Using and managing multiple passwords: A week to a view

Grawemeyer, B. & Johnson, H.

Original citation & hyperlink:
Grawemeyer, B & Johnson, H 2011, 'Using and managing multiple passwords: A week to a view', Interacting with Computers, vol. 23, no. 3, pp. 256-267.
https://dx.doi.org/10.1016/j.intcom.2011.03.007

DOI 10.1016/j.intcom.2011.03.007
ISSN 0953-5438
ESSN 1873-7951

Publisher: Oxford University Press

Using and managing multiple passwords: A week to a view

Beate Grawemeyer & Hilary Johnson

Human Computer Interaction Group, Department of Computer Science,
University of Bath, Bath, BA2 7AY, United Kingdom

Abstract

Security policies are required that protect information from unauthorised access, and also respect challenges users face in creating, and particularly managing, increasing numbers of passwords. This paper investigates *real* password use in the context of daily life. It presents the results of an empirical study where participants completed a password diary over seven days, followed by debrief interviews to gain further knowledge and understanding of user behaviour. The results reported relate to how many passwords are in use, the types of passwords participants created, the relationships between different passwords and to sensitive services, how participants retrieved their passwords and finally, the different strategies adopted by users in their management of passwords. The paper concludes by providing a high level set of password guidelines, along with suggestions for mechanisms to support creating, encoding, retrieving and executing multiple passwords.

Keywords:
Password management, User authentication, Security.

1. Introduction

Research indicates that attacks on organisational, networked systems are increasing, (Harreld, 2001; BBC, 2010). Even in the context of increasing attacks, passwords remain the most commonly used mechanism to enable authentication. It is therefore unsurprising (Keith et al., 2007) that one means for attackers to penetrate systems is through the misuse of passwords. The threat of attack has resulted in security policies evolving with the primary aim of improving information security (Goguen and Meseguer, 1982; Casassa Mont and Thyne, 2008). Additionally, password security guidelines to be followed by users have been developed, such as advice on secure password creation, (Morris and Thompson, 1997; Zviran and Haga, 1999). However, the vulnerability of passwords remains, and research on the effect of security policies upon user behaviour (Adams and Sasse, 1999) has shown that certain security policies, enforced to increase security, have had the opposite effect. For example, an enforcement to frequently change passwords can lead to users writing passwords down to aid recall, thus detrimentally impacting security.

Despite known problems with passwords as a means of authentication, Bonneau and Preibusch (2010) provide an argument as to why passwords are still the most common authentication mechanism. They postulate that efforts to replace passwords with more secure protocols, or federated identity systems, may fail because they do not recreate the entrenched ritual of password authentication. According to Bonneau and Preibusch, these rituals persist independent of their utility on the modern web, and establishing a feeling of trust with respect to web applications, may be a primary function of passwords for users. Accepting this argument means that passwords have (at least) dual roles - *authentication* and *engendering trust*, and these might have different consequences for organisations, and for individual users and their subsequent security behaviour. Changing authentication mechanisms from passwords in order to improve security therefore also means considering other mechanisms for establishing trust, and both establishing and maintaining trust is a major concern for web application designers, and their clients.

Due to the rapid growth of the Internet, users now need more passwords, or some other means by which to satisfy an associated growth in instances of password authentication. An increasingly large number of online accounts, for social networking, shopping, etc. require users to create and remember an increasingly large number of different passwords. However, whilst users have to cope with this demand, the security policies of individual organisations or applications make no allowance for the cognitive load on users of maintaining multiple passwords, and act as though users only have a single password: that used for their particular service. Policies in general take no account of any other passwords the user must memorize. Consequently, password management is the sole responsibility of the user. Policies advise how to create 'more secure' passwords but not how to create multiple, secure passwords which are sufficiently different and do not interfere with each other, but which are nevertheless easy to retrieve from memory, and execute. Given the likelihood of the continuing popularity of passwords, following the argument of Bonneau and Preibusch, and the increasing number of passwords needed, it is necessary to know about password use in reality. How many passwords must users manage and how is this currently achieved? The need for accurate information arises since it is necessary to generate more effective security policies. These policies must provide guidelines for users which not only help to ensure that passwords are resilient to attacks, but also take into account actual user behaviour in relation to passwords, particularly in the context of multiple use of passwords within users' everyday lives.

In this paper the results of an empirical study conducted over a 7-day period to identify *actual user behaviour* in relation to passwords, are reported. Data were collected and analysed concerning the number of passwords in use, the user activities and services they are associated with, and how the different passwords are related, and managed.

In the next section an overview of relevant background literature is provided. Section 2 outlines the empirical study and the study results are reported in section 3. In section 4 the implications of the results are discussed, whilst section 5 concludes the paper.

1.1. Background

Whilst passwords are used to prevent unauthorised access, research has shown that people can show very lax security behaviour. For example, the BBC (2004) published study results demonstrating that over 70% of people would reveal their passwords for chocolate, and 79% would give away personal information (such as their mother's maiden name, or their date of birth). Additionally, Schneier (2006) found that although people are creating more secure passwords than ten years ago, with regard to length and nature, new cracking methods are able to hack passwords easily.

Different studies have investigated users' generation and selection of passwords. One early example comes from Zviran and Haga (1999) who investigated the effect of data importance and sensitivity upon password generation. They found that users frequently change their passwords if the data is perceived as important or sensitive. However, no association was found between the nature of user-generated passwords, such as character length or composition, and the perceived importance or sensitivity of the data. The study reported in this paper will address the question as to whether this is still the case a decade later, whether it is common across users, and what might this be due to? Explanations may relate to the lack of perception about sensitivity of the service, about the need to ensure passwords to sensitive services are secure, and/or a lack of knowledge about what constitutes a secure password for these services.

Both a lack of relevant security knowledge and absence of unpleasant security breaches have been used to explain users' lax security behaviour. For instance, whilst Zviran argues that users are 'naive' and ignore the importance or sensitivity of the data whilst generating passwords, Adams and Sasse (1999) observed that users' knowledge of password security (such as procedures, content, cracking etc.) impacts password behaviour. Similarly, Bubas (2008) found that experience with security problems affected people's behaviour towards security. The assumption was that 'relaxed' use of the Internet (including password usage) without fear (e.g. a disbelief in security threats) was due to an absence of previous unpleasant experiences.

Lax security behaviour and relaxed use of the Internet is typified by users creating passwords with the primary function of enabling easy retrieval from memory, rather than with security as a priority. Investigations of password memorability and security found that users' ability to create secure passwords that are also memorable is difficult. For example, Avarne (1988), Zviran and Haga (1990), Zviran and Haga (1999), Yan et al. (2004), and Keith et al. (2007), all demonstrated that certain characteristics (composition, selection/generation and lifetime) of passwords affect their memorability. Clearly, naively selected, user-generated passwords consisting of a common word or name, have a memorability advantage, but also the disadvantage that they are weak from a security perspective. This can be contrasted with random system-generated passwords, which are usually strong, but typically difficult to remember. Passwords that are difficult to remember (due to their nature or lifetime) have a high probability of being written down and therefore disclosed (e.g. Highland, 1991; Adams and Sasse, 1999; Zviran and Haga, 1999).

Balancing the often opposing security needs of the organisation, and the users' need to create memorable passwords, has given rise to a series of studies where password characteristics are systematically manipulated, and their implications for security and memorability assessed. In this vein, Yan et al. (2004) investigated different types of passwords users generated when given instructions to either (a) select a password with low security constraints; (b) create random passwords (randomly selected letters A-Z and numbers 1-9); or (c) select a mnemonic phrase/passphrase. They analysed the memorability and strength of the passwords. Participants had difficulties in remembering random passwords, while mnemonic phrases (passphrases) were no harder to remember than naively selected passwords, but nevertheless as secure as a random password. From the results of the study reported here we will be able to indicate whether there is any use of mnemonic or meaningful phrases in user-generated passwords, in real life.

Similar to the afore-mentioned study, Keith et al. (2007) found that passphrases were no more difficult to recall from memory than simple passwords. Additionally, they investigated authentication failures due to typographical errors. They found that passphrases result in significantly higher login failures than naively selected passwords, due to typographical errors, but that these errors reduce over time.

