with Tetrad. Furthermore, even if controlled evaluations were possible, it is debatable whether they would have been any less complex or costly to manage as compared to field investigations. Nevertheless, the primary concern is determining the root of the aforementioned problems. There could be several explanations for the problems uncovered but a primary explanation may be the reliance on images.

8.1.4 Summary

The reality is that controlled evaluations are typically less costly and complex than equivalent field investigations. Therefore, controlled evaluations may be more desirable in many situations, purely in terms of cost and complexity. Furthermore, a controlled evaluation could be just as effective as a similar field investigation. Indeed, a reasonable argument is that controlled evaluations could have uncovered many of the aforementioned issues that were exposed in the field evaluations of Tetrad.

Nevertheless, field investigations are advantageous in that they force thought on many aspects, that often go unsubstantiated in controlled investigations. A clear example in the field evaluations of Tetrad is the registration process. The design and development of a self-supporting registration process is arguably not necessary for a controlled evaluation. The cost and complexity spent in developing such a self-supporting process can be avoided in a controlled investigation and replaced with a simple manual process involving an experimenter.

Therefore, controlled evaluations are arguably favourable in many situations, purely as they are less expensive and complex than field investigations. Nevertheless, field evaluation of software is reaching a point where it is less expensive and complex than it ever has been. The distribution of software has been streamlined

and many individuals, especially smartphone users, are trained to download and install software. Furthermore, individuals are increasingly purchasing sophisticated devices that are connected to the Internet. Moreover, a healthy economy of systems and services are emerging to support a growing base of sole traders and garage developers in analysing and assessing software. Consequently, developing and deploying a software solution can arguably be assessed with actual users for little more energy than it costs to create a controlled experiment.

Nonetheless, the field evaluations of Tetrad were still problematic and arguably the evaluations would have benefited from listening to users more. The reality is that more effort could have been spent in engaging with users to determine their experience and problems. The use of social channels, such as Twitter and Facebook, could have been harnessed to gain rapid feedback from users. Furthermore, traditional evaluations techniques, such as think alouds and surveys, could have been used to understand more about the experience of users. These techniques could have been used to get a quick and firm grip on what users were experiencing as well as the problems they were encountering.

In short, field evaluations are a strong direction for the evaluation of alternative authentication mechanisms but more effort must be spent to engage with users and to set boundaries on what users are willing to use and experiment with.

8.2 Thesis Statement Review

The thesis statement, initially outlined in the Introduction chapter, is:

> The viability of a recognition-based graphical authentication mechanism can only be evaluated in the wild.
> **Thesis Statement**

There are two keywords in the thesis statement, namely **viability** and **wild**. The viability of the authentication mechanism was assessed using a series of research questions outlined in the Introduction and discussed over the various discussion chapters for each of the field investigations.

The alternative authentication mechanism, Tetrad, is clearly not viable. The primary reasons are:

- The authentication mechanism was inconvenient, authentication times were simply too long. Moreover, authentication times did not improve over time.
- The memorability of the authentication secret did not appear to improve with increased use or inline with the time taken to create the authentication secret.
- There was a difference in the time taken to complete a successful and unsuccessful authentication attempt.
- Lastly, while expecting users to select distractors as well as target images may not have impeded performance, cycling through images did impact on performance.

Unfortunately, it became clear that the root of all these problems essentially stemmed from the observation-resilient design. There are two potential explanations for the lack of convenience: (1) searchmetric design and (2) the observation-resilient entry. The searchmetric design required users to locate images among distractors every time they made an authentication attempt. The position of distractors and targets were randomly positioned for each attempt. Moreover, the observation-resilient entry approach required users to strategise to determine how to align the targets they uncovered. Consequently, the authentication approach always had a sense of uncertainty associated with the user locating the pieces of the authentication secret and aligning them.

Furthermore, the difference in time between successful and unsuccessful authentication attempts is arguably due to users simply not understanding how to interact with, and use, the authentication mechanism. Many of these problems stem from the registration process. An interaction language was introduced and then an entirely different set of actions required to complete the authentication process.

The registration process is clearly a far more important component of a novel alternative authentication mechanism than previously understood, as it essentially introduces the authentication mechanism to the user. Moreover, registrations processes are rarely discussed, never mind designed and implemented, as alternative authentication mechanisms are rarely evaluated in the field.

This brings us to the second keyword in the thesis statement, namely **wild**. The authentication mechanism was evaluated using a controlled investigation as well as several field investigations. The field investigations did have strong ecological validity for the following reasons:

- The authentication task was considered and evaluated for risk and an independent client and user base was sourced.
- The various elements of authentication, such as the application and registration process, were designed and implemented.
- The application that incorporated the authentication mechanism was placed against competition that offered the same resources. Users were not held to ransom with perverse incentives such as course credit.
- The authentication mechanism was deployed on target devices to non-technical users.

The argument could be made that similarly structured controlled evaluations may have produced similar outcomes to the field investigations. Indeed, these arguments were made and discussed in the previous section, §8.1, with the conclusion being that even if controlled evaluations were possible, they would probably not produce similar results as users would feel inclined to focus and perform actions.

Furthermore, the controlled evaluation of the authentication mechanism indicated that users did not encounter any dramatic usability challenges with the observation-resistant entry approach while all the field investigations of the approach indicated that users struggled with the mechanism. Therefore, the outcomes of the evaluations were very different, clearly the field evaluations more

accurately reflect that of an actual application using authentication. Users undeniably encountered problems with the approach.

The difference in performance suggests that the authentication mechanism evaluated in the field investigation was practically fully formed. The approach was designed to be an authentication solution for actual mobile applications. Unfortunately, the various elements required to make the recognition-based graphical authentication approach practical ensured it was simply not viable. The various elements required, such as the registration process and image source, are considerable hurdles for what appears to be very little gain.

Therefore, the thesis statement is supported as while the controlled investigation indicated minimal user problems, several field investigations indicated otherwise. Consequently, while field investigations effectively killed the mechanism, they also indicated how valuable it is to construct the many pieces of authentication and assess with users in the wild before deciding that an alternative authentication approach is viable.

8.3 Sources of Problems

There are potentially many sources for the problems encountered using the authentication mechanism. However, two primary areas that are arguably the root of many of the problems are the images themselves and observation-resilient entry.

8.3.1 Images

A graphical alternative authentication mechanism does not necessarily need to rely on images, at least not images that have to be sourced or previously generated. A graphical alternative authentication mechanism could rely on an individual actually drawing the authentication secret, e.g. DAS. However, Tetrad relied on images of faces and the reality is that there are many problems associated with such images, in regards authentication.

The primary problem with using images was *bootstrapping* the authentication mechanism with them, i.e. essentially sourcing suitable images for the authentication mechanism to use. The complexity of sourcing image sets is exacerbated by the fact that users make predictable choices based on attraction, race and familiarity. Furthermore, sourcing images is difficult due to concerns of image quality and ownership. Therefore, the images in an image set need to be picked and pruned to ensure they are suitable. Moreover, the image set must be scalable and practical in actual use.

However, while these aspects can be undesirable in some respects, such as predicability, they can be desirable in other respects, such as memorability. For example, increased familiarity with an image can actually improve retention. Consequently, a desirable image set should not elicit predictable user choices but also comprise of familiar images. Therefore, the initial controlled evaluation of Tetrad relied on images of celebrities. The assumption was that such images would be familiar to all individuals.

However, the image set was deemed unsuitable for use in the field evaluations. The solution was deemed not scalable due to concerns of attaining permission for use of each of the celebrity images. Furthermore, there is the complexity of celebrities potentially being viewed as endorsing the application as well as approving of the use of the authentication mechanism.

Furthermore, there was concern that celebrity images were not a distinct set that any one application could necessarily own. Consequently, similar image sets could be used across several applications, affording individuals the ability to create the same authentication secret across applications. Password reuse, a primary password problem, would persist with the alternative authentication mechanism.

Therefore, an image set that an organisation could retain control over, but was familiar to users, was deemed the most desirable image set. Therefore, the application used an image set of staff profile pictures from the School of Psychology. The assumption was the images were essentially of 'local celebrities' to the user base. Furthermore, permission was obtainable for the images and the solution was scalable as other organisations could simply offer similar profile pictures. Consequently, there would be no opportunity for individuals to reuse authentication secrets as they would not be available outside the application. Nevertheless, there was concern that users would make predictable choices based on attraction, race and familiarity. However, the assumption was that the image set was not particularly diverse in terms of appearance and all individuals would have likely be familiar to all individuals using the application.

Unfortunately, the reality is that users did make similar choices and some images were considerably more popular than others. The outcome is not particularly surprising as individuals favoured images of staff that were prominent in the first year teaching programme. Users simply selected images they had confidence in, that they could easily remember, to ensure they were spared the inconvenience of not being able to authenticate. Therefore, while the image set was of familiar images, the set clearly contained images that were far more familiar or famous than others. Therefore, an alternative image set was needed for the authentication mechanism to be viable.

Consequently, the second iteration of authentication relied on users essentially providing the image set for Tetrad. It required users to connect their Facebook account to the application. The image collection was all the public profile pictures of a user's friends on the popular social network service. The images were downloaded and analysed to extract face images for users to select for use in authentication. Users were expected to select the entire image set, distractors and target images. However, there was concern that authentication times may ballon as users may struggle to quickly discern target images from distractor images.

Fortunately, authentication times were not impacted by the inclusion of user selected distractor images, although the sample was small and further investigation was deemed necessary. Therefore, there is promise in an individual providing their own image collection for use in authentication. Nevertheless, the reason the sample was small was partly due to the complex and lengthy registration process required to create the authentication secret. The entire process of downloading, analysing and extracting images was lengthy and all this was before an individ-

ual selected images for use in authentication. Asides from these problems there was also the concern that close associates may be able to discern the friends an individual had selected for their authentication secret.

Therefore, while expecting users to select distractor and target images may be viable, the registration process was less than optimal. Consequently, the next iteration of the authentication approach tried to reduce the time taken to complete the registration process and the energy spent sourcing images from Facebook. The decision was taken to use tags of individuals in images from Facebook. This negated the need for the analysis step on the device as Facebook could provide the location of tags within images. The location of each tag was used to extract images for each friend. The user would then create a personal collection of friends for use in the authentication approach both distractors and targets. Moreover, since the analysis step had been removed, the decision was taken to create three images for each individual, if possible. The authentication mechanism could then cycle through different images for each friend in the image set.

There were two reasons for cycling through images. The first reason was to tackle the concern that close associates might able to discern the friends selected as part of the authentication secret. Cycling through images was assumed to make it harder for close associates to make the connection. The second reason was that observing faces from different poses and positions can improve the retention of the authentication secret.

However, while removing the analysis step may have reduced the burden of the registration process, it still took considerable time to complete the registration process. The assumption was that although a longer registration process might take time, it would lead to memorable and observation-resilient authentication. Unfortunately, users also struggled with the third iteration. Moreover, several attempts were incorrect due to sequence, i.e. users were able to group some, if not all, the images together but not in the correct sequence.

Nevertheless, the third iteration was still less than optimal and it seems simply too intensive for accessing lecture recordings. Furthermore, while it may appear that many of the problems of the authentication mechanism may be due to images, arguably many of the problems stemmed from the need to protect entry of the authentication secret.

8.3.2 Observation-resilience

The need to offer an observation resilient authentication mechanism was necessary not because the approach was *graphical* or image-based but because it was *recognition*-based.

