Assessing the User Experience of Password Reset Policies in a University

Simon Parkin(✉), Samy Driss, Kat Krol, and M. Angela Sasse

Department of Computer Science, University College London, London, UK
{s.parkin,samy.driss.14,kat.krol.10,a.sasse}@ucl.ac.uk

Abstract. Organisations often provide helpdesk services to users, to resolve any problems that they may have in managing passwords for their provisioned accounts. Helpdesk logs record password change events and support requests, but overlook the impact of compliance upon end-user productivity. System managers are not incentivised to investigate these impacts, so productivity costs remain with the end-user. We investigate how helpdesk log data can be analysed and augmented to expose the user's personal costs. Here we describe exploratory analysis of a university's helpdesk log data, spanning 30 months and 500,000 system events for approximately 10,000 staff and 20,000-plus students. The scale of end-user costs was identified in log data, where follow-on exploratory interviews and NASA-RTLX assessments with 20 students exposed issues which log data did not adequately represent. The majority of users reset passwords before expiration. Log analysis indicated that the online self-service system was vastly preferred to the helpdesk, but that there was a 4:1 ratio of failed to successful attempts to recover account access. Log data did not capture the effort in managing passwords, where interviews exposed points of frustration. Participants saw the need for security but voiced a lack of understanding of the numerous restrictions on passwords. Frustrations led to adoption of diverse coping strategies, for example deliberately waiting to reset a password after reaching the post-expiry grace period. We propose ways to improve support, including real-time communication of reasons for failed password creation attempts, and measurement of timing for both successful and failed login attempts.

1 Introduction

An organisation may have a password policy, which dictates rules about the character composition of passwords for provisioned accounts and how these passwords should be managed (e.g., how often they should be changed). Password policies in organisations can be unusable – and consequently insecure – if the end-user experience is not considered [9]. Policies may dictate that users regularly change their passwords, complicating attempts to remember a password. It is already common to forget passwords [32], especially when managing many at once [5].

To reduce the impact of forgetting passwords, password recovery mechanisms are often provided. These typically take two forms [20]: (1) self-service,

F. Stajano et al. (Eds.): PASSWORDS 2015, LNCS 9551, pp. 21–38, 2016.
DOI: 10.1007/978-3-319-29938-9_2

authenticating through an alternate factor such as security questions, and/or; (2) helpdesk, where a representative will reset a password on behalf of a user.

IT security policies must accommodate regulatory and economic considerations [15]. Policy decisions are typically based on intuition and security expertise, and consider business impact ahead of any effects upon end-users [18]. This does not necessarily lead to poor choices, but is subjective and enacted without evidence of productivity impacts for end-users who have jobs to do.

For the individual, a password reset can appear to lack a personal benefit for the amount of effort required [19]. This can encourage negative attitudes towards computer security, leading users to subvert security mechanisms [4]. Assertions about the effectiveness of password reset policies are then weakened if end-users feel pressured to respond to the burden in unanticipated ways. It may not be possible for those responsible for maintaining security to observe these unanticipated responses (or otherwise seem obvious for them to monitor behaviours if they have no evidence that policy is not being enacted) [2].

Here we explore what the data collected for typical password support mechanisms can and cannot indicate about the end-user experience, and where they may be augmented to improve both usability and security. A university provided access to helpdesk log data for password reset activity, spanning 30 months, for all users of an authentication system for email and other centrally-managed IT services. This data was analysed using Exploratory Data Analysis (EDA) techniques to find out what it shows of the end-user experience for (un)successful reset or recovery of passwords over time. Where analysis did – or did not – describe user experiences, this informed questions posed during exploratory interviews with 20 students, structured to explore attitudes towards specific aspects of password policy (such as responding to password expiry notifications and adhering to password composition rules). Interview analysis relates the end-user experience to numbers in the helpdesk logs, with an aim to identify candidate extensions to both the logging process and the password support mechanism itself. A contribution of this paper is the application of holistic thinking to the way security infrastructure supports and monitors users and their security behaviours.