A related study to those of Yan et al. and Keith et al., was conducted by Wiedenbeck et al. (2005), who introduced 'passpoints' as an alternative to passwords and passphrases. Passpoints was considered to be a more secure graphical password system. Participants in their study created an alphanumeric or graphical password and used this for over 6 weeks. The results showed that in the graphical passwords condition users created a valid password with fewer difficulties than in the alphanumeric passwords condition, but whilst practicing their passwords they took longer and made more invalid password inputs than the alphanumeric users. In the longitudinal trials, the two groups performed similarly but the graphical group took more time on inputting their passwords due to their lack of perception about the precision needed in their interaction with the system. One question raised by the authors concerns whether *multiple* passpoints suffer from the same level of interference as multiple passwords.

Other research investigating alternatives to passwords is described in Weir et al. (2010), who also considered problems from the perspective of possession of multiple passwords. These authors argue that 2 factor authentication could overcome the problems with managing multiple passwords. However, there is reluctance from new users to migrate to 2 factor authentication, regardless of how easy it may be to use. The pertinent difference relates to the distinction between having to carry something with them as opposed to having to learn another password. Weir et al. partially echo the ritual and entrenchment argument provided by Bonneau and Preibusch (2010) described earlier in the paper, by arguing that new methods may either be enhanced or suffer due to being new. In addition to problems of the cultural effects of changing from passwords as the sole means of authentication, the authors also note that usability of these systems could be a problem, and some 2 factor solutions do not always extend to multiple uses (i.e. more than one use as opposed to re-use).

Re-use (BBC, 2004; Florencio and Herley, 2007) is one obvious means by which users manage the increasing demand to create online identities that require secure passwords to protect users from unauthorised access. The assumption is that people reuse the same insecure passwords frequently. Thus, hackers who obtain access to a password from a popular site might be able to use the same user-ID and password for different sites (Ives et al., 2004). Shay et al. (2010) noted significant amounts of password reuse in their comprehensive study, with over 80% of participants (381/470) reporting that they reused sets of passwords in different places, and over two thirds of these participants reused one password with slight modification for different accounts.

The survey by Shay et al. comprises a unique comparison involving the same users operating under two different security policies, instituted by Carnegie Mellon University. Their study aimed to advance understanding of factors that make creating and following password policies difficult. They discovered that users (i) were annoyed at the change to a stricter password policy, (ii) did not find it difficult to comply, and (iii) thought that the newer, complex passwords improved security. The authors note that users were 'neutral' as to whether the change was worth the effort, and whether or not the organisation should revert to the previous, more lax policy. The results are particularly telling in that even though it was not considered difficult to comply and users perceived the result to be greater personal security they were still unsure whether it was worth the extra effort. It is not known whether users reported higher perceived security as a result of heightened awareness of security/more knowledge of security as an issue (perhaps as a result of being questioned about security), or because of the psychological need to see a benefit due to the additional mental effort expended.

The finding in Shay et al.'s study that over 80% of users reuse passwords, and the small incidence (13%) of writing passwords down could potentially signal a shift in users' coping strategies towards dealing with multiple passwords by reuse, rather than by writing down. This also explains to some degree the reported ease of creating new passwords and the ease of compliance found in the study. However, the authors do not indicate whether the reported 19% who struggled to comply were users who did not adopt reuse as a coping strategy. Even more interesting is the statistic that in creating a new password 52% of people modified the old password for that service, whilst only 11% modified a password from elsewhere, and only 3.7% reused a whole password from elsewhere. There are memory implications associated with creating new passwords. We would argue that many of Shay et al.'s study participants demonstrated meta-cognitive awareness of (and attempted to limit) the cognitive burden and effort required to learn a new password, by making a clear mnemonic association between the old and new passwords for this service. Thus the old password functions as a prime or becomes a hint for the retrieval of the new password. This could partially explain why users reported it was not difficult to comply with the new stricter policy but were still unsure as to whether the small amount of cognitive effort of modifying a password was worth any gain in security.

A study by Dhamija and Perrig (2000) specifically investigated password management by reuse. They interviewed 30 people and estimated that 1-7 passwords were used for 10-50 websites. Similarly, in Brown et al.'s (2004) survey users were asked to login and count the number of passwords with the reported result that users had 8.18 password uses with only 4.45 of these being unique. Further, Gaw and Felten (2006) report that having to cope with multiple passwords is positively correlated with password reuse. They found that participants reused passwords over multiple accounts. Forty-nine undergraduates participated in their study which was part laboratory exercise and part survey. The majority of participants had three or more passwords and these were reused twice. Gaw and Felten contend that accumulation of accounts means more instances of reuse not more passwords being created. Specifically, they found that reuse rates were positively correlated with the number of accounts the participants held, but also even with few accounts passwords were reused. The study reported in this paper will specifically investigate the degree of reuse in the context of the number of passwords users possess.

The previous research, and that reported by Gaw and Felten is thorough. The latter involves the investigation of different solutions users could employ – e.g. the use of different types of password, and/or the use of tools to manage passwords. The authors report that using themed passwords was relatively unpopular as the median use of related passwords was zero, unlike the results reported by Shay et al. (2010). Technological solutions were found to be unpopular, but the authors note that these solutions have disadvantages anyway. Participants in their study relied on their memory for password management. Again, it will be interesting to see if these findings are replicated in the present study.

Gaw and Felten consider solutions to the problem that could be recommended in future. These solutions relate mainly to individual passwords. The first concerns changing the way in which authentication occurs, although the arguments of Bonneau and Preibusch, and Weir et al., still apply. The second concerns password creation: this could involve the system helping users in the task of secure password generation. The third solution involves shifting the burden of recalling passwords from the user to the system by allowing computers to store and retrieve passwords. However, the authors rightly argue that whilst this reduces the user's burden it also results in the password being hidden from the user thus hindering learning of the password. Here the problem concerns encoding the password and not executing the password, which was the problem participants faced in the studies of Wiedenbeck et al., and Weir et al.

We believe that repeated exposure to a password and the inherent context is one means by which an automatic stage of password learning, and depth of password processing, occurs. The problem still existing, after learning and storage have occurred, is successful retrieval among potential alternative password candidates. The complexity of the problem lies in the multiplicity of passwords to learn and then encode in some distinctive manner. The user's task is to *successfully create a series of secure, distinctive passwords that prime distinctive associations with particular services, and efficient and parsimonious mnemonics for remembrance*. This is cognitively demanding even in isolation and it is therefore unsurprising that users adopt one or more coping strategies, such as password reuse or cognitively off-loading, often using paper as an insecure cognitive handle. Moreover, this complex task is not performed in isolation, it is performed against a backdrop of competing user goals concerned with work and leisure tasks and activities that frequently take precedence over security goals. Devising and using passwords is another task the user must perform while interleaving other work and leisure activities. This is potentially one reason why the participants in the Shay et al. study are unsure of the utility of increased security (which they are unable to measure) against the needs of work productivity, or the valence of leisure activities (which they can place a value upon).

In the next section we describe details of an empirical study into actual password use.

2. Empirical study

2.1. Research aims

The overarching research aim is to investigate how people create, use and manage a plethora of passwords in the reality of their daily lives. Specifically, a number of questions that we aim to address in our study have arisen from outlining the prior studies of password authentication. These include:

- What is the instance of actual password use over the study period, and what activities require password authentication during the day?
- How many passwords do participants need to manage, what types of password are created, and how often are they changed?
- Is there a relationship between the estimated strength of the passwords created, and participant perceptions of service/data sensitivity?
- How do people manage multiple passwords, and what are the incidences of failure to authenticate and manage passwords?

To address these questions, an appropriate methodology needs to be applied which produces rich, valid, qualitative and quantitative data about real password use without any investigative manipulation or influence by the researchers. The next section discusses the basis for the methodology chosen.

2.2. Methodology.

The past decade has seen the deployment of a variety of different data collection methods in empirical studies gathering password data. These methods range from experiments where password characteristics have been systematically manipulated and the effects observed and assessed, to more qualitative data from questionnaires and surveys (e.g. Shay et al., 2010). Surveys (e.g. Stanton et al., 2004) have sought the attitudes and reactions of people to passwords in specific contexts, and have attempted to understand what motivates security behaviour (Adams and Sasse, 1999; Novakovic et al., 2009). Other research has focused on investigating single web applications, or on passwords used on an individual computer. For instance, Schneier, (2006) investigated people's passwords for the popular networking site 'MySpace', and Florencio and Herley (2007) gathered *real* data about people's password usage based upon participants' authentications from a particular computer.