The elements of the authentication secret were literally presented on-screen. Therefore, distractor elements were needed to confuse attackers but ideally this should have little impact on users. The distractors come at a cost though since a user has to filter and ignore distractors and locate targets, a process that comes at the cost of time and energy. The process is necessary as the authentication mechanism is recognition-based. Therefore, arguably there may be little benefit in offering a recognition-based approach, if it comes at the cost of having to filter

through distractors to locate targets. The process is costly in terms of time and energy.

Furthermore, many of the problems associated with the approach stem from the use of the elaborate and complex process of entry. Consequently, another argument could be that the process was simply incorrect, that an alternative observation-resilient process could have been used to enter the images. There is certainly an argument that an alternative process could have been used but it would have ultimately relied on a collection of distractor and target elements. There is the possibility an alternative process could not use distractors and rely solely on targets but that would still require some sort of procedure to deter attackers at some cost, including time, training and education. Moreover, while a superior process is entirely possible and not suffer from the same problems, it could potentially suffer from a range of other problems.

Thus, software-based observation-resilience still comes at some cost to the user in terms of time and energy. Consequently, it may be worth considering the need for observation-resilience techniques in the first instance.

Argument against observation-resilient entry

The biggest mistake security researchers can make is assuming that user effort is free [116]. The benefit of shielding an authentication secret, for example, must not come at an exorbitant cost. Therefore, the real question becomes: How big a threat is shoulder-surfing? The threat needs to be fairly big to be worth the hassle of dealing with a mechanism such as Tetrad, instead of alphanumerics. The prototypes forced thought on deployment issues, such as how the mechanism would be bootstrapped with personal images and what sort of application it would protect.

The original scenario we envisioned was a user purchasing a movie in front of friends and family. The reality, however, is that most users do not care if these people know the password they use to purchase movies. They want to share a movie in their living room, the password is secondary and less important. If a friend or family member does use someone else's password, then they will work it out. The same way they would, for example, if someone took the last beer or slice of pizza. Shoulder-surfing in this context is a non-issue, something we failed to recognise.

Therefore, for the prototypes we envisioned the scenario of an individual accessing a digital store, being exposed to the wider world. For example, purchasing a movie on a smartphone while having a coffee at Starbucks or waiting for train in a station. These places are bursting with strangers and a user would not want any of them to observe entry of their password. However, users are likely to authenticate on a 4-inch screen, not a 40-inch one. Therefore, they can easily position themselves and the screen to prevent observation or use a privacy screen.

Lets say a stranger does, somehow, manage to observe the password being entered in such a space. What are the actual consequences, in the case of a digital store? If an attacker does manage to purchase several items from a service, such as iTunes, Apple can de-activate the purchases.

Therefore, the focus on digital stores might have been the wrong scenario.

Tetrad could have been used to protect email or a financial service. However, no individual would check their personal email account or personal finance account on a television in front of friends and family. Not necessarily because the information is sensitive, but because it is boring and uninteresting. Large screens are for sharing and although some users might want to read email and check balances on a big-screen, they would wait until the television were free, rather than interrupting a movie.

Constructing prototypes that are meant to be the manifestation of a practical, deployable product require us to focus on the actual application used, the people using it and the places where it will be used. Authentication is a package, it is not something that is spooned on top of a finished product: it is part of the product. The time and energy taken to create a Tetrad secret is not necessarily worth the hassle because shoulder-surfing is not really something to be concerned about in many contexts.

In the end we did not actually tackle an authentication problem: we tackled a password problem. Tetrad was not designed thinking about the users and the tasks they wanted to complete but in order to right the wrongs of passwords. Tetrad was burdened with the legacy of the password approach. The design process did not start with: people want to buy movies in the living room. Instead it started with: these are the problems with passwords in the living room.

8.4 Access and Accountability

The aforementioned ideas were presented and discussed at BCS HCI 2012 [174]. Many ideas were put forward but the consensus suggested that authentication could essentially be improved by removing users from the authentication loop, probably using biometric-based authentication. Therefore, a graphical authentication approach may have been entirely the wrong research direction. A different direction may be to remove users from the authentication loop or rather remove their awareness of authentication. The strategy has been used by various technology companies, such as Apple and Google, as to improve the usability of software, one such example is management of files.

In the personal computer era, users were expected to manage the files generated by applications. However, a different direction has been taken on mobile devices and on the web, where users do not manage files, instead applications manage files. The approach was common on personal computers but not widespread, e.g. Apple iTunes was an early desktop application that managed music files, rather than the user. The need to manage files, construct hierarchies and monitor storage space still continues but the user's involvement is essentially curtailed to the creation or purchase of files. The deletion of files is a command users can still issue but seems increasingly associated with hiding a file from sight rather than its actual destruction.

Similarly, biometric sensors could afford subtle, almost invisible, authentication. An example would be the unlock button on a mobile phone. The unlock button could double as a biometric sensor. Consequently, the authentication process would begin when the users comes into contact with the sensor. If the user

is the owner of the device it will unlock, otherwise the individual will be denied access. Therefore, ideally the user would never encounter the authentication approach.

The main barrier to the deployment of such biometrics approaches has been the specialised hardware required to use them: users simply do not want to be weighed down with equipment specifically for authentication. However, several companies are in a position to deliver such solutions as part of their hardware offerings. Apple, Google, Microsoft and Amazon could all deliver subtle authentication to millions of individuals at a very rapid pace. Moreover, there is also a concern surrounding speed and accuracy, the situation could be improved by using the output of several sensors embedded in watches, eye-glasses and clothes. Again, several large technology companies would be best positioned to deliver such ecosystems to individuals.

However, while removing users from the authentication loop may be acceptable for the purposes of regulating access, it is arguably unthinkable for the purposes of accountability. How can an individual be accountable for an action, when they did not deliberately authorise it?

8.4.1 Consequences of Authentication

An authentication mechanism often embodies a decision. If an individual purchases a digital product, such as a film or song, they are instantly prompted for their password. The user is confirming their decision to purchase a product. They are accountable for their actions. An authentication mechanism is not necessarily required for such actions, a dialogue box could just as easily substitute for the authentication mechanism but users may not understand the weight of the consequences from a mere dialogue box.

Users may simply not read text or absorb actions, they may not realise their are consequences. An authentication mechanism can communicate *consequences*. The mere presentation of an authentication mechanism can communicate to an individual that they are making a decision, that actions will follow that they will be accountable for. An example would be the presentation of the password prompt on the Apple iPhone. Apple have made a concentrated effort to only display the password prompt when an individual is being charged hard currency. The password prompt rarely appears otherwise, expect when you install software etc.

The benefit to the approach is that the password prompt could appear at any time and it indicates to the individual that entry of the password will result in a charge. An individual could be watching a film, for example, and if a password prompt appears, the user understands that they will be charged once they enter their password. However, arguably, once again, an authentication mechanism is not necessary for such actions, a dialogue box could suffice.

However, many users may disagree and appreciate the weight and gravitas of an authentication mechanism. Apple recently settled a lawsuit with users who complained that it was too easy and not entirely clear that they were making chargeable purchases [11]. The complaint was caused, in part, by a design decision.

The password prompt is presented when an individual is required to make a payment. However, when users enter their password a 15 minute authorisation window is also activated. Consequently, chargeable purchases made within the window are automatically authorised without subsequent password entry. The approach merely presents a dialogue box, asking users to confirm payment.

8.4.2 Client vs Users

A business may want to use biometrics, for example, to provide seamless and secure access to enterprise files but users may feel it is too intrusive. In the consumer market users are free to abandon applications and services where they feel the authentication process is simply too intrusive or intensive. Indeed, all the field investigations of Tetrad illustrate that users are quite happy to abandon applications that are too demanding.

An important aspect of all the field investigations was that users were not forced to use the application. Users could attend lectures without entering any authentication process and could access a web portal offering the same lecture recordings using a traditional light-weight password. Consequently, users had options, our application and authentication mechanism, essentially had real competition in the field. Users were not carolled into using our authentication mechanism, arguably if we offered course credit or were the only source of lecture recordings, users may not have been so quick to dismiss the application.

However, the reality is that developers creating applications do face competition from many sources and if an authentication mechanism dissuades users, then developers will drop it. There is no art or science in protecting information. The problem is not protecting information. If that was the case, destruction of information is the best direction, as information that is destroyed is not accessible to anyone. An authentication process is necessary to protect information from attackers but make it accessible to users.

Biometric-based authentication is probably the best solution for the aforementioned problem. The process has the potential to be invisible to users but inconvenient for attackers. Nevertheless, the legal case with Apple indicates that at least some users feel it necessary to be deliberately involved in the authentication loop. The Apple example speaks to the other use of authentication, namely accountability and arguably users need to be involved in the authentication process to be accountable.

Furthermore, there is arguably no authentication approach better for accountability than a knowledge-based approach. The authentication secret literally resides within the individual and they have to participate in the authentication process to complete it. Moreover, users can pass knowledge to others to act on their behalf, the user can remain accountable but assign another individual to act of their behalf. An example would be a pensioner entrusting another individual to collect cash from an ATM on their behalf, they simply tell the other individual their PIN.

Unfortunately, passwords and PINs are the only viable knowledge-based authentication approaches. There are simply too many unknowns with the numerous proposed knowledge-based alternatives. The lack of progress is due to the

lack of investigation of the various components of authentication, such as the registration process and application as well as field investigations with strong ecological validity.

The field evaluations of Tetrad were deployed on actual devices, such as iPhones and iPads, to non-technical students. There was no preserve incentive such as course credits, online assessments or exclusive access to teaching resources. The application was deployed against competing products and options, no user had to use the application or the authentication mechanism. Moreover, the application and authentication was delivered to a client. The investigators and evaluators were separate from the client requesting the application. Consequently, the application and authentication mechanism had strong ecological validity and arguably closely mirrored an actual scenario.

However, arguably the field evaluations of Tetrad were timid and did not go far enough. The application had potential but it was simply not capitalised on. The reality is that tried and true metrics, e.g. time taken, that had been used in controlled evaluations of other alternative authentication mechanisms were essentially relied upon in the field investigations of Tetrad. However, there was so much more potential in assessing different aspects of how users actually use authentication in the wild. The location of where they authenticated could have been recorded, devices nearby, whether users are in motion or not and the time spent on the entire task etc. Unfortunately, while the field investigations were novel the metrics were not and clearly more effort is required to outline frameworks and metrics for adequately assessing alternative authentication mechanisms in the field.

There must be more investigation into the actual performance of authentication mechanisms in their entirety, rather than reporting simply on the process itself. Many alternative authentication mechanisms are evaluated and reported much like stars in films. Scripts can be written to make the solution sound wonderful but when considered in actual use, it is not clear the characters could function.

Tetrad was terrible and creepy, undesirable adjectives that fell from the mouths of users. The focus has arguably been assessing authentication in a vacuum, addressing a problem with an ageing mechanism that is increasingly becoming irrelevant not because of research but because users are shifting away from desktop computers.

The only direction for alternative authentication mechanisms is to refocus on the user, refocus on the problems the user is encountering and give due consideration to context. The time is quickly approaching where users will be unable to perform certain tasks not because the password approach has so many problems but because there is simply not a viable alternative.