The remainder of the paper is organised as follows: In Sect. 2, we give an overview of related research on password policies in organisations and end-user behaviours around authentication. Section 3 summarises the methodological approaches in our research. Section 4 presents our findings from the analysis of helpdesk log data while Sect. 5 focuses on the qualitative results from user interviews. We discuss our findings in Sect. 6 and build on them to provide recommendations for practitioners. Finally, Sect. 7 summarises our findings and outlines avenues for future research.

2 Related Work

A number of works have analysed organisation log data and engaged with system users to develop understanding of behaviours around password policies and related authentication mechanisms. Adams and Sasse [4] stress that IT security managers must consider that (i) users can decide whether to comply with security policy, and (ii) the decision to comply is shaped by goals, perspectives and attitudes, and the norms that govern behaviours. Security is not at the forefront of users' minds [31], and should not be a barrier to a user's primary task or it will create frustration and promote negative attitudes towards security [9]. Tightening security can not only render systems less accessible and make users less productive – it might also undermine security as users will employ workarounds to carry out their tasks [23]. The same can be said of demands that occur repeatedly or in sum over time [9].

The usability and productivity impacts of security mechanisms are rarely considered when devising policies [9]. The user burden of security can be measured as the impacts they experience, for instance delays to task completion, restrictions to their ability to complete their work, and in the case of passwords the memorability of new information becomes critical [7]. A user may begrudgingly commit effort to compliance with security, if only to avoid being held accountable for a security compromise [9]. Shay et al. [3] look specifically at user behaviours and attitudes towards passwords and password change at a university, providing recommendations for how to manage user responses to policy expectations. Strategies for managing passwords were also identified, where a number of survey respondents remarked that they wrote down their passwords, or otherwise shared them with others or followed a repeatable set of rules for creating a new password such as reusing their previous password or a password from another account. Here we consider the impacts upon users which may be hidden from the view of system managers, even where user behaviour as observed in system logs appears to be compliant with the password policy.

It has been suggested that the design of security systems be based on task and work-flow analysis [25], and that security restrictions be routinely tested with various users to minimise interference and enhance (rather than hinder) productivity [23]. Elsewhere productivity impacts for groups of users have been considered alongside business support costs and recognised efforts to reduce security breaches, in a decision-support paradigm [8] and complementary proposal for tool support [22], with a focus on password policies. The paradigm was well-received by security managers, especially where it was able to relate to decisions and data that managers were familiar with.

Organisations may employ a helpdesk team to help users who need to replace forgotten passwords. Alternatively, users may manage passwords personally through self-service portals, or by using a registered contact point (e.g., having information sent to a registered email account) [24]. The number and complexity of password policies plays a role in increasing the frequency of helpdesk visits for password resets [4], where it is estimated that roughly 30 % of help requests relate to password resets [29], the majority as a result of passwords being forgotten. Password reset requests are most common after holidays or after long periods of inactivity [12]. Additionally, helpdesk calls may result from users being locked out, as it is common for organisations to lock user accounts after three failed login attempts (where increasing this limit can reduce the number of reset requests [12]). In some organisations, the cost of maintaining helpdesks leads security staff to consider the writing down of passwords as a minor infraction [27]. Many organisations then seek to reduce the number of calls to the helpdesk [13]. Self-service solutions however require investment and support to be appropriately integrated into designated systems [28]. When selecting or changing a password, users should be provided with instructions for constructing and memorising strong passwords [26]. Without such information, users make up their own rules to devise more memorable passwords [4]. Training users to create secure and memorable passwords has been identified as a more plausible solution to easing the burden on the helpdesk [16]. Here we also consider where communication of policy can be improved without necessarily changing the rules of the password policy itself.

3 Methodology

A mixed methods approach combined both quantitative and qualitative data. Exploratory Data Analysis (EDA) [30] was performed on the anonymised helpdesk log data. Second, semi-structured interviews were conducted to understand end-user perceptions of system use and related policies. Interviewees completed NASA-RTLX workload assessments [17] for regular tasks, thereby quantifying perceived workload. Interviews and workload responses were then analysed relative to the findings of the log analysis. The study received ethics approval from the Research Ethics Committee at UCL. All quantitative data was anonymised. Participants in interviews were assigned participants numbers and no personally-identifying information was stored during data collection.