In addition to experiments and surveys, diary studies are frequently employed within HCI research. Brown, Sellen and O'Hara, (2000) conducted a diary study related to the development of a new handheld scanner, and reported that over a 7 day period 22 users made 381 diary entries. Czerwinski et al. (2004) undertook a diary study of task switching and interruption involving 11 participants, and Inglesant and Sasse (2010) conducted a diary study of password use over five days.

A comprehensive review of the use, utility, advantages and disadvantages of diary studies is presented in Lazar et al. (2010). They quote Alaszewski (2006) who argued that diaries are more accurate than other research methods, and Hyldegard (2006) who states that diaries fill the gaps in HCI research methods between observation in naturalistic settings or laboratories, and surveys. In many cases it is not feasible to bring users into laboratories, or observe them in their naturalistic settings at different times of the day and night. Neither can the detrimental effects of observation be overlooked. Surveys often fill this gap but as Lazar et al. (2010) indicate, surveys can lead to biased data, do not provide an understanding or explanation of behaviour, and are reliant on participant recall. Using diaries in conjunction with other methods is considered an ‘ideal’ alternative. Diaries can result in increased overall validity given the limited time lapse between an event occurring and it being recorded. Further advantages of diaries as a means of collecting data stem from their excellent facility for recording the existence and quantity of incidents that Lazar et al. term ‘user-defined’. Diaries are also useful for examining situations where users’ behaviour in different locations or settings may be of interest, or an inherent part of the study, and where behaviours are not well understood. One disadvantage of using diaries as a means of collecting data is that the entries constitute self-reports. Whilst the validity of the entries is high due to the limited time lapse between an event occurring and it being recorded, it is hard to externally verify that the entry is accurate. However, it is clear that in gathering real data about the nature of users’ actual passwords, their use and management, there is no alternative but to rely on self-report, since this information is privy to the user alone, and not open to direct observations which would confirm accuracy.

Diaries are typically kept for a short period of time – 1 or 2 weeks to ensure their completion (Rieman, 1993). Carter and Mankoff (2005) distinguish two types of diary – ‘feedback’, where the diary feeds back information to the researcher, and ‘elicitation’, where the data recorded is used for prompting purposes and interviews take place at a later date, with users being encouraged to elaborate on diary entries.

For the above reasons, the complementary data collection methods chosen for the study to be reported in this paper, consists of a diary study coupled with a debrief session/structured interview based upon the diary entries. The diary in this instance will provide feedback to the researchers, but also be used as the basis for the debrief/interview sessions where the users will be asked to reflect on and provide further details and explanations of their behaviour. Consequently, the diary will satisfy the characteristics of a hybrid feedback and elicitation diary, (Carter and Mankoff, 2005). The data of interest and the units of study are the authentications users make with passwords over the period of the diary.

2.3. Participants

The first question to address in conducting a diary study relates to the participants. Lazar et al. (2010) argue that strict representation is not as important for diaries as it is for large-scale survey or experimental design. Nevertheless diaries can produce rich, qualitative and quantitative data, and according to Alaszewski (2006) it is particularly important to involve participants who can furnish valuable and reliable information.

Taking heed of Alaszewski’s point outlined above, we aim to embrace some breadth of the general population, by carrying out the study on a population of willing volunteers, but restricted to two different organisations. The first cohort originates from a commercial business organisation, and the second from an educational establishment.

A total of 22 participants (13 female/9 male) were recruited from HP Laboratories in Bristol, and the computer science and other departments, at the University of Bath. These include:

- 6 female, *administrative support staff* (1 in the age range 20-29, 2 in the age range 30-39, and 3 between 40 and 49 years) and with between 18 and 25 years of experience in using computers (average 21.8 years), 2 were educated in IT;
- 1 female (in the age range 30-40), and 3 male *researchers* (all in the 40-49 age range) each having an IT background and with 24-30 years of experience in using computers;
- 1 female *lecturer* with an educational background in IT – early 30s with 18 years of experience in using computers;
- 1 male *systems engineer* in his 20s, with 16 years of computing experience and a background in IT, and;
- 7 *PhD students* and 3 *MSc students*, (4 female/6 male) from a variety of educational backgrounds (6 in the age range 20-29, 2 between 30-39, and 2 between 40-49 years of age), and with between 14 and 28 (average 17.8) years of computing experience.

As the above details demonstrate, the participants undertake a wide selection of roles and work activities with the organisations. They range in age from early 20s to late 40s, with 14 to 30 years of experience using computers, and with just less than half the participants having an IT background. Consequently, the data collected originates from a small sample from a wide selection of computer users, and is not restricted to a specific user group. Whilst the size of the sample is relatively small for an empirical study, it is in line with other diary studies.

2.4. Procedure

The diary study was conducted during November and December 2008, over a 7-day period, deliberately kept short for three reasons: first, to ensure their organisations were agreeable for participants to take part; second, to encourage self-report by participants, ensuring that motivation remained high, and valid and reliable data resulted; and finally, we wanted to know the magnitude of password use in just 1 week chosen at random, even though we concede that there will be additional services that are used on a monthly or yearly basis. The chosen period of diary completion – 7 days – is consistent with the views of Rieman (1993) noted earlier.

The study was concerned with collecting real data about actual password usage over the study period. We did not ask the users what their actual passwords were but the characteristics of their passwords for different services, used for different activities and tasks, at different times of the day, were deemed of interest. The participants were made aware of the aims of the study and clearly understood the purposes of maintaining the diary. They were also aware that at the end of the period they would be allowed more time and freedom to reflect on their use of passwords and elaborate on any points arising from the diary entries, in debrief and structured interview sessions. The participants were also made aware that the purpose of the debrief and structured interviews was to gather additional information about the properties of passwords.

The participants were requested to take notes (whilst undertaking tasks and activities with services) of different factors related to the usage and management of passwords, including failures. The diary consisted of a form that included detailed questions about the authentication process. These included the time of authentication; which password was used (the actual password was not divulged); location (home, office, mobile or other); activity before and after authentication; whether any hints were provided; the estimated time taken to authenticate; whether the authentication was successful or failed; if it failed, what the nature of the problem was for example, whether the user knew if the failure originated from either mistyping (a typographical error) or misremembering, and what participants did in order to recover from failure (for example, try again or call the helpdesk). Participants were asked to complete the relevant details on the form (multiple copies were provided) each time they authenticated over the 7 days.

After participants had completed their password diary a debrief/in-depth interview was conducted. Here, detailed information concerning the passwords and their use was recorded. The information gathered included the *nature* of the passwords (the passwords themselves were not disclosed), whether they were shared between services, whether they were known by a third person, how they were generated and remembered, how often they were used, and how often they were changed.

At the end of the debrief all participants received a £20 bookstore gift voucher as a reward for taking part in the study. All stored data has been anonymized.

In the following section the combined results of the diary study, debrief and structured interview, are reported.

3. Results

Over the 7-day period, 991 diary entries/password authentications were recorded, suggesting that the task of diary entry was taken seriously, and the motivation to comply was high. Each of the following sections will be structured to address the research questions outlined in section 2.1.

3.1. What is the instance of actual password use over the seven-day period, and what activities require password authentication during the day?

3.1.1. Password use

The 991 authentications reported provide an average of 45.05 authentications per person over the 7-day period, (N=22, min=17, max=173, SD=34.08, median=37.50). The range and standard deviation indicate large individual differences between participants. Table 1 shows the magnitude and pattern of password authentications per day.

Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
165	178	168	177	150	66	87

Table 1: Magnitude and pattern of password authentications per day.

3.1.2. Activities needing authentication

This section is concerned with the different activities needing authentication, their authentication frequency, and also the times of day when services are used. The rationale for including activities and times of day within the diary study, is twofold. First, from a user, work and productivity perspective, the data may be interesting as a basis for future investigations into when participants are interrupted, what activities are being disrupted, and consequently whether users are likely to comply. Second, if security professionals have some knowledge of the pattern of sensitive service use, and also whether those services are instances where insecure passwords are reused from other services (whether sensitive or not), then it may be possible to speculate when security breaches are likely, and when to be extra cautious and vigilant.