Chapter 9

Conclusion

The Background chapter opened with the notion that an alternative knowledge-based authentication solution had the potential to solve the problems of passwords. Consequently, Tetrad, an authentication solution that relied on face images, was proposed as such an alternative. The initial controlled evaluation suggested that we might have taken the correct path as Tetrad was deemed a success. Users did not struggle with the shared-space prototype. Moreover, the approach stimulated discussions and conversations about the need for such an observation-resilient method when presented at conference. However, several subsequent field investigations revealed Tetrad was not only suboptimal, it was awful. Therefore, Tetrad is not a solution to the problems of passwords.

The reality is that the password will probably be around, in the same format, a decade from now [116]. Herley argues that researchers have failed to realise that the password is actually the best authentication solution in many contexts [116]. The password is powerful, accessible and versatile [194]. Passwords, in their first incarnation, performed exactly as planned. Many of the problems we have today are a consequence of indiscriminate and excessive usage thereof, more so than the nature of the mechanism itself. Consequently, rather than attempting to replace the password, a near-term solution may be to accept it but to find a better way of mediating its use, of tackling the gratuitousness of password usage.

Therefore, the future direction of authentication research may not be the replacement of the password approach but rather by minimising the footprint of authentication on task and devices [222]. This topic is discussed further in the Future Work section, see §9.2.

Nevertheless, there are many problems with passwords and the list will only grow longer as computers take increasingly different forms. Passwords simply do not make sense on eyeglasses or watches as there is no physical keyboard and limited screen space. Furthermore, while biometrics may be the solution on such devices, the solution must not be invisible as accountable actions need to keep the user in the authentication loop. The strength of knowledge-based authentication approaches is the suitability to accountable authentication, primarily as the authentication secret has to be extracted from the individual.

Secret sharing is often considered a problem with passwords: that an individual can share knowledge used to authenticate. However, in certain instances, the ability to share a password can be incredibly important and is a strength of

an authentication approach. Nevertheless, many organisations would view this as a concern as they need to ensure accountability. This typifies the divide between what users want from authentication and what the developer deploying the mechanism is trying to achieve.

Therefore, arguably there is still a need for a knowledge-based authentication approach replacement for passwords. Tetrad is just not that replacement.

9.1 Contributions

Nevertheless, while Tetrad was not a viable alternative authentication mechanism, its research and development provided many lessons and contributions:

1. *An authentication mechanism can not be studied in isolation*
   The research demonstrated that even carefully considered components can be flawed and impact on the process of authentication itself. An authentication mechanism may simply be a series of steps but the ingredients and products of each step can have an impact on the overall performance of the process. Consequently, the viability of an authentication mechanism can not be determined in a vacuum with other aspects removed. Every aspect and element must be designed and developed to assess performance.

2. *A controlled evaluation can erroneously suggest that an impracticable alternative authentication mechanism is viable.*
   Controlled evaluations are incredibly important at determining specific aspects of an authentication mechanism. However, in determining the viability of an authentication mechanism for actual use, controlled evaluations must be carefully crafted to mimic actual conditions. Nevertheless, controlled evaluations can be structured and performed to suggest an authentication mechanism is fit for purpose when in actuality many caveats ensure an authentication approach is not practical. Indeed the shared-space prototype assessed using a controlled evaluation was published and presented, concluding that the observation-resilience authentication approach was promising. However, after field investigations, we found that the mechanism was awkward and suboptimal, arguing that observation-resilience was the root of problems.

The research and development of Tetrad has been a journey in itself. The starting point was a successful shared-space prototype that showed promise. The observation-resilient authentication approach was well received when presented and discussed at BCS HCI 2009. However, several years later after intensive development and field investigations we presented at the BCS HCI 2012 conference only to announce that observation-resilience may not have been as important as initially considered. These ideas were outlined and discussed in the Discussion chapter.

The reality is that by focusing on observation-resilience we delivered a suboptimal authentication approach that was provably not viable. However, it is still not clear how a viable a recognition-based authentication approach can be without protecting the authentication secret from the wandering gaze of others.

Our journey to tackle the problems of passwords continues, albeit in a different direction, as discussed in the next section.

9.2 Future Work

Researchers have spent a great deal of effort coming up with new authentication mechanisms e.g. alternative alphanumeric approaches and graphical alternatives. The motivation appears to be to find a more memorable alternative to supplant the password. Having found one, researchers publish details, with evidence of the alternative mechanism's superiority. Yet the password persists.

The widespread use of passwords has accustomed users to access control as a necessary evil. Unfortunately, the very familiarity of the mechanism has also been its undoing. A clear indication of the failure of the password is the prevalence of policies and procedures that use words such as comply, sanction and disciplinary. Unfortunately these policies, instead of convincing users to walk the developer's intended path, often do more harm than good [116].

As authentication researchers, we have often been obsessed by the mechanism itself, rather than its placement within a workflow, task or path. Developers define such workflows and lay down paths for navigating them, users merely react. Hence developers determine the position, use and frequency of authentication challenges. We need first to understand their perspective, the paths they create, and how these can better be designed to overlap with the paths users prefer.

Paths in the grass

Siracusa talks about users making paths in the grass: paths which allow humans to achieve their goals while expending as little extraneous effort as possible [252]. He says:

> "Any viable solution must work within the (often inconvenient) bounds of reality. It must be constructed in such a way that the motivations and actions of the participants — both the good and the bad...especially the bad — serve tobalancethe system as a whole. Suggesting that all would be well, if only certain people would act differently or alter their desires in some way is wishful thinking, not an actual solution."
> **John Siracusa [252]**

This strikes a chord. Wishing that people "would only behave securely" has almost been the mantra of security practitioners world-wide for at least the last decade. Siracusa argues that socialism, communism and libertarianism have failed because they attempt to create artificial systems which ignore the nature of participants, or which insist that they change their nature. Siracusa relates that when the University of California at Irvine campus was first built, they did not lay sidewalks: they planted grass. The next year, they returned and laid the sidewalks where the trails were in the grass. We must find ways to consider how users want to authenticate, and design to accommodate their "paths". Trying to coerce them into walking down our paths is futile.

Consider Wikipedia, a collaborative encyclopaedia launched shortly after the millennium. The articles on the website are generated by the man and woman in the street, not by experts. Wikipedia is the 6th most visited website on the planet. The authors are not owners of articles per se: instead articles are generated through user collaboration, by crowd-sourcing. Users are not vetted; any individual is able to generate or edit an article. In summary, Wikipedias developers had people (everyone), problem (building a knowledge resource), and context (accountability is less important than encouraging contributions). They chose a novel approach: no authentication.

Malicious users are always a concern and the obvious solution would have been to require authentication. Yet that would probably have discouraged contributors, and Wikipedia seemingly wanted to be sure that no obstacles impeded the creation of a knowledge resource, so they created a new path of unhindered access. It turns out that the developers path coincided with the path users preferred, as demonstrated by the efforts of millions of contributors. Investigations have found the pages to be of high quality [99] similar to traditional encyclopaedias. Vandalism attacks do occur [282] but users are unaware of this because tools, such as watch lists [186], make contributors aware of changes so that they can be corrected. Wikipedia has essentially outsourced both creation and policing. Their approach perfectly matches the risk problems of their context, and it works. Wikipedias approach is unusual.

Nevertheless, a future direction of our authentication research could be the development of usage patterns for developers that would serve to reduce the deployment of authentication to sufficiency rather than extravagance. It is worth considering how many times an individual is required actively to unlock their device during any given day. They may infrequently perform actions with serious or risky side-effects. Nevertheless, whether or not they want to use their device to check the time or to stop music playing, they are presented with an authentication hurdle. This traditional approach can be referred to as "exhaustive access control".

Therefore, a future direction could be to consider the numerous tasks performed on devices and determine when authentication should be presented and when it should be avoided. These ideas were presented and discussed at HAISA 2013 [222].

9.3 Concluding Remarks

The research and development of Tetrad was a long and arduous journey. The alternative authentication mechanism was designed for a context rather than attempting to be another pseudo silver bullet password replacement applicable in all scenarios. This research has concluded that there is no longer a battle for an authentication solution on desktop and laptop computers, passwords won that war.

However, there is a new frontier: the next generation of computers that no longer look like computers, namely: watches, eyeglasses, cars and clothing. The iPhone is not so much a phone as a highly portable computer as is the iPad. These

devices will be *technically* the same. However, the introduction of such devices into a wide range of contexts changes everything, as everyday tasks increasingly incorporate Internet access and computers.

Passwords simply do not make sense on such devices. It is not clear that passwords can make the leap to such devices and fulfil authentication needs for so many different contexts. Nevertheless, what is the alternative? There are few choices but a clear authentication alternative for such devices has yet to appear. In many ways this is due to lack of focus, in authentication research, on context.

There are many lessons to be drawn from the research and development of Tetrad. Many of these lessons apply to all authentication mechanisms, not simply recognition-based graphical authentication mechanisms. The context of an alternative authentication mechanism must be designed, implemented and evaluated for a given context. This dissertation outlines the extensive work done to create evaluations that target specific contexts and had strong ecological validity. Nevertheless, the field investigations were complex, hard work and did not push far enough.

Tetrad's journey should act as an inspiration and a warning. Alternative authentication mechanisms must be evaluated in the field and while researchers may not enjoy the response to their hard work, such criticism must be taken on board and used to develop solutions for the next generation of computers.

If alternative, viable authentication solutions can not be created for the next generation of computers, then everyday users will not be able to complete their own journeys.

Passwords have brought humanity this far and empower users everyday. The concern is that the password is struggling and nothing new is emerging to take over the reins.

Bibliography

[1] Anne Adams and Martina Angela Sasse. Users are not the enemy. *Commun. ACM*, 42(12):40–46, December 1999. ISSN 0001-0782. doi: 10.1145/322796.322806. URL http://doi.acm.org/10.1145/322796.322806.

[2] F. Aloul, S. Zahidi, and W. El-Hajj. Two factor authentication using mobile phones. In *Computer Systems and Applications, 2009. AICCSA 2009. IEEE/ACS International Conference on*, pages 641–644. IEEE, 2009.

[3] Mark I Alpert, Judy I Alpert, and Elliot N Maltz. Purchase occasion influence on the role of music in advertising. *Journal of Business Research*, 58(3):369–376, 2005.

[4] J.R. Anderson and G.H. Bower. *Human associative memory: A brief edition*. Lawrence Erlbaum, 1980.

[5] R. Anderson, M. Bond, and S.J. Murdoch. Chip and spin. *Computer Security Journal*, 22(2):1–6, 2006.

[6] Apple Press Release. Apple Launches the iTunes Music Store, April 2003. URL http://www.apple.com/pr/library/2003/04/28Apple-Launches-the-iTunes-Music-Store.html.

[7] Apple Press Release. Apple's App Store Downloads Top 25 Billion, March 2012. URL http://www.apple.com/pr/library/2012/03/05Apples-App-Store-Downloads-Top-25-Billion.html.

[8] Marco Arment. Instapaper, July 2008. URL http://www.instapaper.com/iphone.

[9] Marco Arment. Requiring email and passwords for new accounts, December 2010. URL http://blog.instapaper.com/post/2318776738.

[10] Charles Arthur. iOS v Android: app revenues, downloads and country breakdowns, December 2012. URL http://www.guardian.co.uk/technology/appsblog/2012/dec/04/ios-android-revenues-downloads-country.

[11] Charles Arthur. Apple faces multimillion US settlement over 'in-app purchases' by children, February 2013. URL http://www.theguardian.com/technology/2013/feb/26/apple-settlement-children-in-app-purchases.

[12] ASDA. Low prices and a friendly welcome: what you'll find in every Asda store, 2013. URL http://your.asda.com/our-stores.