3.1 Systems Under Analysis

Analysis was carried out with a university support team's password reset log data. The data comprises monthly reports of the frequency distribution for various password reset activities (e.g., resetting an expired password) from September 2012 to April 2015, for approximately 10,000 staff and 20,000-plus students. The data contains logs from two systems: Personal Password Management and Personal Password Recovery, where each system has both an online and in-person helpdesk channel. The main dataset consisted of monthly aggregations, to both manage the scale of the data and to preserve anonymity.

The **Personal Password Management (PPM)** is a web-application allowing users at the university to change their account passwords. Here a *password change* is the act of replacing an existing known password with a new one. *Reminder emails* are sent to system users in advance of a password being set to expire – a user will receive a series of timed emails at set intervals until they reset their password. There is a *grace period* of several days after expiry when a user can still reset their password. New passwords must conform to set rules, which define character composition and prohibit similarity to previous passwords. A password change may take time to propagate across all university IT systems.

The **Personal Password Recovery (PPR)** system allows users to register memorable phrases and associated hints so that they can authenticate themselves to the system if they have forgotten their password or it has expired – a user can then enact a *password reset*, to replace their password. The process can be enacted online personally, or via the helpdesk in-person or over the phone. Users must also provide a limited number of personal details. When answering memorable questions, users are prompted to only enter selected characters from their memorable answers. A *password reset* is not the same as *password recovery*, where a user may indicate on a webpage that they have forgotten their password and follow a direct link (or indirect link, perhaps via email) to a webpage to change their password – such a feature was not part of the system under analysis.

3.2 Helpdesk Log Analysis

Exploratory Data Analysis (EDA) can identify patterns within a dataset to visually summarise its primary characteristics [30]. The inherent advantages of EDA include that it allows researchers to uncover patterns and structures within the data that are not readily discerned. EDA is applied to the password helpdesk data to carry out (i) data munging, (ii) examine key variables and relationships that may exist between them, (iii) visually describe the data and, (iv) determine temporal patterns and trends that may be of interest to IT security managers. In this study, EDA findings also informed the design of a subsequent user interview exercise.

3.3 User Interviews

Semi-structured interviews were conducted, to facilitate probing of responses where they related to unanticipated factors or issues identified from log analysis. Thematic analysis was used here to identify coding themes and sub-themes based on topics emerging from the interviews. The analysis followed the six phases identified in [11]: familiarisation with data, generating initial codes, searching for themes, reviewing themes, defining and naming themes and producing the report. The objective of the analysis was to document and reason about the end-user experience with password resets, helping to explain findings in the log data.

The student population of the university was the focus of the interviews – staff at the university had a wider range of password management options available to them, but all users shared access to the same core set of functions considered in this paper. Also, staff may have had clerical, support, academic, or research roles, which would influence the systems they regularly accessed (where students had access to provisioned systems such as timetabling and teaching resources).

A pilot study was conducted to assess the interview questions. The final set of questions consisted of 9 open-ended questions, related directly to metrics in the helpdesk log data. A pre-screen survey determined eligibility for the study proper – the pre-screen determined that all participants were students at the subject university, and also recorded whether the participant had ever forgotten their password and “lost access to [their university account] for any amount of time as a result”. An information sheet was provided at the start of each interview, alongside a consent form to be signed. Each interview lasted approximately 20 min and was audio-recorded. Each participant received £5 for their participation.

3.4 NASA Raw Task Load Index (NASA-RTLX)

Interviews were complemented with a NASA-RTLX (Raw Task Load Index) workload measurement [14] for password use, where participants assess a task by providing ratings for perceived workload [17]. The scheme employs six subscales to assess user workload, on a scale of 0 to 100 for each subscale: mental, physical and temporal demands as well as performance, effort and frustration. Users assign a score for each individual factor (with all but performance being measured with a high score equating to a negative impact upon the respondent).