In completing the password diary, participants described the activity that needed password authentication. These activities include: working (e.g. beginning a 'work' task) with 422 instances; email (checking, reading or responding to email) with 285 instances; leisure (e.g. browsing internet sites for personal use) with 115 instances; login (logging into a network/computer) with 74 instances; money (e.g. making payments or checking bank accounts) with 40 instances; communicating (authenticating for a communication service such as 'MSN' or 'Skype') with 39 instances; unknown (an activity that was not described) with 12 instances, and meeting (e.g. authenticating for a conference call over the phone) with 4 instances.

The different time zones stipulated were: morning from 6:00 till 11:59am; midday, 12:00-13:59; afternoon, 14:00-17:59; evening, 18:00-22:59; and night 23:00-5:59. The majority of the authentications for the different activities occur in the morning potentially indicating this is when the most disruption to work occurs. An exception to this pattern includes meeting authentications that mainly occur in the afternoon (3 out of 4). Additionally, the highest proportion of money-related activities occurred in the evening (13 out of 40). Table 2 shows the different activities needing password authentication within the different time zones, whilst Table 3 indicates the number of different services needing authentication over the 7-day period.

	Morning	Midday	Afternoon	Evening	Night
Working	229	51	97	44	1
Email	133	24	50	71	7
Leisure	48	13	18	33	3
Login	48	6	8	12	0
Money	11	7	7	13	2
Communicating	23	4	5	5	2
Unknown	9	1	1	1	0
Meeting	1	0	3	0	0

Table 2: Different activities that needed authentication within different time zones.

	Mon	Tue	Wed	Thu	Fri	Sat	Sun
Working	72	89	89	81	69	9	13
Email	47	48	47	45	43	29	26
Leisure	17	14	11	15	11	12	35
Login	11	14	9	17	10	8	6
Money	9	8	3	9	6	4	1
Communicating	6	5	5	8	7	3	5
Unknown	3	1	1	1	4	1	1
Meeting	0	0	3	1	0	0	0

Table 3: Number of services (175 in total) used for different activities over the seven-day period.

Given the extent of password use across the 7-day period, the next question to address relates to the number and type of passwords users created.

3.2. How many passwords do participants need to manage, what types of password are created, and how often are they changed?

3.2.1. Number of passwords to manage

The mean number of passwords to manage is 7.95 (SD=2.46, min=4, max=13, median=8.00) across 175 reported services. Again the range and standard deviation indicate individual differences between the participants.

3.2.2. Password types created

With respect to the types of passwords created, the passwords reported by participants, to authenticate for a particular service, were divided into seven types. These include, (i) a single/common word, or names (e.g. a single name, 'Lucy'); (ii) meaningful phrases (a simple sentence where the words are connected, e.g. 'WhoHasSentMeNewMail'); (iii) variation/abbreviation on a meaningful phrase (the initial letters of a sentence, e.g. the first letters of the sentence 'Who has sent me new mail?' would be 'Whsmnm?'); (iv) a meaningful combination of letters and numbers (e.g. initials and birthdays); (v) a meaningful number pattern (e.g. '0845'); (vi) random characters; or vii) some other pattern (a combination of e.g. words and letters).

As described in Yan et al. (2004) the most secure passwords are random passwords and variations of a meaningful phrase. In contrast, single/common word or name passwords, even if strengthened with numeric suffixes or prefixes are not secure, (Schneier, 2006). In contrast, number pattern passwords or passwords that contain some other meaningful combination can be seen as fairly secure, depending on their pattern or combination.

Figure 1 shows the frequency of the different password types created. Participants created 27.7% of passwords that were a variation on meaningful phrases, e.g. initial letters of meaningful sentences. This was followed by 25.9% of passwords that contained a single/common word or name; random passwords, 12.7%; some other pattern (for example, a combination of words and letters), 12.0%; a meaningful combination of letters and numbers, 9.0%; a meaningful phrase, 7.2%; and number pattern, 5.4%.

In relation to matching password types to activities, money-related activities contained mainly letters and numbers, and the least secure passwords containing letters only included leisure, and a subset of work-related activities, (both 10 out of 35).

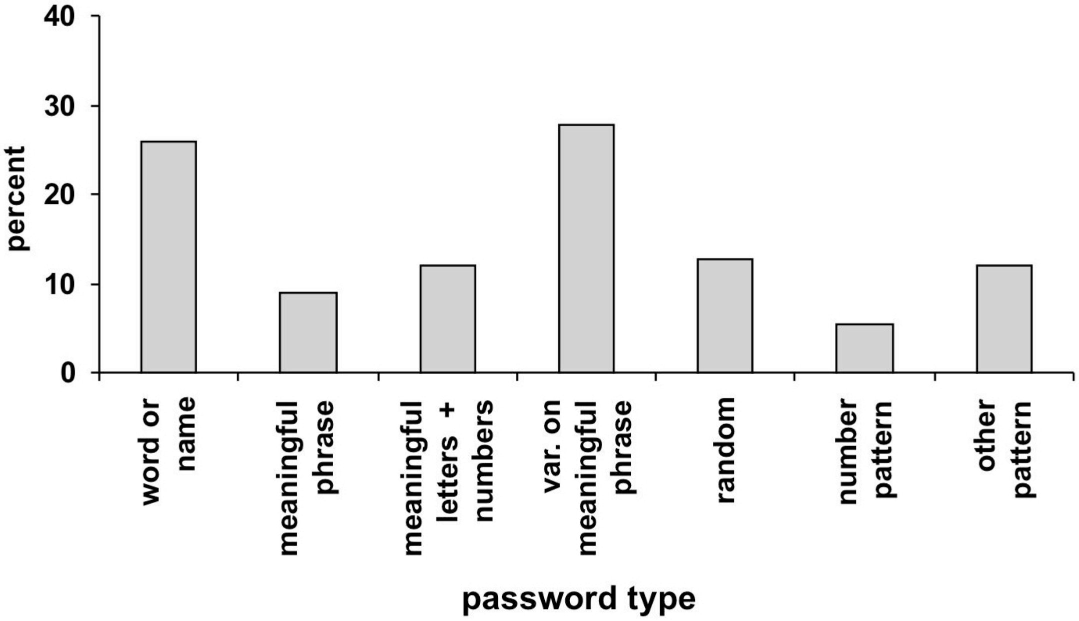

Figure 1: Frequency of the different password types created.

3.2.3. Changing passwords

Participants reported how often they changed passwords. Out of 175 passwords that participants reported for the different services, 139 were never changed; 17 were changed more than once per year; 16 were changed up to once per year; and 3 had been changed once in total.

A chi-square test revealed a significant association between the different password types participants created and whether or not the password was never changed, $X^2(7) = 42.14, p < .001$. The main reason for this significant overall effect lies in the fact that none of the passwords containing random characters or meaningful combinations of letters and numbers were ever changed.

Additionally, there was a significant association between the different password types and whether or not passwords were changed more than once per year, $X^2(7) = 23.73, p < .01$. The main reason for this significant overall effect relates to the fact that the odds of changing a password more than once per year were 7.54 times higher if the password contained a variation on a meaningful phrase than if the password consisted of a word or name.

The next section addresses the question as to whether there was any attempt to match secure passwords to sensitive services.

3.3. Is there a relationship between the estimated strength of the passwords created, and participant perceptions of service/data sensitivity?

Figure 2 shows the frequency of perceived sensitivity of the services. Most services (39.4%) were perceived by the participants to be ‘very sensitive’. This was followed by services perceived as ‘fairly sensitive’, 26.3%; ‘sensitive’, 25.1%; ‘not very sensitive’, 6.3%; and, ‘not at all sensitive’, 2.9%. The passwords were rated as ‘highly secure’ (uncrackable), 17.7%; ‘secure’ (hard to crack), 39.4%; ‘fairly secure’, 27.4%; ‘not very secure’, 10.9%; and ‘insecure’ (easy to guess), 4.6%.