[13] H.P. Bahrick, P.O. Bahrick, and R.P. Wittlinger. Fifty Years of Memory for Names and Faces: A Cross-sectional Approach. *Journal of Experimental Psychology: General; Journal of Experimental Psychology: General*, 104 (1):54, 1975.

[14] Denis Baldisserra, Annalisa Franco, Dario Maio, and Davide Maltoni. Fake fingerprint detection by odor analysis. In *Advances in Biometrics*, pages 265–272. Springer, 2005.

[15] B.F. Barton and M.S. Barton. User-friendly password methods for computer-mediated information systems. *Computers & Security*, 3(3):186–195, 1984.

[16] BBC. Apple iPhone 5 unveiled with taller screen and 4G, September 2012. URL http://www.bbc.co.uk/news/technology-19572820.

[17] BBC News. Napster use slumps 65%, February 2001. URL http://news.bbc.co.uk/1/hi/business/1449127.stm.

[18] BBC News. Cards 'more popular than cash', July 2004. URL http://news.bbc.co.uk/1/hi/business/3872817.stm.

[19] BBC News. Chip and pin 'cutting' card fraud, October 2005. URL http://news.bbc.co.uk/1/hi/business/4320072.stm.

[20] BBC News. UK consumers see card fraud rise, October 2008. URL http://news.bbc.co.uk/1/hi/business/7646670.stm.

[21] BBC News. Inventor of cash machine, John Shepherd-Barron, dies, May 2010. URL http://news.bbc.co.uk/1/hi/scotland/highlands_and_islands/8691747.stm.

[22] BBC News. Mobile phone £6,875 bill for decorator Chris Wilson, December 2012. URL http://www.bbc.co.uk/news/uk-wales-south-west-wales-20567165.

[23] A. Beautement and M.A. Sasse. Gathering realistic authentication performance data through field trials. In *SOUPS USER Workshop*, 2010.

[24] A. Beautement, M.A. Sasse, and M. Wonham. The Compliance Budget: Managing Security Behaviour in Organisations. In *Proceedings of the 2008 workshop on New security paradigms*, pages 47–58. ACM, 2009.

[25] P.J. Benson and D.I. Perrett. Perception and recognition of photographic quality facial caricatures: Implications for the recognition of natural images. *European Journal of Cognitive Psychology*, 3(1):105–135, 1991.

[26] T. Berson, P. Capek, J. Schweitzer, and C. Weissman. Identity verification(authentication) working group. *ACM SIGSAC Review*, 6(1):2–9, 1988.

[27] A. Bhargav-Spantzel, A.C. Squicciarini, S. Modi, M. Young, E. Bertino, and S.J. Elliott. Privacy preserving multi-factor authentication with biometrics. *Journal of Computer Security*, 15(5):529–560, 2007.

[28] A. Bianchi, I. Oakley, V. Kostakos, and D.S. Kwon. The phone lock: audio and haptic shoulder-surfing resistant pin entry methods for mobile devices. In *Proceedings of the fifth international conference on Tangible, embedded, and embodied interaction*, pages 197–200. ACM, 2011.

[29] Andrea Bianchi, Ian Oakley, and Dong Soo Kwon. The secure haptic keypad: a tactile password system. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1089–1092. ACM, 2010.

[30] Irving Biederman. On the Semantics of a Glance at a Scene. In M. Kubovy and J. R. Pomerantz, editors, *Perceptual Organization*, pages 213–263, 1981.

[31] Irving Biederman and Margaret M Shiffrar. Sexing day-old chicks: A case study and expert systems analysis of a difficult perceptual-learning task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13 (4):640, 1987.

[32] Irving Biederman, Thomas W Blickle, Richard C Teitelbaum, and Gary J Klatsky. Object search in nonscene displays. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14(3):456, 1988.

[33] Emmanuel Bigand and Bénédicte Poulin-Charronnat. Are we "experienced listeners"? a review of the musical capacities that do not depend on formal musical training. *Cognition*, 100(1):100–130, 2006.

[34] M. Bishop. Proactive password checking. In *4th Workshop on Computer Security Incident Handling*. Citeseer, 1992.

[35] G.E. Blonder. Graphical password, 1996. US Patent 5,559,961.

[36] J. Bodamer. Die prosop-agnosie. *European Archives of Psychiatry and Clinical Neuroscience*, 179(1):6–53, 1947.

[37] H. Bojinov and D. Boneh. Mobile token-based authentication on a budget. In *Proceedings of the 12th Workshop on Mobile Computing Systems and Applications*, pages 14–19. ACM, 2011.

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

[39] G.H. Bower and M.B. Karlin. Depth of Processing Pictures of Faces and Recognition Memory. *Journal of Experimental Psychology*, 103(4):751, October 1974.

[40] G.H. Bower, M.B. Karlin, and A. Dueck. Comprehension and memory for pictures. *Memory & Cognition*, 3(2):216–220, 1975.

[41] J. Brainard, A. Juels, R.L. Rivest, M. Szydlo, and M. Yung. Fourth-factor authentication: somebody you know. In *Conference on Computer and Communications Security: Proceedings of the 13 th ACM conference on Computer and communications security*, volume 30, pages 168–178, 2006.

[42] E. Brunswik. Organismic achievement and environmental probability. *Psychological Review*, 50(3):255, 1943.

[43] R. Bruyer. Covert face recognition in prosopagnosia: A review. *Brain and Cognition*, 15(2):223–235, 1991.

[44] J. Bunnell, J. Podd, R. Henderson, R. Napier, and J. Kennedy-Moffat. Cognitive, associative and conventional passwords: Recall and guessing rates. *Computers & Security*, 16(7):629–641, 1997.

[45] G. Byatt and G. Rhodes. Recognition of own-race and other-race caricatures: implications for models of face recognition. *Vision research*, 1998.

[46] A.W. Carroo. Recognition of faces as a function of race, attitudes, and reported cross-racial friendships. *Perceptual and Motor Skills*, 64(1):319–325, 1987.

[47] D.S. Carstens, P.R. McCauley-Bell, L.C. Malone, and R.F. DeMara. Evaluation of the human impact of password authentication practices on information security. *Informing Science: International Journal of an Emerging Transdiscipline*, 7:67–85, 2004.

[48] S. Chakrabarti, G.V. Landon, and M. Singhal. Graphical passwords: drawing a secret with rotation as a new degree of freedom. In *The Fourth IASTED Asian Conference on Communication Systems and Networks (AsiaCSN 2007)*, pages 561–173. Citeseer, 2007.

[49] K. Chalkias, A. Alexiadis, and G. Stephanides. A multi-grid graphical password scheme. In *Proceedings of the 6th International Conference on Artificial Intelligence and Digital Communications*. Citeseer, 2006.

[50] S. Chiasson, PC Van Oorschot, and R. Biddle. A usability study and critique of two password managers. In *15th USENIX Security Symposium*, pages 1–16, 2006.

[51] S. Chiasson, R. Biddle, and P.C. van Oorschot. A Second Look at the Usability of Click-based Graphical Passwords. In *Proceedings of the 3rd Symposium on Usable Privacy and Security*, pages 1–12. ACM, 2007.

[52] S. Chiasson, P. van Oorschot, and R. Biddle. Graphical Password Authentication Using Cued Click Points. *European Symposium on Research in Computer Security*, 4734:359–374, September 2007.

[53] S. Chiasson, A. Forget, R. Biddle, and P.C. van Oorschot. Influencing Users Towards Better Passwords: Persuasive Cued Click-Points. In *Proceedings of the 22nd British HCI Group Annual Conference on People and Computers: Culture, Creativity, Interaction*, volume 1, pages 121–130. British Computer Society, 01-05 September 2008.

[54] M. Cole, L. Hood, and R.P. McDermott. Concepts of ecological validity: Their differing implications for comparative cognitive research. *Mind, culture, and activity: Seminal papers from the Laboratory of Comparative Human Cognition*, pages 49–56, 1997.

[55] Corbató, F.J. On building systems that will fail. In *ACM Turing Award Lectures*. ACM, 1990.

[56] Corbató, F.J. and Merwin-Daggett, M. and Daley, R.C. An Experimental Time-sharing System. In *Proceedings of the Spring Joint Computer Conference*, pages 335–344. ACM, May 01-03 1962.

[57] Nelson Cowan. The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and brain sciences*, 24(1): 87–114, 2001.

[58] M. Crabb. Password security in a large distributed environment. In *Proceedings of the UNIX Security Workshop II*, pages 17–30, 1990.

[59] F.I.M. Craik and R.S. Lockhart. Levels of Processing: A Framework for Memory Research. *Journal of Verbal Learning and Verbal Behavior*, 11(6): 671–684, 1972.

[60] F.I.M. Craik and E. Tulving. Depth of Processing and the Retention of Words in Episodic Memory. *Journal of Experimental Psychology: General*, 104(3):268–294, September 1975.

[61] J.F. Cross, J. Cross, and J. Daly. Sex, race, age, and beauty as factors in recognition of faces. *Attention, Perception, & Psychophysics*, 10(6):393–396, 1971.

[62] D. Davis, F. Monrose, and M.K. Reiter. On User Choice in Graphical Password Schemes. In *Proceedings of the 13th USENIX Security Symposium*, pages 151–164, 2004.

[63] A. De Angeli, L. Coventry, G. Johnson, and K. Renaud. Is a picture really worth a thousand words? exploring the feasibility of graphical authentication systems. *International Journal of Human-Computer Studies*, 63(1): 128–152, 2005.

[64] A. De Luca, E. Von Zezschwitz, and H. Hußmann. Vibrapass: secure authentication based on shared lies. In *Proceedings of the 27th international conference on Human factors in computing systems*, pages 913–916. ACM, 2009.

[65] A. De Luca, K. Hertzschuch, and H. Hussmann. Colorpin: securing pin entry through indirect input. In *Proceedings of the 28th international conference on Human factors in computing systems*, pages 1103–1106. ACM, 2010.

[66] M. Dell'Amico, P. Michiardi, and Y. Roudier. Password strength: an empirical analysis. In *INFOCOM, 2010 Proceedings IEEE*, pages 1–9. IEEE, 2010.

[67] Travis Deyle and Volker Roth. Accessible authentication via tactile pin entry. *Computer Graphics Topics*, 2:24–26, 2006.

[68] R. Dhamija and A. Perrig. Déjà Vu: A User Study Using Images for Authentication. In *Proceedings of the 9th Conference on USENIX Security Symposium*, 2000.

[69] F.J. Di Vesta and G.S. Gray. Listening and note taking. *Journal of Educational Psychology*, 63(1):8–14, 1972.

[70] R. Diamond and S. Carey. Why Faces Are and Are Not Special: An Effect of Expertise. *Journal of Experimental Psychology: General*, 115(2):107, 1986. URL `http://psycnet.apa.org/journals/xge/115/2/107/`.

[71] A. Diddi and R. LaRose. Getting hooked on news: Uses and gratifications and the formation of news habits among college students in an internet environment. *Journal of Broadcasting & Electronic Media*, 50(2):193–210, 2006.

[72] A.E. Dirik, N. Memon, and J.C. Birget. Modeling user choice in the PassPoints graphical password scheme. In *Proceedings of the 3rd Symposium on Usable Privacy and Security*, pages 20–28. ACM, 18-20 July 2007.

[73] W.F. Dukes and W. Bevan. Stimulus variation and repetition in the acquisition of naming responses. *Journal of experimental psychology*, 74(2p1): 178, 1967.

[74] P. Dunphy and J. Yan. Do Background Images Improve "Draw a Secret" Graphical Passwords? In *Proceedings of the 14th ACM Conference on Computer and Communications Security*, pages 36–47. ACM, October 2007.

[75] P. Dunphy, A. Fitch, and P. Olivier. Gaze-contingent passwords at the ATM. In *4th Conference on Communication by Gaze Interaction (COGAIN)*, 2008.

[76] P. Dunphy, A.P. Heiner, and N. Asokan. A Closer Look at Recognition-based Graphical Passwords on Mobile Devices. In *Proceedings of the 6th Symposium on Usable Privacy and Security*. ACM, 2010.

[77] T. Eisenberg, I.D. Gries, J. Hartmanis, D. Holcomb, and M.S. Lynn. The cornell commission: on morris and the worm. *Communications of the ACM*, 32(6):706–709, 1989.

[78] Y. Eisenthal, G. Dror, and E. Ruppin. Facial Attractiveness: Beauty and the Machine. *Neural Computation*, 18(1):119–142, 2006. URL `http://www.mitpressjournals.org/doi/abs/10.1162/089976606774841602`.

[79] HD Ellis. Introduction to Aspects of Face Processing: Ten Questions in Need of Answers. *Aspects of Face Processing*, pages 3–13, 1986.

[80] Federal Trade Commission. Federal Trade Commission, Plaintiff, v. Wyndham Worldwide Corporation; Wyndham Hotel Group, LLC; Wyndham Hotels & Resorts, LLC; and Wyndham Hotel Management, Inc., Defendants (United States District Court for the District of Arizona), August 2012. URL `http://www.ftc.gov/os/caselist/1023142/120809wyndhamcmpt.pdf`.

[81] Bernhard Fink and Ian Penton-Voak. Evolutionary Psychology of Facial Attractiveness. *Current Directions in Psychological Science*, 11(5):154–158, October 2002. doi: 10.1111/1467-8721.00190. URL `http://cdp.sagepub.com/lookup/doi/10.1111/1467-8721.00190`.

[82] S.T. Fiske and S.E. Taylor. *Social cognition* . Mcgraw-Hill Book Company, 1991.

[83] I. Flechais, M.A. Sasse, and S. Hailes. Bringing security home: a process for developing secure and usable systems. In *Proceedings of the 2003 workshop on New security paradigms*, pages 49–57. ACM, 2003.

[84] J.J. Fleishman, M.L Buckley, M.J. Klosinsky, N. Smith, and B. Tuck. Judged Attractiveness in Recognition Memory of Women's Faces. *Perceptual and Motor Skills*, 43(3):709–710, 1976.

[85] D. Florêncio, C. Herley, and B. Coskun. Do strong web passwords accomplish anything? In *Proceedings of the 2nd USENIX workshop on Hot topics in security*, page 10. USENIX Association, 2007.

[86] A. Forget, S. Chiasson, PC Van Oorschot, and R. Biddle. Improving text passwords through persuasion. In *Proceedings of the 4th symposium on Usable privacy and security*, pages 1–12. ACM, 2008.

[87] A. Forget, S. Chiasson, and R. Biddle. Shoulder-surfing resistance with eye-gaze entry in cued-recall graphical passwords. In *Proceedings of the 28th international conference on Human factors in computing systems*, pages 1107–1110. ACM, 2010.

[88] G.H. Forman and J. Zahorjan. The challenges of mobile computing. *Computer*, 27(4):38–47, 1994.

[89] R.L. Fowler and A.S. Barker. Effectiveness of highlighting for retention of text material. *Journal of Applied Psychology*, 59(3):358, 1974.

[90] S.M. Furnell, PS Dowland, HM Illingworth, and P.L. Reynolds. Authentication and supervision: A survey of user attitudes. *Computers & Security*, 19(6):529–539, 2000.

[91] R.E. Galper and J. Hochberg. Recognition memory for photographs of faces. *The American journal of psychology*, pages 351–354, 1971.

[92] H. Gao, X. Guo, X. Chen, L. Wang, and X. Liu. YAGP: Yet another graphical password strategy. In *Computer Security Applications Conference, 2008. ACSAC 2008. Annual*, pages 121–129. IEEE, 2008.

[93] H. Gao, Z. Ren, X. Chang, X. Liu, and U. Aickelin. A new graphical password scheme resistant to shoulder-surfing. In *Proceedings of the 2010 International Conference on Cyberworlds*, pages 194–200. IEEE Computer Society, 2010.

[94] Yongsheng Gao, Maylor KH Leung, Siu Cheung Hui, and Mario W Tananda. Facial expression recognition from line-based caricatures. *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*, 33(3):407–412, 2003.

[95] I. Gauthier and N.K. Logothetis. Is face recognition not so unique after all? *Cognitive Neuropsychology*, 17(1-3):125–142, 2000.

[96] R. Gavriloaie, W. Nejdl, D. Olmedilla, K. Seamons, and M. Winslett. No registration needed: How to use declarative policies and negotiation to access sensitive resources on the semantic web. *The Semantic Web: Research and Applications*, pages 342–356, 2004.

[97] D. Gehrke and E. Turban. Determinants of successful website design: relative importance and recommendations for effectiveness. In *System Sciences, 1999. HICSS-32. Proceedings of the 32nd Annual Hawaii International Conference on*, pages 8–pp. IEEE, 1999.

[98] Marcia Gibson, Karen Renaud, Marc Conrad, and Carsten Maple. Musipass: authenticating me softly with my song. In *Proceedings of the 2009 workshop on New security paradigms workshop*, pages 85–100. ACM, 2009.

[99] Jim Giles. Internet encyclopaedias go head to head. *Nature*, 438(7070): 900–901, 2005.

[100] Peter N. Glaskowsky. The casio cassiopeia e-10 palm-sized pc: Some first impressions, December 2005. URL `http://www.jerrypournelle.com/reports/peter_g/Peter1.html`.

[101] Von Goethe and Johann Wolfgang. *Theory of Colours*. Cambridge, MA: MIT Press, 1970.

[102] J. Goldberg, J. Hagman, and V. Sazawal. Doodling Our Way to Better Authentication. In *CHI'02 Extended Abstracts on Human Factors in Computing Systems*, pages 868–869. ACM, 2002.

[103] A.G. Goldstein and J.E. Chance. Visual recognition memory for complex configurations. *Attention, Perception, & Psychophysics*, 9(2):237–241, 1971.

[104] Google. Introducing Google Play: All your entertainment, anywhere you go, March 2012. URL `http://googleblog.blogspot.co.uk/2012/03/introducing-google-play-all-your.html`.

[105] C.C. Goren, M. Sarty, and P.Y.K. Wu. Visual following and pattern discrimination of face-like stimuli by newborn infants. *Pediatrics*, 56(4):544–549, 1975.

[106] V. Griffith and M. Jakobsson. Messin'with texas deriving mother's maiden names using public records. In *Applied Cryptography and Network Security*, pages 91–103. Springer, 2005.

[107] J. Grossklags and A. Acquisti. When 25 cents is too much: An experiment on willingness-to-sell and willingness-to-protect personal information. In *Workshop on the Economics of Information Security (WEIS)*, 2007.

[108] C. Gutwin and C. Fedak. Interacting with big interfaces on small screens: a comparison of fisheye, zoom, and panning techniques. In *Proceedings of Graphics Interface 2004*, pages 145–152. Canadian Human-Computer Communications Society, 2004.

[109] W.J. Haga and M. Zviran. Question-and-answer passwords: an empirical evaluation. *Information systems*, 16(3):335–343, 1991.

[110] N. Haller, C. Metz, P. Nesser, and M. Straw. A one-time password system, 1996.

[111] David Harrison. Digital Rights Management on the HD Freeview platform. Technical Report PE370(09), Ofcom, December 2009.

[112] J. Hartley and I.K. Davies. Note-taking: A critical review. *Programmed Learning and Educational Technology*, 15(3):207–224, 1978.

[113] E. Hayashi, R. Dhamija, N. Christin, and A. Perrig. Use your illusion: secure authentication usable anywhere. In *Proceedings of the 4th symposium on Usable privacy and security*, pages 35–45. ACM, 2008.

[114] Z. He, L. Jin, L. Zhen, and J. Huang. Gesture recognition based on 3D accelerometer for cell phones interaction. In *Circuits and Systems, 2008. APCCAS 2008. IEEE Asia Pacific Conference on*, pages 217–220. IEEE, 2008.

[115] P.A. Henry. Two-factor authentication–a look behind the headlines. *Network Security*, 2006(4):18–19, 2006.

[116] C. Herley. So long, and no thanks for the externalities: the rational rejection of security advice by users. In *Proceedings of the 2009 workshop on New security paradigms workshop*, pages 133–144. ACM, 2009.

[117] Lily E Hirsch. Weaponizing classical music: Crime prevention and symbolic power in the age of repetition. *Journal of Popular Music Studies*, 19(4): 342–358, 2007.

[118] B. Hoanca and K. Mock. Secure graphical password system for high traffic public areas. In *Proceedings of the 2006 symposium on Eye tracking research & applications*, pages 35–35. ACM, 2006.

[119] P. Hoonakker, N. Bornoe, and P. Carayon. Password authentication from a human factors perspective: Results of a survey among end-users. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 2009.

[120] IBM Technical Disclosure Bulletin. Menu Icon with Hidden Geometric Password. Technical Report 10B, IBM, March 1990.

[121] M Ida Gobbini, Ellen Leibenluft, Neil Santiago, and James V Haxby. Social and emotional attachment in the neural representation of faces. *NeuroImage*, 22(4):1628–1635, August 2004. doi: 10.1016/j.neuroimage.2004.03.049. URL `http://linkinghub.elsevier.com/retrieve/pii/S1053811904002058`.

[122] Information Commissioner's Office. Data Protection Principles, 2013. URL `http://www.ico.org.uk/for_organisations/data_protection/the_guide/principle_7`.

[123] P.G. Inglesant and M.A. Sasse. The true cost of unusable password policies: password use in the wild. In *Proceedings of the 28th international conference on Human factors in computing systems*, pages 383–392. ACM, 2010.

[124] H. Intraub and S. Nicklos. Levels of Processing and Picture Memory: The Physical Superiority Effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 11(2):284–298, 1985.

[125] Phillip Isola, Jianxiong Xiao, Antonio Torralba, and Aude Oliva. What makes an image memorable? In *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, pages 145–152. IEEE, 2011.

[126] B. Ives, K.R. Walsh, and H. Schneider. The domino effect of password reuse. *Communications of the ACM*, 47(4):75–78, 2004.

[127] A.K. Jain, A. Ross, and S. Prabhakar. An Introduction to Biometric Recognition. *IEEE Transactions on Circuits and Systems for Video Technology*, 14(1):4–20, January 2004.

[128] A.K. Jain, A. Ross, and S. Pankanti. Biometrics: a tool for information security. *IEEE Transactions on Information Forensics and Security*, 1(2): 125–143, June 2006.

[129] Petr Janata, Stefan T Tomic, and Sonja K Rakowski. Characterisation of music-evoked autobiographical memories. *Memory*, 15(8):845–860, 2007.

[130] Lutz Jäncke. Music, memory and emotion lutz jäncke. *Journal of biology*, 7:21, 2008.

[131] C. Jennett, S. Brostoff, M. Malheiros, and M.A. Sasse. Investigating loan applicants' perceptions of alternative data items and the effect of incentives on disclosure. *Privacy and Usability Methods Pow-wow (PUMP) workshop at 24th BCS Conference on Human Computer Interaction*, 2010.