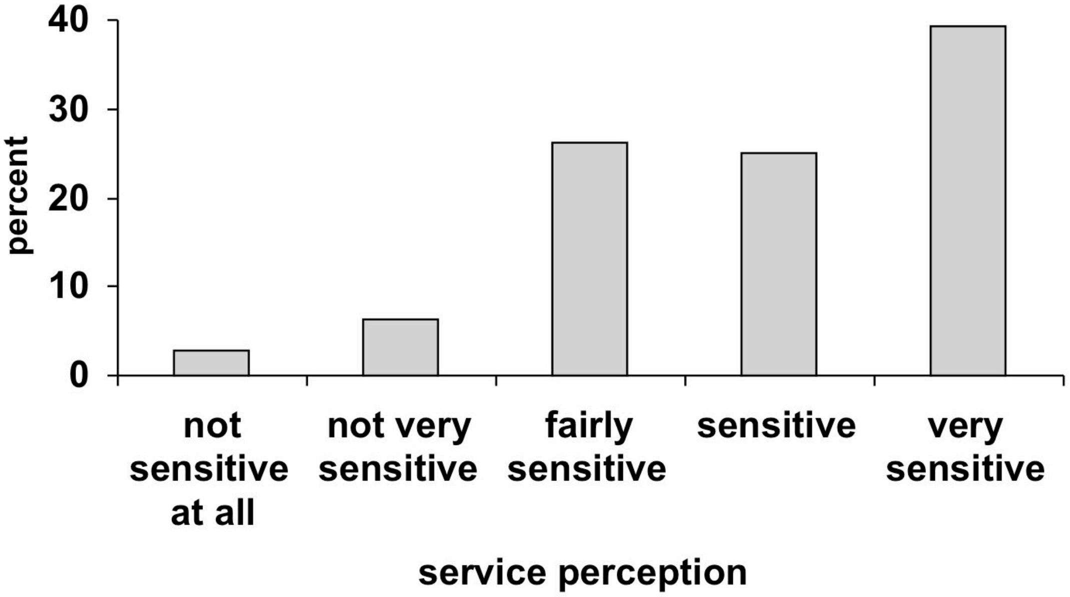

Figure 2: Frequency of perceived sensitivity of the service.

The participants made some attempts to match passwords and services to their perceived strength/security level. For example, participants matched passwords that ‘they considered to be secure’ to services perceived to be very sensitive; fairly secure passwords to fairly sensitive services; and insecure passwords to not very sensitive services. There was a significant positive correlation between participants’ security perception of the service and their estimated password strength (r =0.34, p<0.01). However, this does not mean the passwords were in reality more secure: This is dependent on the participants’ understanding of what constitutes a secure password.

Not only did participants attempt to match secure passwords to sensitive services, they were also concerned with password reuse on very sensitive services. A chi-square test revealed a significant association between perceiving a service as very sensitive and whether or not the password was unique ($X^2(1) = 6.51, p < .05$). The odds of using a unique password were 2.25 times higher for a service that was perceived as very sensitive than if the service was perceived as not very sensitive.

There was no significant association between security perceptions of the service and the reuse of whole or partial passwords.

In the next section we report findings regarding user strategies for successful password management, and where password failures occurred.

3.4. How do people manage multiple passwords, and what are the incidences of failure to authenticate and manage passwords?

The studies by Dhamija and Perrig (2000) and Gaw and Felten (2006) led us to investigate how people manage the need to create and use more passwords. Specific questions which arose are: (i) what is the number of unique passwords and the extent of reuse; (ii) with respect to memorisation, what use is made of hints and how much cognitive off-loading (i.e. writing down and sharing of passwords) to aid memory occurs; and finally, (iii) what is the incidence of failure to authenticate and manage passwords?

3.4.1. Unique passwords and the extent of reuse

Out of the 175 detected services that needed passwords to authenticate, 69 passwords were unique (not reused in a different service and did not share any parts of other passwords), 86 (whole) passwords were reused in other services (up to 4 times), and 20 passwords reused parts of another password. In relation to service sensitivity and reuse, we reported in the previous section that the odds of using a unique password were 2.25 times higher for a service perceived by participants as very sensitive compared with services which were not perceived as very sensitive.

3.4.2. Managing passwords by cognitive off-loading - writing down and sharing with others

Clearly, managing multiple passwords requires effort for creation, encoding, retrieval and execution. None of the participants reported using password management tools. They relied mainly on memory to retrieve their unique and reused passwords. Some participants reported the use of hints to aid memory of passwords (12/991). For the reuse of whole passwords, the hints primarily indicated relationships between passwords, i.e. the password was either the same as their ‘main’ password, or their ‘home’ password. Interestingly, no participant recorded any hints regarding passwords that reused parts of another password.

The main strategies used for managing passwords other than reuse relate to cognitively offloading by either writing passwords down or sharing with a third party. From the debrief, it was found that 11 passwords were written down out of a total of 175. This number may appear quite low but this is set in context given the extent of password reuse. The results indicate that the main factor that affects whether or not a password is written down relates to the password type and characteristics, with other factors such as frequency of use also affecting participants’ behaviour. Our calculations show that the odds of writing a password down were 17.84 times higher if it was unique, and 11.03 times higher if it contained only numbers. Additionally, the odds of writing a password down were 10.20 times higher if it was used occasionally rather than frequently. There was no association between the sensitivity of the service and writing down passwords.

There was some sharing of passwords with third parties, and Bonneau and Preibusch (2010) argue that this is one mechanism by which intimacy and trust between people is displayed. Responses to questions in the debrief about why passwords were shared predominantly stated that the service was also shared. Participants also reported the use of hints relating to shared passwords and these constituted who else knew the shared password. However, further questioning of participants as to who shared their passwords, we deemed to be overly intrusive.

3.4.3. Unsuccessful password authentication, failure to manage passwords and recovery from failure

Over the 7-day period 48 failures to enter a password were reported. The different types of failures reported by participants include: mistyping the password (19); misremembering the password (15); uncertainty about which password to chose (6); forgetting the password altogether (4); some other problem (3); and, being interrupted during the authentication process (1). It could be argued that of these categories, mistyping and interruption (20/48) may be more of an indication of difficulty in executing password authentication. On the other hand, misremembering passwords, interference between passwords evidenced by confusion over which password was appropriate for the service, and forgetting the passwords altogether (25/48) may all be instances of memory failure of one type or another, and relate specifically to difficulty in managing memorisation of multiple passwords.

There were relationships between failures and password ‘uniqueness’. A significant association was found between a unique password and whether or not participants mistyped the password/failed to enter the password correctly, $X^2(1) = 12.97, p < .001$. The odds of mistyping a password were 3.20 times higher if the password was unique than if it was reused. There was also a significant association between the type of password and whether or not participants failed to enter the password correctly $X^2(7) = 18.72, p < .01$. The odds of mistyping a password was 8.42 times higher if it contained a number pattern than if it contained a meaningful phrase. There was also a significant association between changing a password and whether or not an authentication failure occurred, $X^2(1) = 5.98, p < .05$. The odds of failing to enter the password correctly were 2.05 times higher if the password was changed than if it was never changed.

In addition to the incidence of mistyping unique passwords, a chi-square test revealed a significant association between a unique password and whether or not the password was misremembered, $X^2(1) = 4.26, p < .05$. The odds of misremembering a password were 2.86 times higher if it was a unique password than if it was not unique. There was also a significant association between the type of password and whether or not participants misremembered a password, $X^2(7) = 29.71, p < .001$. The odds of misremembering a password were 9.52 times higher if the password contained a number pattern than if it was a variation on a meaningful phrase.

Participants reported 24 recoveries from failures to authenticate. These reported recoveries include: trying to re-enter the password (14); requesting a new password (5); giving up or trying to authenticate later (3); and, locating and using a hint to enable authentication (2). A chi-square test revealed a significant association between the different password types and whether or not participants tried to re-enter a password, $X^2(7) = 61.33, p < .001$. The odds of trying to re-enter a password were 46.08 times higher if it contained a number pattern than if it was a variation on a meaningful phrase. There was also a significant association between unique passwords and whether or not a new password was requested as recovery from failure, $X^2(1) = 4.52, p < .05$. No reused password was recovered from failure due to requesting a new password.