[132] I. Jermyn, A. Mayer, F. Monrose, M.K. Reiter, and A.D. Rubin. The Design and Analysis of Graphical Passwords. In *Proceedings of the 8th USENIX Security Symposium*, pages 1–14. Washington DC, 23-26 August 1999.

[133] A.T.B. Jin, D.N.C. Ling, and A. Goh. Biohashing: two factor authentication featuring fingerprint data and tokenised random number. *Pattern recognition*, 37(11):2245–2255, 2004.

[134] D.L. Jobusch and A.E. Oldehoeft. A Survey of Password Mechanisms: Weaknesses and Potential Improvements. Part 1. *Computers & Security*, 8 (7):587–604, 1989.

[135] D.L. Jobusch and A.E. Oldehoeft. A Survey of Password Mechanisms: Weaknesses and Potential Improvements. Part 2. *Computers & Security*, 8 (8):675–689, 1989.

[136] M.H. Johnson, S. Dziurawiec, H. Ellis, and J. Morton. Newborns' preferential tracking of face-like stimuli and its subsequent decline. *Cognition*, 40 (1):1–19, 1991.

[137] Pierre Jolicoeur, Mark A Gluck, and Stephen M Kosslyn. Pictures and names: Making the connection. *Cognitive psychology*, 16(2):243–275, 1984.

[138] Jonathan Dann. Under the hood: Rebuilding Facebook for iOS, August 2012. URL `http://www.facebook.com/notes/facebook-engineering/under-the-hood-rebuilding-facebook-for-ios/10151036091753920`.

[139] M. Jones, G. Marsden, N. Mohd-Nasir, K. Boone, and G. Buchanan. Improving web interaction on small displays. *Computer Networks*, 31(11): 1129–1137, 1999.

[140] M. Jones, G. Buchanan, and H. Thimbleby. Sorting out searching on small screen devices. *Human Computer Interaction with Mobile Devices*, pages 555–567, 2002.

[141] A. Josang, B. AlFayyadh, T. Grandison, M. AlZomai, and J. McNamara. Security usability principles for vulnerability analysis and risk assessment. In *Computer Security Applications Conference, 2007. ACSAC 2007. Twenty-Third Annual*, pages 269–278. IEEE, 2007.

[142] A. Jøsang, M.A. Zomai, and S. Suriadi. Usability and Privacy in Identity Management Architectures. In *Proceedings of the 5th Australasian symposium on ACSW frontiers*, volume 68, pages 143–152, Ballarat, Australia,

June 2007. Australian Computer Society, Inc. URL http://dl.acm.org/citation.cfm?id=1274548.

[143] M. Just. Designing and evaluating challenge-question systems. *Security & Privacy, IEEE*, 2(5):32–39, 2004.

[144] M. Just. Designing authentication systems with challenge questions. *Security and Usability: Designing Secure Systems That People Can Use*, pages 143–155, 2005.

[145] M. Just and D. Aspinall. Challenging challenge questions. In *Paper for the Trust 2009 Conference, Socio-Economic Strand*, pages 6–8, 2009.

[146] M. Just and D. Aspinall. Personal choice and challenge questions: a security and usability assessment. In *Proceedings of the 5th Symposium on Usable Privacy and Security*, page 8. ACM, 2009.

[147] S. Kallio, J. Kela, P. Korpipää, and J. Mäntyjärvi. User independent gesture interaction for small handheld devices. *International Journal of Pattern Recognition and Artificial Intelligence*, 20(04):505–524, 2006.

[148] T. Kamba, S.A. Elson, T. Harpold, T. Stamper, and P. Sukaviriya. Using small screen space more efficiently. In *Proceedings of the SIGCHI conference on Human factors in computing systems: common ground*, pages 383–390. ACM, 1996.

[149] N. Kanwisher, J. McDermott, and M.M. Chun. The Fusiform Face Area: A Module in Human Extrastriate Cortex Specialized for Face Perception. *The Journal of Neuroscience*, 17(11):4302–4311, 1997.

[150] L. Kärkkäinen and J. Laarni. Designing for small display screens. In *Proceedings of the second Nordic conference on Human-computer interaction*, pages 227–230. ACM, 2002.

[151] D. Katre. One-handed thumb use on smart phones by semi-literate and illiterate users in india. *Human Work Interaction Design: Usability in Social, Cultural and Organizational Contexts*, pages 189–208, 2010.

[152] K.A. Kiewra. Investigating notetaking and review: A depth of processing alternative. *Educational Psychologist*, 20(1):23–32, 1985.

[153] Maria M King. Rebus passwords. In *Computer Security Applications Conference, 1991. Proceedings., Seventh Annual*, pages 239–243. IEEE, 1991.

[154] H. Kinjo and J.G. Snodgrass. Is there a picture superiority effect in perceptual implicit tasks? *European Journal of Cognitive Psychology*, 2000.

[155] D.V. Klein. "Foiling the Cracker": A Survey of, and Improvements to, Password Security. In *Proceedings of the 2nd USENIX Security Workshop*, pages 5–14, 1990.

[156] Talia Konkle, Timothy F Brady, George A Alvarez, and Aude Oliva. Conceptual distinctiveness supports detailed visual long-term memory for real-world objects. *Journal of Experimental Psychology: General*, 139(3):558, 2010.

[157] Gregg Kreizman and Ant Allan. Toolkit: Evaluating Enterprise Options for Managing Passwords, November 2006. URL `http://www.gartner.com/id=498322`.

[158] W.C. Ku and M.J. Tsaur. A Remote User Authentication Scheme Using Strong Graphical Passwords. In *Proceedings of the IEEE Conference on Local Computer Networks*, pages 351–357. IEEE, 2005.

[159] R. Kuber and W. Yu. Tactile vs graphical authentication. *Haptics: Generating and Perceiving Tangible Sensations*, pages 314–319, 2010.

[160] Ravi Kuber and Shiva Sharma. Toward tactile authentication for blind users. In *Proceedings of the 12th international ACM SIGACCESS conference on Computers and accessibility*, pages 289–290. ACM, 2010.

[161] Ravi Kuber and Wai Yu. Authentication using tactile feedback. *Interactive Experiences, HCI, London, UK*, 2006.

[162] M. Kumar, T. Garfinkel, D. Boneh, and T. Winograd. Reducing Shoulder-surfing by Using Gaze-based Password Entry. In *Proceedings of the 3rd symposium on Usable Privacy and Security*, pages 13–19. ACM, 2007.

[163] J. G. Künzler. OS X Lion Provides Multi-Touch Tutorial On First Boot, July 2011. URL `http://www.mactrast.com/2011/07/os-x-lion-provides-multi-touch-tutorial-on-first-boot/`.

[164] S.A. Kurzban. Easily remembered passphrases: a better approach. *ACM SIGSAC Review*, 3(2-4):10–21, 1985.

[165] L. Lamport. Password authentication with insecure communication. *Communications of the ACM*, 24(11):770–772, 1981.

[166] J.H. Langlois, L. Kalakanis, A.J. Rubenstein, A. Larson, M. Hallam, and M. Smoot. Maxims or myths of beauty? a meta-analytic and theoretical review. *Psychological bulletin*, 126(3):390, 2000.

[167] Joseph E LeDoux. Emotion as memory: Anatomical systems underlying indelible neural traces. *The handbook of emotion and memory: Research and theory*, pages 269–288, 1992.

[168] M.B. Lewis. Are caricatures special? evidence of peak shift in face recognition. *European Journal of Cognitive Psychology*, 11(1):105–117, 1999.

[169] Jim Liddell, Karen Renaud, and Antonella De Angeli. Using a combination of sound and images to authenticate web users. In *17th Annual Human Computer Interaction Conference: Designing for Society, Bath England*, 2003.

[170] D. Lin, P. Dunphy, P. Olivier, and J. Yan. Graphical Passwords & Qualitative Spatial Relations. In *Proceedings of the 3rd Symposium on Usable Privacy and Security*, pages 161–162. ACM, 18-20 July 2007.

[171] R.W. Lindner et al. Highlighting text as a study strategy: Beyond attentional focusing. Technical report, Annual Meeting of the American Educational Research Association, April 8–12 1996.

[172] S. Liu and M. Silverman. A practical guide to biometric security technology. *IT Professional*, 3(1):27–32, 2001.

[173] M. Long and U. Blumenthal. Manageable one-time password for consumer applications. In *Consumer Electronics, 2007. ICCE 2007. Digest of Technical Papers. International Conference on*, pages 1–2. IEEE, 2007.

[174] Joseph Maguire and Karen Renaud. You only live twice or the years we wasted caring about shoulder-surfing. In *Proceedings of the 26th Annual BCS Interaction Specialist Group Conference on People and Computers*, pages 404–409. British Computer Society, 2012.

[175] Paul Mahrer and Christopher Miles. Memorial and strategic determinants of tactile recency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25(3):630, 1999.

[176] Paul Mahrer and Christopher Miles. Recognition memory for tactile sequences. *Memory*, 10(1):7–20, 2002.

[177] Miguel Malheiros, Sascha Brostoff, Charlene Jennett, and Angela Sasse. Would You Sell Your Mother's Data? Personal Data Disclosure in a Simulated Credit Card Application. In *Workshop on the Economics of Information Security (WEIS)*, 2012.

[178] D.R. Malone, H.H. Morris, M.C. Kay, and H.S. Levin. Prosopagnosia: a double dissociation between the recognition of familiar and unfamiliar faces. *Journal of Neurology, Neurosurgery & Psychiatry*, 45(9):820–822, 1982.

[179] Christian A. Meissner and John C. Brigham. Thirty Years of Investigating the Own-race Bias in Memory for Faces: A Meta-analytic Review. *Psychology, Public Policy, and Law*, 7(1):3–35, 2001.

[180] P Meissner. Guidelines on Evaluation of Techniques for Automated Personal Identification. Technical Report FIPS PUB 48, US Department of Commerce, National Bureau of Standards, April 1977.

[181] Susanna Millar. Memory in touch. *Psicothema*, 11(4):747–767, 1999.

[182] G.A. Miller. The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological review*, 63(2):81, 1956.

[183] R. Morris and K. Thompson. Password security: A case history. *Communications of the ACM*, 22(11):594–597, November 1979.

[184] S.J. Murdoch, S. Drimer, R. Anderson, and M. Bond. Chip and pin is broken. In *Security and Privacy (SP), 2010 IEEE Symposium on*, pages 433–446. IEEE, 2010.

[185] D. Nali and J. Thorpe. Analyzing user choice in graphical passwords. *School of Computer Science, Carleton University, Tech. Rep. TR-04-01*, 2004.

[186] D. Nasaw. Meet the 'bots' that edit Wikipedia., July 2012. URL `http://www.bbc.co.uk/news/magazine-18892510`.

[187] R.M. Needham and M.D. Schroeder. Using encryption for authentication in large networks of computers. *Communications of the ACM*, 21(12):993–999, 1978.

[188] D.L. Nelson, V.S. Reed, and J.R. Walling. Pictorial superiority effect. *Journal of Experimental Psychology: Human Learning and Memory*, 2(5):523, 1976.

[189] J. Nicholson, P. Dunphy, L. Coventry, P. Briggs, and P. Olivier. A security assessment of tiles: a new portfolio-based graphical authentication system. In *Proceedings of the 2012 ACM annual conference extended abstracts on Human Factors in Computing Systems Extended Abstracts*, pages 1967–1972. ACM, 2012.

[190] R.S. Nickerson. Short-term memory for complex meaningful visual configurations: A demonstration of capacity. *Canadian Journal of Psychology*, 19 (2):155, 1965.

[191] R.S. Nickerson. A note on long-term recognition memory for pictorial material. *Psychonomic Science*, 1968.

[192] Jakob Nielsen. Why You Only Need to Test with 5 Users, March 2000.

[193] Ofcom. The Ofcom Broadcasting Code, 2003. URL `http://stakeholders.ofcom.org.uk/broadcasting/broadcast-codes/broadcast-code/protecting-under-18s/`.

[194] L. O'Gorman. Comparing passwords, tokens, and biometrics for user authentication. *Proceedings of the IEEE*, 91(12):2021–2040, 2003.

[195] M. Oka, K. Kato, Y. Xu, L. Liang, and F. Wen. Scribble-a-Secret: Similarity-based Password Authentication Using Sketches. In *Proceedings of the 19th International Conference on Pattern Recognition*, pages 1–4. IEEE, 2008.

[196] Aude Oliva. Gist of the scene. *Neurobiology of attention*, 696:64, 2005.

[197] M. Orozco, B. Malek, M. Eid, and A. El Saddik. Haptic-based sensible graphical password. In *Proceedings of the 2006 Virtual Concept*. Citeseer, November 2006.

[198] A. Paivio. *Mental representations: A dual coding approach*. Oxford University Press, USA, 1990.

[199] S. Pankanti, R.M. Bolle, and A. Jain. Biometrics: The future of identification [guest eeditors' introduction]. *Computer*, 33(2):46–49, 2000.

[200] Devi Parikh, Phillip Isola, Antonio Torralba, and Aude Oliva. Understanding the intrinsic memorability of images. *Journal of Vision*, 12(9):1082–1082, 2012.

[201] R.J. Peper and R.E. Mayer. Note taking as a generative activity. *Journal of Educational Psychology*, 70(4):514, 1978.

[202] Isabelle Peretz, Anne J Blood, Virginia Penhune, and Robert Zatorre. Cortical deafness to dissonance. *Brain*, 124(5):928–940, 2001.

[203] T. Pering, M. Sundar, J. Light, and R. Want. Photographic authentication through untrusted terminals. *Pervasive Computing, IEEE*, 2(1):30–36, 2003.

[204] D. Perito, C. Castelluccia, M. Kaafar, and P. Manils. How unique and traceable are usernames? In *Privacy Enhancing Technologies*, pages 1–17. Springer, 2011.

[205] DI Perrett, ET Rolls, and W. Caan. Visual neurones responsive to faces in the monkey temporal cortex. *Experimental brain research*, 47(3):329–342, 1982.

[206] D.I. Perrett, A.J. Mistlin, and A.J. Chitty. Visual neurones responsive to faces. *Trends in Neurosciences*, 10(9):358–364, 1987.

[207] Peter Farago. For Generating App Revenue, Amazon Shows Google How to Play, March 2012. URL `http://blog.flurry.com/bid/83604/For-Generating-App-Revenue-Amazon-Shows-Google-How-to-Play`.

[208] A. Pirhonen, S. Brewster, and C. Holguin. Gestural and audio metaphors as a means of control for mobile devices. In *Proceedings of the SIGCHI conference on Human factors in computing systems: Changing our world, changing ourselves*, pages 291–298. ACM, 2002.

[209] K. Pitts and N. Hurst. How much do people prefer widescreen (16× 9) to standard ntsc (4× 3)? *Consumer Electronics, IEEE Transactions on*, 35 (3):160–169, 1989.

[210] L.Y. Por and X.T. Lim. Multi-grid Background Pass-Go. *WSEAS Transactions on Information Science and Applications*, 5(7):1137–1148, 2008.

[211] L.Y. Por, X.T. Lim, and F. Kianoush. Background Pass-Go (BPG), a New Approach for GPS. In *Proceedings of the 12th WSEAS international conference on Computers*, pages 369–374. World Scientific and Engineering Academy and Society (WSEAS), 2008.

[212] S.N. Porter. A Password Extension for Improved Human Factors. *Computers & Security*, 1(1):54–56, 1982.

[213] S. Prabhakar, S. Pankanti, and A.K. Jain. Biometric recognition: Security and privacy concerns. *Security & Privacy, IEEE*, 1(2):33–42, 2003.

[214] A. Rabkin. Personal knowledge questions for fallback authentication: Security questions in the era of Facebook. In *Proceedings of the 4th symposium on Usable privacy and security*, pages 13–23. ACM, 2008.

[215] K. Rao and S. Yalamanchili. Novel shoulder-surfing resistant authentication schemes using text-graphical passwords. *International Journal of Information and Network Security (IJINS)*, 1(3):163–170, 2012.

[216] Abdullah Rashed and Henrique Santos. Odour user interface for authentication: Possibility and acceptance: Case study. In *The International MultiConference of Engineers and Computer Scientists*, 2010.

[217] C. Rathgeb and A. Uhl. Two-factor authentication or how to potentially counterfeit experimental results in biometric systems. *Image Analysis and Recognition*, pages 296–305, 2010.

[218] J. Rekimoto. Tilting operations for small screen interfaces. In *Proceedings of the 9th annual ACM symposium on User interface software and technology*, pages 167–168. ACM, 1996.

[219] K. Renaud. On user involvement in production of images used in visual authentication. *Journal of Visual Languages & Computing*, 20(1):1–15, 2009.

[220] K. Renaud and A. De Angeli. My password is here! An investigation into visuo-spatial authentication mechanisms. *Interacting with computers*, 16 (6):1017–1041, 2004.

[221] K. Renaud and M. Just. Pictures or questions?: examining user responses to association-based authentication. In *Proceedings of the 24th BCS Interaction Specialist Group Conference*, pages 98–107. British Computer Society, 2010.

[222] K. Renaud and J. Maguire. Shrinking the Authentication Footprint. *International Symposium on Human Aspects of Information Security & Assurance*, May 2013.

[223] Karen Renaud. Web authentication using mikon images. In *Privacy, Security, Trust and the Management of e-Business, 2009. CONGRESS'09. World Congress on*, pages 79–88. IEEE, 2009.

[224] Karen Renaud and Joseph Maguire. Armchair authentication. In *BCS-HCI '09: Proceedings of the 23rd British HCI Group Annual Conference on People and Computers: Celebrating People and Technology*. British Computer Society, September 2009. URL http://portal.acm.org/citation.cfm?id=1671011.1671061&coll=DL&dl=ACM&CFID=81475699&CFTOKEN=79301160.

[225] KV Renaud. Guidelines for designing graphical authentication mechanism interfaces. *International Journal of Information and Computer Security*, 3 (1):60–85, 2009.

[226] G. Rhodes. When do caricatures look good? *New Zealand Journal of Psychology*, 1993.

[227] G. Rhodes and R. Wooding. Laterality effects in identification of caricatures and photographs of famous faces. *Brain and cognition*, 9(2):201–209, 1989.

[228] G. Rhodes, S. Brennan, and S. Carey. Identification and ratings of caricatures: Implications for mental representations of faces. *Cognitive psychology*, 19(4):473–497, 1987.

[229] G. Rhodes, G. Byatt, T. Tremewan, and A. Kennedy. Facial distinctiveness and the power of caricatures. *Perception*, 26:207–224, 1997.

[230] G. Rhodes, W.G. Hayward, and C. Winkler. Expert face coding: Configural and component coding of own-race and other-race faces. *Psychonomic Bulletin & Review*, 13(3):499–505, 2006.

[231] B.L. Riddle, M.S. Miron, and J.A. Semo. Passwords in use in a university timesharing environment. *Computers and Security*, 8(7):569–578, 1989.

[232] Uri Rivner. Anatomy of an Attack, April 2011. URL http://blogs.rsa.com/rivner/anatomy-of-an-attack/.

[233] V. Roth, K. Richter, and R. Freidinger. A pin-entry method resilient against shoulder surfing. In *Proceedings of the 11th ACM conference on Computer and communications security*, pages 236–245. ACM, 2004.

[234] RSA. RSA SecurID Solution Named Best Third-Party Authentication Device by Windows IT Pro Magazine Readers' Choice 2004, September 2004. URL http://www.rsa.com/press_release.aspx?id=5028.

[235] Richard Saintvilus. 40 Billion Reasons To Still Love Apple, But Only 1 Really Matters, January 2012. URL http://www.forbes.com/sites/richardsaintvilus/2013/01/07/heres-40-billion-reasons-to-still-love-apple-but-only-1-really-matters/.

[236] A. Salehi-Abari, J. Thorpe, and PC Van Oorschot. On Purely Automated Attacks and Click-Based Graphical Passwords. In *Proceedings of the 2008 Annual Computer Security Applications Conference*, pages 111–120. IEEE, 2008.

[237] H. Sasamoto, N. Christin, and E. Hayashi. Undercover: authentication usable in front of prying eyes. In *Proceedings of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*, pages 183–192. ACM, 2008.

[238] M.A. Sasse. Computer security: Anatomy of a usability disaster, and a plan for recovery. In *Proceedings of CHI 2003 Workshop on HCI and Security Systems*. Citeseer, 2003.

[239] M.A. Sasse, S. Brostoff, and D. Weirich. Transforming the 'weakest link'—a human/computer interaction approach to usable and effective security. *BT Technology Journal*, 19(3):122–131, 2001.

[240] S. Schechter, A.J.B. Brush, and S. Egelman. It's no secret. measuring the security and reliability of authentication via "secret" questions. In *Security and Privacy, 2009 30th IEEE Symposium on*, pages 375–390. IEEE, 2009.

[241] S. Schechter, S. Egelman, and R.W. Reeder. It's not what you know, but who you know: a social approach to last-resort authentication. In *Proceedings of the 27th international conference on Human factors in computing systems*, pages 1983–1992. ACM, April 2009.

[242] Klaus R Scherer and Marcel R Zentner. Emotional effects of music: Production rules. *Music and emotion: Theory and research*, pages 361–392, 2001.

[243] B. Schneier. Two-factor authentication: Too little, too late. *Communications of the ACM*, 48(4), 2005.

[244] M. Schubin. Searching for the Perfect Aspect Ratio. *Society of Motion Picture & Television Engineers Journal*, 105(8):460–478, 1996.

[245] Scott Lowe. Google Play celebrates 25 billion downloads with 25 cent apps, discounted books, music, and movies, September 2012. URL `http://www.theverge.com/2012/9/26/3409446/google-play-25-billion-downloads-sale`.

[246] W.B. Scoville and B. Milner. Loss of Recent Memory after Bilateral Hippocampal Lesions. *Journal of Neurology, Neurosurgery & Psychiatry*, 20 (1):11–21, 1957.

[247] Seeking Alpha. Apple's CEO Discusses Q2 2012 Results — Earnings Call Transcript, April 2012. URL `http://finance.yahoo.com/news/apples-ceo-discusses-q2-2012-011004832.html`.

[248] S.U. Shah, A.A. Minhas, et al. New factor of authentication: Something you process. In *Future Computer and Communication, 2009. ICFCC 2009. International Conference on*, pages 102–106. IEEE, 2009.

[249] R. Shay, S. Komanduri, P.G. Kelley, P.G. Leon, M.L. Mazurek, L. Bauer, N. Christin, and L.F. Cranor. Encountering Stronger Password Requirements: User Attitudes and Behaviors. In *Proceedings of the Sixth Symposium on Usable Privacy and Security*, page 2. ACM, 2010.