In summary, it is clear that the need to manage multiple passwords has led to different strategies to aid password use and remembrance. There were 991 authentications to 175 services, however, only 69 of these were unique (not reused in a different service and did not share any parts of other passwords), and 11 in total, including passwords that were reused, were written down. The likelihood of writing down was highest if the password was unique, followed by if it contained numbers only and finally, if it was used infrequently. Further insecure security behaviour occurs in the sharing of passwords. There were 25 failures of authentication that relate to difficulties in managing the memorisation of passwords. Whether or not the password is unique, and the type of password affects both mistyping and successful or unsuccessful remembrance.

In the next section, we discuss the study results in the context of the previous literature and with particular regard to managing multiple passwords.

4. Discussion

The primary research aim was to investigate the *reality* of password creation, use and management in the context of an increasing need for password authentication. The research referred to earlier in the paper by Bonneau and Preibusch, suggests that passwords are here to stay for sometime in the future given their dual roles of engendering trust, and as a means of authentication. In this regard, it may be beneficial to undertake further research into how password use can be made more secure, and password policies be devised that could maximize compliance by making it easier for users to comply. The review of the literature, taken together with the results of the empirical study reported in this paper show that currently insecure behaviour persists, and there is a lack of appropriate policies related to creating and managing multiple passwords. A significant problem underpinning insecure behaviour relates to the distinction between the magnitude of attention paid to individual password creation and use, as opposed to the paucity of attention paid to the collection of passwords a user possesses.

A further research aim is to use the study results to improve user behaviour. The study data indicate three reasons for poor security from a user perspective; a lack of knowledge about security, erroneous knowledge, and poor strategies for coping with password overload. To improve security, two goals need to be accomplished. First we need to identify what knowledge is lacking or erroneous, and provide appropriate education. Second, we need to change users' insecure strategies for coping with multiple passwords. In order to achieve this second goal we need to replace existing user strategies with effective secure strategies that capitalize on users' existing real-world knowledge, cognition and memory processes and mechanisms, making it easier to comply. These strategies in future should provide the basis for developing policies and support structures for managing password collections.

The high level guidelines provided later in this discussion have been specifically devised in an attempt to replace users' poor strategies for coping with password overload – reuse, writing down and sharing as evidenced in the study results – with effective and secure strategies. The substance of these strategies needs to be dictated by an understanding of memory capabilities and limitations.

One incidence of insecure behaviour to be discussed relates to the study a decade ago by Zviran and Haga (1999). They found that users changed their passwords frequently to sensitive services but there was no association between the type/security of a user-generated password, and the sensitivity of the service. We postulated that this could be due to lack of perception of i) the sensitivity of the service, ii) the need to match security of the password to service sensitivity, or iii) a lack of knowledge about what constitutes a secure password for these services. The results of the study reported here, indicate that participants did perceive the respective sensitivity of the services they were using. Of these many, but not all, saw the necessity of matching password security to service sensitivity: some participants used both stronger and unique, as opposed to reused, passwords for sensitive services. However, problems with user behaviour still exist in that the unique passwords created were very seldom changed. Moreover, there were incidences of a complete lack of knowledge about what constituted a secure or insecure password. The argument by Adams and Sasse (1999) appears to still apply - users were not averse to adhering to security policy but they did not know how. This indicates the rather obvious point, reiterated many times before that users need better education about, and technical support for, (cf. Gaw and Felten, 2006) secure password creation and use. In this case the education would be designed to overcome a lack of knowledge rather than better strategic use of existing knowledge.

Further examples of insecure behaviour relate to password reuse in an attempt to satisfy the need for multiple passwords. The results of the study reported here on actual password use, corroborate the findings of Dhamija and Perrig (2000), Brown et al. (2004), Florencio and Herley (2007), Gaw and Felten (2006) and Shay et al. (2010). Users frequently reused passwords to satisfy the demand for more passwords, and the proportion of unique passwords is generally quite low in relation to the number of services used. Also there were instances of password reuse to sensitive services by participants and anecdotal evidence from conversations with some of the participants indicated that they do not generally reflect upon their approach to security. The substance of related comments from two participants was that, 'it had just occurred to them that reusing passwords across banking, social networking sites, and work environments was not a good idea'. It appeared that it may have been the first time they had taken the opportunity to reflect on their behaviour and this was as a result of being questioned, and not of their own volition.

In Gaw and Felten's study the median use of 'themed' passwords for reuse was zero. This is due to their participants reusing whole passwords. In the study reported here, 86 passwords were reused up to four times but also 20 passwords reused parts of other passwords. These results are in contrast to those reported in Gaw and Felten. This finding could be linked to some reported incidences of authentication failure where participants were confused about which one of a collection of related passwords to chose.

Participants in our study relied on memory for retrieving passwords. They also shared, wrote down and reused passwords, with a small number of hints for memory retrieval, all forming part of their coping strategies for dealing with multiple passwords. In the past some security policies (Adams and Sasse, 1999) have lessened rather than improved security. The example they provide, as outlined in the introduction, concerns the fact that frequent requirements to change passwords leads to passwords being written down. The crux of the problem is that users will devise strategies, including cognitive offloading, to avoid problems of not being able to successfully authenticate. Unsuccessful attempts at authentication are costly to users and their productivity, and password recovery is also a cost to service providers, (see the economic and strategic discussion in Bonneau and Preibusch (2010) regarding the provider efforts and steps that can be taken to recover passwords). According to the Gartner Group (http://www.mandylionlabs.com/PRCCalc/PRCCalc.htm) between 20% and 50% of all help desk calls are for password resets, and Forrester Research indicate that the average help desk labour cost for a single password reset is about $70 in the USA.

The study data revealed 48 failures to authenticate, reported by participants. There were both different types of failure and different failure rates depending upon the uniqueness and type of password. For example, the odds of *mistyping* a password were 3.2 times higher if the password was unique, 8.42 times higher if they contained a number pattern than meaningful phrases, and 2.05 higher if they had been changed. Additionally, the odds of *misremembering* a password were 2.86 times higher if it was unique and 9.52 times higher if it was a number pattern. In relation to overcoming authentication failures, the odds of re-entering a password were 46.08 times higher if it was a number pattern, unique passwords were primarily those where new passwords/resets were requested, and interestingly no failure involving a reused password was ever recovered by requesting a new password/reset.

The cognitive processes and user behaviour that underlie failures are complex. In the absence of hypotheses and systematic study, it is not possible to do more than speculate concerning the reasons for failure. For instance, it is not clear from the data, whether the activity of mistyping is independent of memory retrieval problems even though participants categorise incidences of mistyping and misremembering differently. Unique passwords for example which take more effort to retrieve, have a higher chance of being mistyped. Are they retrieved wrongly because they are less frequently used, or does less frequent use lead to inadequate practice of the motor skill of typing the specific password, i.e. an execution error? Unique passwords also have a higher chance of being reported as misremembering failures and the odds of them being written down are over 17 times higher than reused passwords. This clearly shows that unique passwords are particularly problematic for users, both to remember correctly and to execute when remembered. Users may well be aware of this and adopt the strategy of writing them down as a precaution against authentication failure. If they are not written down or they cannot be located, then participants request another replacement unique password. By contrast no failures of reused passwords were ever recovered by requesting a new password.

The number of participants in our study was relatively low. This means making strong claims and generalisations from our data is problematic. We are not able to extrapolate from the representatives of the organisations to other individuals in those organisations, since they were a small number of volunteers, and also we cannot extrapolate to other organisations. It is important that the tentative predictions we now make from our data are understood within the limited scope of the study, and that these predictions may not hold for larger and different populations, and further research with these populations is needed.

One of a number of tentative predictions we can make from our study data relates to the effects of following security policies advocating the use of unique passwords, whilst not providing some support structure for their retrieval in the light of multiple password use. *Users who must create unique passwords will suffer higher incidences of mistyping, higher incidences of misremembering and will adopt two coping strategies – they will request new passwords thus costing password providers, and they will write the password down thus weakening security.*

It has long been known that different password characteristics affect their memorability and our data in particular show that numbers, when used in isolation, cause users problems. However they are far less problematic when used in a meaningful combination with letters or words. Both retrieving and execution of number pattern passwords is poor. The fact that the odds of password re-entry are over 46 times higher for number pattern passwords demonstrates how disruptive these are for the user, their work productivity and daily life. Another tentative recommendation from the study findings suggests that *using numbers in isolation is not to be recommended by security policies and supporting users in creating passwords that combine meaningful numbers with other secure meaningful material would be beneficial for memory retrieval, for execution and to avoid frequent occurrences of password re-entry*. Passwords containing only numbers and unique passwords are not entirely independent of one another. Therefore, it is possible that if numbers with specific meanings and other secure meaningful material are used for unique passwords (as opposed to numbers only unique passwords), that some of the problems with unique passwords may also be addressed.

It is of fundamental importance that human capabilities and strategies for dealing with overload are understood and taken into account in devising future security policies, guidelines and education. The reason some policies are not successful is that they are not founded on a sufficient theoretical understanding of the users' tasks, and the different psychological processes inherent in creating, encoding, retrieving and executing passwords. Consequently, it is not always possible to predict user behaviour in adhering to, or working around security policies, nor does it take the stance of maximising the users' chances of following security policies. In the next section we consider how to overcome problems with password overload by devising guidelines and mechanisms for managing multiple passwords. Our ultimate goal in the future is to improve the management of unique (but possibly related) passwords, reduce password reuse and the specific aspects of cognitive offloading that lessen security.

4.1. Guidelines to support password creation and management

The motivation for developing guidelines for multiple password creation and management is that currently policies exist for individual as opposed to multiple passwords, and the near absence of policies for creating and remembering collections of passwords is clearly a failing.

The bases for the guidelines relate to both the problems with insecure password behaviour outlined in prior literature, and also in our study results. For instance, with respect to the prior literature, Shay et al. demonstrate that despite little effort to comply and a perception of greater security, users do not know whether it is worth the extra mental effort of complying with a policy requiring password change. Similarly, in the study reported here, there is clear evidence of users limiting the cognitive effort in devising and memorizing large numbers of passwords by reusing, sharing and writing them down. Given that the main problems users face is with memorisation, there needs to be a means by which memory of passwords, is improved. Specifically, in order to improve security, poor and insecure strategies for creating, encoding, and retrieving collections of passwords must be replaced with effective, secure strategies.

Earlier in this paper we argued that the users' 'job' with respect to passwords is to 'successfully create a series of secure, distinctive passwords that prime distinctive associations with particular services, and efficient and parsimonious mnemonics for remembrance', against a backdrop of competing work/leisure goals. Therefore, we need to develop appropriate policies, guidelines, and education to support users' psychological processes for managing multiple passwords. This is instead of inflicting demands and goal competition onto users, which means security goals have a poor chance of being met when set against work goals with tight deadlines with clear and immediate consequences, when they are not met.

How do users cope with overload or the need to make sense of infinite demands on their memory and attention? The answer is that they make consummate use of memory processes, such as recall and recognition. They also problem solve, apply guesswork and trial and error, and they infer in making sense of these activities from past experience. Humans categorise in order to limit and handle vast arrays of stimuli. This supports inference and generalisation. The process of categorisation aids recall and provides structure and relationships between stimuli that underpins informed guesses. The stimuli of interest here are numbers of passwords, and users' present insecure coping strategies involve limiting instances of passwords to be categorised, by reusing them.

We postulate that the process of categorisation could help in creating, forging relationships between and recalling passwords. Moreover, it would be particularly helpful in recalling a collection of passwords. Therefore, the purpose of the guidelines is to support users in creating and managing password collections (as opposed to individual passwords about which sufficient literature exists).

In the past security policies have recommended the provision of hints that supposedly prime memory associations to aid recall of passwords. However, the format of these hints need to be tightly specified, clearly associated with the password(s) in question, and in their implementation must not be the passwords themselves. From our study data, it is clear that the hints written down by participants involved either complete disclosure of the passwords, or names of people who share their passwords.

We think we can add to high level guidelines related to memory, and the use of hints and primes. For instance, security policies, guidelines and education should support the user to:

(i) harness their ability to categorize to support the encoding and retrieval of multiple passwords;
(ii) provide guidelines to support users in mapping and representing categories of passwords to enable creation, encoding and retrieval of appropriate passwords;
(iii) create cognitive handles: allow users to/or demonstrate to users how to, write down *encrypted* categories/hints related to their passwords;
(iv) allow users to make guesses, use trial and error, and problem solving, in retrieving candidate passwords.

We are aware that some of these recommendations *in implementation* could be very contentious. For instance, one means of complying with the guideline to allow users to guess, use trial and error, and problem solve, is to allow users more password input attempts. This would allow users to try different candidate (those with a high likelihood of matching) passwords when they are unsure which one is associated with a service. However, Florencio and Herley specifically point out that lockouts are an effective way of curtailing password hacking. Moreover, we do not possess enough knowledge of user strategies to know when they guess as opposed to reuse, how many guesses they might make before they give up, or whether the benefit to users' cognitive load is paid for by disproportionately more hacking attempts that are successful.

However, we suspect that security could be improved as a result of supporting users in creating encryptions and writing them down as opposed to actual passwords. It could potentially be better also if designs could be devised to lead users through a process of discovery that recreates the context, by means of 'transfer appropriate processing', (Morris et al., 1977) in which the passwords were created and encoded.

The recommendations above could attract the criticism of being too general and vague. With this in mind, we now try to apply the recommendations/guidelines in a specific manner. We also attempt to overcome some of the problems, reported in our data, with unique passwords. In addition, the diary study indicates that people reuse parts of, and whole, passwords, and there is a need to reduce (even if we cannot eliminate) password reuse. Passwords are typically considered in isolation but from the perspective of users there is actually a collection of them. Rather than remembering a single password, users have to remember a number of passwords, and this requires different methods. Consequently, methods are needed to support memory for a collection of passwords instead of just one - hence the advantage of considering categorisation as a means to aid recall of passwords.

However, a fundamental problem exists in utilising a single category of passwords given the need to distinguish between individual passwords for particular authentication events, to curtail interference, and retrieve the appropriate passwords, (category members). One generic solution is to have a number of categories, and favourite instances, (along perhaps with related features and events to make up meaningful phrases or their abbreviations) so that they are the first to be recalled. For some (natural, artificial, event or adhoc) categories specific instances naturally come to mind, these instances being core to the individual's category (based on prior experiences), and these are the quickest and easiest to recall. Different categories would be associated with particular services. The problem can then become remembering the category/service associations. This clearly needs further investigation in a future research programme, but it does not seem to be any more of a challenge than currently remembering which password (possibly made up of parts of other passwords) to reuse for specific services.

When categories are formed they are structured taking account of shared, overlapping and common features and attributes. The core of the categorical structure lends itself to representing the instances with the most shared, common or overlapping features. The category core and outwards to the periphery have frequently been represented as multidimensional spaces, or visually as a series of ever widening circles from the core to the periphery. Coincidentally, within the social sciences, a method known as socio-mapping, (Bahbouh, 1993) has been used to represent relationships between data, e.g. social network data, as a series of concentric circles. More recently, this method has been used by individuals in representing their feelings of 'belongingness', with other people/communities (real or virtual) – families, friends, friends of friends, acquaintances, etc. These socio-maps consist of concentric circles of 'belonging' with close family members represented as nodes at the centre spiralling outwards to the periphery where acquaintances may be represented. A potential solution to associating sensitivity with secure passwords is to marry the process of categorisation (in this case, 'people') with the representational formalism of socio-maps with the closest relationships at the core. The idea is to use the closeness of the relationships as a mnemonic for service sensitivity – following this principle the closest family members would be the first to be protected, and therefore require the strongest password. For instance, close family members could form part of a meaningful phrase which includes them, related events and details about them, and the nearer they are to the centre the more likely they are to be used for a highly sensitive service. Clearly, 'people' is only one of many exemplar categories that could be employed in this way – the process is a general rather than specific solution.