[250] R. Shay, P.G. Kelley, S. Komanduri, M.L. Mazurek, B. Ur, T. Vidas, L. Bauer, N. Christin, and L.F. Cranor. Correct horse battery staple: Exploring the usability of system-assigned passphrases. In *Proceedings of the Eighth Symposium on Usable Privacy and Security*, page 7. ACM, 2012.

[251] W.R. Shockley. Identification and Authentication when Users have Multiple Accounts. In J. Bret Michael, V. Ashby, and C. Meadows, editors, *Proceedings on the 1992-1993 workshop on New security paradigms*, pages 185–191. ACM, 1992. ISBN 0818654309. URL `http://dl.acm.org/citation.cfm?id=283846`.

[252] J. Siracusa. Paths in the grass. We have created, for the first time in all history, a garden of pure ideology, February 2006. URL `http://arstechnica.com/staff/2006/02/2918/`.

[253] S.L. Smith. Authenticating Users by Word Association. *Computers & Security*, 6(6):464–470, 1987.

[254] Merrielle Spain and Pietro Perona. Some objects are more equal than others: Measuring and predicting importance. In *Computer Vision–ECCV 2008*, pages 523–536. Springer, 2008.

[255] Y. Spector and J. Ginzberg. Pass-sentence–a new approach to computer code. *Computers & Security*, 13(2):145–160, 1994.

[256] JC Spender. Identifying computer users with authentication devices (tokens). *Computers & Security*, 6(5):385–395, 1987.

[257] L. Standing. Learning 10000 pictures. *The Quarterly journal of experimental psychology*, 25(2):207–222, 1973.

[258] Jonathan Stempel. Amazon seeks to throw out Apple "app store" advertising claim, September 2012. URL `http://uk.reuters.com/article/2012/09/27/uk-apple-amazon-appstore-lawsuit-idUKBRE88Q00620120927`.

[259] G. Stoneburner, A. Goguen, and A. Feringa. Risk Management Guide for Information Technology Systems. Technical Report 30, National Institute of Standards and Technology, 2002.

[260] D.S. Tan, P. Keyani, and M. Czerwinski. Spy-resistant Keyboard: more secure password entry on public touch screen displays. In *Proceedings of the 17th Australia conference on Computer-Human Interaction: Citizens Online: Considerations for Today and the Future*, pages 1–10. ACM, 2005.

[261] J.W. Tanaka. The Entry Point of Face Recognition: Evidence for Face Expertise. *Journal of Experimental Psychology: General*, 130(3):534, 2001.

[262] H. Tao and C. Adams. Pass-Go: A Proposal to Improve the Usability of Graphical Passwords. *International Journal of Network Security*, 7(2): 273–292, 2008.

[263] F. Tari, A. Ozok, and S.H. Holden. A Comparison of Perceived and Real Shoulder-surfing Risks between Alphanumeric and Graphical passwords. In *Proceedings of the 2nd Symposium on Usable Privacy and Security*, pages 56–66. ACM, 12-14 July 2006.

[264] E. Tatarunaite, R. Playle, K. Hood, W. Shaw, and S. Richmond. Facial attractiveness: A longitudinal study. *American Journal of Orthodontics and Dentofacial Orthopedics*, 127(6):676–682, June 2005. doi: 10.1016/j.ajodo.2004.01.029. URL `http://linkinghub.elsevier.com/retrieve/pii/S0889540605000594`.

[265] S. Teitelbaum and R.E. Geiselman. Observer mood and cross-racial recognition of faces. *Journal of Cross-Cultural Psychology*, 28(1):93–106, 1997.

[266] The Passfaces Corporation. The Science Behind Passfaces. *White Paper, June*, 2004.

[267] R. Thornhill and S.W. Gangestad. Human facial beauty. *Human Nature*, 4(3):237–269, 1993. URL `http://www.springerlink.com/index/G52263H761671619.pdf`.

[268] R. Thornhill and S.W. Gangestad. Facial attractiveness. *Trends in Cognitive Sciences*, 3(12):452–460, October 1999.

[269] J. Thorpe and PC Van Oorschot. Graphical dictionaries and the memorable space of graphical passwords. In *13th USENIX Security Symposium*, pages 135–150, 2004.

[270] J. Thorpe and PC Van Oorschot. Towards Secure Design Choices for Implementing Graphical Passwords. In *Proceedings of the 20th Annual Computer Security Applications Conference*, pages 50–60. IEEE, 06-10 December 2004.

[271] J. Thorpe and P.C. van Oorschot. Human-Seeded Attacks and Exploiting Hot-Spots in Graphical Passwords. In *Proceedings of the 16th USENIX Security Symposium*, page 8. USENIX Association, 06-10 August 2007.

[272] J. Thorpe, PC Van Oorschot, and A. Somayaji. Pass-thoughts: Authenticating with Our Minds. In *Proceedings of the 2005 workshop on New Security Paradigms*, pages 45–56. ACM, 2005.

[273] S. Thorpe, D. Fize, C. Marlot, et al. Speed of processing in the human visual system. *nature*, 381(6582):520–522, 1996.

[274] Amar Toor. Teardown reveals Android smartphone inside Entertainment Weekly's print ad, October 2012. URL `http://www.theverge.com/2012/10/4/3452890/entertainment-weekly-cw-ad-smartphone-teardown`.

[275] E. Tulving and M.J. Watkins. Continuity between recall and recognition. *The American Journal of Psychology*, pages 739–748, 1973.

[276] B. Tversky and D. Baratz. Memory for faces: Are caricatures better than photographs? *Memory & cognition*, 13(1):45–49, 1985.

[277] B.J. Underwood. Interference and forgetting. *Psychological review*, 64(1): 49, 1957.

[278] B. Ur, P.G. Kelley, S. Komanduri, J. Lee, M. Maass, M. Mazurek, T. Passaro, R. Shay, T. Vidas, L. Bauer, et al. How does your password measure up? the effect of strength meters on password creation. In *Proc. USENIX Security*, 2012.

[279] T. Valentine. Upside-down faces: A review of the effect of inversion upon face recognition. *British Journal of Psychology*, 79(4):471–491, 1988.

[280] T. Valentine and V. Bruce. The effect of race, inversion and encoding activity upon face recognition. *Acta psychologica*, 61(3):259–273, 1986.

[281] Tim Valentine, Vicki Bruce, et al. The effects of distinctiveness in recognising and classifying faces. *Perception*, 15(5):525–535, 1986.

[282] Fernanda B Viégas, Martin Wattenberg, and Kushal Dave. Studying cooperation and conflict between authors with history flow visualizations. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 575–582. ACM, 2004.

[283] K.P.L. Vu, A. Bhargav, and R.W. Proctor. Imposing password restrictions for multiple accounts: Impact on generation and recall of passwords. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 47, pages 1331–1335. SAGE Publications, 2003.

[284] Tom Warren. How Microsoft plans to teach the world to use Windows 8, August 2012. URL `http://www.theverge.com/2012/8/2/3214795/windows-8-tutorial-setup-guide`.

[285] J. Wayman, A. Jain, D. Maltoni, and D. Maio. An introduction to biometric authentication systems. *Biometric Systems*, pages 1–20, 2005.

[286] D. Weinshall. Cognitive authentication schemes safe against spyware. In *Security and Privacy, 2006 IEEE Symposium on*, pages 6–pp. IEEE, 2006.

[287] C.S. Weir, G. Douglas, M. Carruthers, and M. Jack. User perceptions of security, convenience and usability for ebanking authentication tokens. *Computers & Security*, 28(1-2):47–62, 2009.

[288] D. Weirich and M.A. Sasse. Pretty Good Persuasion: A First Step towards Effective Password Security in the Real World. In *Proceedings of the 2001 Workshop on New Security Paradigms*, pages 137–143. ACM, 2001.

[289] R. Weiss and A. De Luca. PassShapes: Utilizing Stroke Based Authentication to Increase Password Memorability. In *Proceedings of the 5th Nordic Conference on Human-Computer Interaction: Building Bridges*, pages 383–392. ACM, 2008.

[290] Lance Whitney. Microsoft rivals critical of browser 'ballot screen', September 2009. URL `http://news.cnet.com/8301-10805_3-10363306-75.html`.

[291] S. Wiedenbeck, J. Waters, J.C. Birget, A. Brodskiy, and N. Memon. Authentication Using Graphical Passwords: Effects of Tolerance and Image Choice. In *Proceedings of the 2005 Symposium on Usable Privacy and Security*, pages 1–12. ACM, 06-08 July 2005.

[292] S. Wiedenbeck, J. Waters, J.C. Birget, A. Brodskiy, and N. Memon. PassPoints: Design and Longitudinal Evaluation of a Graphical Password System. *International Journal of Human-Computer Studies*, 63(1):102–127, 2005.

[293] S. Wiedenbeck, J. Waters, L. Sobrado, and J.C. Birget. Design and evaluation of a shoulder-surfing resistant graphical password scheme. In *Proceedings of the working conference on Advanced visual interfaces*, pages 177–184. ACM, 2006.

[294] J.T. Wixted. The Psychology and Neuroscience of Forgetting. *Annual Review of Psychology*, 55:235–269, November 2004.

[295] H.M. Wood. The use of passwords for controlled access to computer resources. Technical Report 500-9, US Department of Commerce, National Bureau of Standards, 1977.

[296] G. Wurster and PC van Oorschot. The developer is the enemy. In *Proceedings of the 2008 workshop on New security paradigms*, pages 89–97. ACM, 2009.

[297] Jianxiong Xiao, James Hays, Krista A Ehinger, Aude Oliva, and Antonio Torralba. Sun database: Large-scale scene recognition from abbey to zoo. In *Computer vision and pattern recognition (CVPR), 2010 IEEE conference on*, pages 3485–3492. IEEE, 2010.

[298] J. Yan, A. Blackwell, R. Anderson, and A. Grant. The memorability and security of passwords-some empirical results. Technical report, University of Cambridge, 2000.

[299] J. Yan, A. Blackwell, R. Anderson, and A. Grant. Password memorability and security: Empirical results. *Security & Privacy, IEEE*, 2(5):25–31, 2004.

[300] J.J. Yan. A note on proactive password checking. In *Proceedings of the 2001 workshop on New security paradigms*, pages 127–135. ACM, 2001.

[301] R.K. Yin. Looking at upside-down faces. *Journal of experimental psychology*, 81(1):141, 1969.

[302] C. Zarcadoolas, M. Blanco, J.F. Boyer, and A. Pleasant. Unweaving the web: an exploratory study of low-literate adults' navigation skills on the world wide web. *Journal of health communication*, 7(4):309–324, 2002.

[303] W. Zhao, R. Chellappa, P.J. Phillips, and A. Rosenfeld. Face Recognition: A Literature Survey. *ACM Computing Surveys (CSUR)*, 35(4):399–458, 2003. URL `http://dl.acm.org/citation.cfm?id=954342`.

[304] M. Zviran and W.J. Haga. Cognitive Passwords: The Key to Easy Access Control. *Computers & Security*, 9(8):723–736, 1990.

[305] M. Zviran and W.J. Haga. A Comparison of Password Techniques for Multilevel Authentication Mechanisms. *The Computer Journal*, 36(3):227–237, 1993.

[306] M. Zviran and W.J. Haga. Password security: an empirical study. *Journal of Management Information Systems*, pages 161–185, 1999.