The idea of marrying categorisation processes, and socio-maps as a means to represent categories of passwords, with the most sensitive associations at the centre, as an aid in multiple password management, needs thorough investigation. To be successful, it must lead to password collections that are easier to remember and distinguish, and lessen instances of insecure user behaviour. Consequently, a future research aim is to experimentally manipulate three factors, (i) different types of categories, (ii) different instructions to participants, and (iii) different representations of passwords and password collections. The effects of these manipulations on improvements to security, password memorability, and accurate retrieval of passwords for specific instances of authentication, would need to be assessed. The issue would be whether using categorisation could support accessing the correct password from the multiple passwords possessed, and whether or not it increased the efficiency of unique passwords whilst also reducing the incidence of reuse and lax cognitive offloading. Moreover, a series of further empirical studies investigating the utility of providing different supports for user-defined encryption and deciphering encryptions, is needed. These encryptions could be number-based, textual, graphical, or even employ maps (including socio-maps) that were involved in the password collection created. Depending on the results of these studies, a long term aim is to develop clear unequivocal guidelines, and technological support for enabling users to take the responsibility of successfully managing multiple passwords.

5. Conclusion

The study reported in this paper has provided provisional but rich data about the reality of password use, and the management of a plethora of passwords, in the context of people’s everyday lives. The results demonstrate that lax security behaviour involving password reuse, writing down and sharing passwords still exists, along with a lack of or erroneous knowledge about what constitutes a secure or insecure password. Tentative predictions have been made from the study data about user behaviour in the light of security policies that advocate the use of unique or number only passwords.

Our main argument has been that it is fundamentally important that human capabilities and strategies for dealing with overload are understood and taken into account in devising future security policies, guidelines and education. We further argue that the reason some policies are not successful is that they are not founded on a sufficient theoretical understanding of the users’ tasks, and the different psychological processes inherent in creating, encoding, retrieving and executing passwords. As a consequence we have suggested the utilisation of categorisation as a mechanism for aiding password creation, encoding, retrieval and execution.

Finally, we outlined a future research programme of experimental studies to test if categorisation aids in creating and remembering individual and collections of passwords, and also if the concept of ‘belongingness’ coupled with socio-mapping has a role to play in creating and using unique, very secure passwords to highly sensitive services.

Acknowledgements

We are especially grateful to the 22 study participants who willingly gave their time, and to Hewlett Packard and the University of Bath as employers who gave consent for their staff and students to take part in this study. The support of the Technology Strategy Board (TSB) is also gratefully acknowledged. Geoff Duggan commented on the paper and we thank him for the thought, time and effort he put in. We also thank Julie Turner Cobb for discussions about belonging and socio-mapping. The authors also wish to acknowledge and thank the Trust Economics team and David Pym in particular as project lead (http://www.trust-economics.org). Special thanks to Philip Inglesant for the initial design of forms and debrief that we modified for our study, and Angela Sasse for discussions on passwords within the Trust Economics project. Finally, thanks to three anonymous reviewers whose suggestions have significantly improved the quality of this paper.

References

Adams, A., Sasse, M.A., 1999. Users are not the enemy. Communications of the ACM 42(12), 40-46.

Alaszewski, A., 2006. Using Diaries for Social Research. Sage Publications, London.

Avarne, S., 1988. How to find out a password. Data Processing & Communication Security 12(2), 16-47.

Bahbouh, R., 1993. Sociomap. <http://www.sociomap.com/>.

BBC News, 2004. Passwords Revealed by Sweet Deal. BBC News, UK Edition. <http://news.bbc.co.uk/1/hi/technology/3639679.stm>.

BBC News, 2010. New Era for Internet Security Amid Increased Attacks. <http://news.bbc.co.uk/1/hi/technology/8544413.stm>.

Bonneau, J., Preibusch, S., 2010. The password thicket: Technical and Markey failures in human authentication on the web. In: Proceedings of WEIS 2010, Boston, MA, USA.

Brown, B., Sellen, A., O'Hara, K., 2000. A diary study of information capture in working life. In: Proceedings of the 2000 ACM Conference on Human Factors in Computing Systems, 438-445.

Brown, A.S., Bracken, E., Zoccoli, S., Douglas, K., 2004. Generating and remembering passwords. Applied Cognitive Psychology 18 (6), 641-651.

Bubas, G., Orehovacki, T., Konecki, M., 2008. Factors and predictors of online security and privacy behavior. Journal of Information and Organizational Sciences 32(2).

Carter, S. and Mankoff, J., 2005. When participants do the capturing: The role of media in diary studies. In: Proceedings of the 2005 ACM Conference on Human Factors in Computing Systems, 899-908.

Casassa Mont, M., Thyne, R., 2008. Privacy policy enforcement in enterprises with identity management solutions. Journal of Computer Security 16(2), 133-163.

Czerwinski, M., Horvitz, E., Wilhite, S., 2004. A Diary Study of Task Switching and Interruptions. In: Proceedings of CHI 2004, Human Factors in Computing Systems. Vienna.

Dhamija, R., Perrig, A., 2000. Déjà vu: A user study. Using Images for Authentication. In: Proceedings of the 9th USENIX Security Symposium.

Florencio, D.A.F., Herley, C., 2007. A large-scale study of web password habits. In: Proceedings of the 16th International Conference on World Wide Web.

Gaw, S., Felten, E.W., 2006. Password management strategies for online accounts. In: Proceedings of Soups '06, 44-55.

Goguen, J.A., Meseguer, J., 1982. Security policies and security models. In: IEEE Symposium on Security and Privacy.

Harreld, H., 2001. Security: An uneasy alliance. Infoworld 23(13), 42-44.

Highland, J.H., 1991. How to prevent the use of weak passwords. EDPACS 18(9), 7-12.

Hyldegard, J., 2006. Using diaries in group-based information behaviour research: A methodological study. In: Proceedings of the Information Interaction in Context, 153-161.

Inglesant, P. & Sasse, A., 2010. The true cost of unusable password policies: Password use in the wild. In: Proceedings of the 2010 ACM Conference on Human Factors in Computing Systems.

Ives, B., Walsh, K.R., Schneider, H., 2004. The domino effect of password reuse. Communications of the ACM 47(4), 75-78.

Keith, M., Shao, B., Steinbart, P.J., 2007. The usability of passphrases for authentication: an empirical field study. International Journal of Human-Computer Studies 65(1), 17-28.

Lazar, J., Feng, J., and Hochheiser, H., 2010. Research Methods in Human-Computer Interaction. John Wiley and Sons, Chichester, UK.

Morris, R., Thompson, K., 1997. Password security: a case history. Communications of the ACM 22(11), 594-597.

Morris, C.D., Bransford, J.D. & Franks, J.J., 1977. Levels of processing versus transfer appropriate processing. Journal of Verbal Learning and Verbal Behaviour, 16, 519-533.

Novakovic, L., McGill, T., Dixon, M., 2009. Understanding user behaviour towards passwords through acceptance and use modelling. International Journal of Information Security and Privacy 3(1).

Rieman, J., 1993. The diary study: A work-place-oriented research tool to guide laboratory efforts. In: Proceedings of the 1993 ACM Conference on Human Factors in Computing Systems, 321-326.

Schneier, B., 2006. MySpace Passwords Aren't So Dumb. Essay on Wired News. <http://www.schneier.com/essay-144.html>.

Shay, R., Komanduri, S., Kelley, P.G., Bauer, L., Leon, P.G., Christin ,N., Mazurek, M.L., Cranor, L.F., 2010. Encountering stronger password requirements: User attitudes and behaviors. In: Proceedings of SOUPS ‘10, July 14-16 Redmond, WA, USA.

Stanton, M.J., Stam, K. R., Mastrangelo, P., Jolton, J., 2004. Analysis of end user security behaviors. Computers and Security, Elsevier.

Weir, C.S., Douglas, G., Richardson, T., Jack, M., 2010. Usable security: user preferences for authentification methods in e-banking and the effects of experience. Interacting with Computers, 22, 153-164.

Wiedenbeck, S., Waters, J., Birget, J.C., Brodskiy, A., and Memon, N., 2005. PassPoints: design and longitudinal evaluation of a graphical password system. International Journal of Human-Computer Studies, Special Issue on HCI Research on Privacy and Security 63, 102-127.

Yan, J., Blackwell, A., Anderson, R., Grant, A., 2004. Password memorability and security: empirical results. Security & Privacy Magazine, IEEE 2(5), 25-31.

Zviran, M., Haga, W.J., 1990. Cognitive passwords: the key for easy access control. Computer and Security 9(8), 723-736.

Zviran, M., Haga, W.J., 1999. Password Security: an empirical study. Journal of Management Information Systems 15(4), 161-185.