Workload measurements rely on users' recollection of completing a task [6]. Here NASA-RTLX served to promote discussion, and was used to explore potential metrics for user perception of password use. Three RTLX forms were prepared for each interview, assessing workload for password reset, PPR registration and PPR authentication. If an interviewee had not registered/authenticated to the PPR system, related RTLX forms were not included. RTLX responses were reviewed during the interview, as cues for discussion.

4 Results: Helpdesk Log Analysis

In this section we discuss the results of the log analysis and user interviews, relating interviews to log analysis to expose links between user experiences the log data. Specifics of the subject university's password policy are not discussed, so as not to identify the institution.

4.1 Results

Figure 1 shows time-plots of monthly summary data. There are noticeable peaks in resets around September, October and January of each year. These months follow longer periods of inactivity (when students and staff return from summer/winter holidays and begin study). Frequencies increase gradually over the timespan of the data, reflecting an increase in users. It is important to understand how support mechanisms are utilised, to manage the impact of increased activity. The month of September 2013 does not align with the same period in other years – where numerous systems interact in complex ways, mapping the relationships between data fields could help to rationalise such one-off events.

Figure 2 shows the different ways in which staff and students have reset their passwords. With a mean of 74.4 % the most-used approach to resetting passwords is via the PPM system, signifying that the majority of users follow the channel preferred for limiting helpdesk costs. However, there are a consistent minority of users resetting their passwords within the expiry "grace period", implying either that reminder emails are not completely effective, or that there is always a small

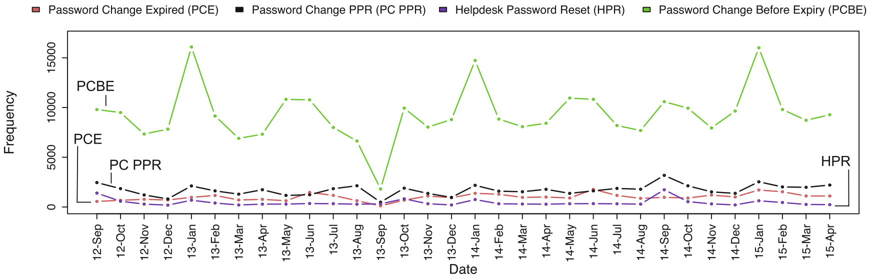

Fig. 1. Time plot for password related fields.

proportion of users who are unable or do not want to reset their passwords before expiry. Reasons for using the PPR and how users perceive it is also important for the interviews, as up to a fifth of resets are enacted through recovery mechanisms.

From Fig. 2, the provisioned recovery channels (i.e., PPR and helpdesk) are – as intended – not used by the majority of registered users. As illustrated in Fig. 2, password changes via the helpdesk are the least common method of choice for users, with an average of 3.6 %.

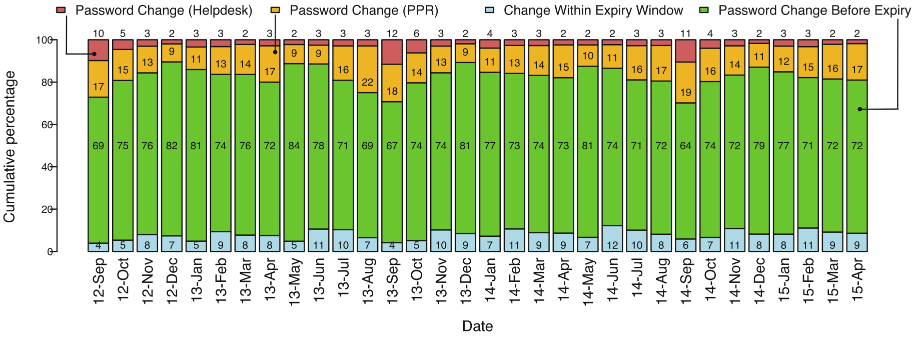

Fig. 2. Stacked bar plot for percentage of password change events.

The system that produced the data logs does not record any data to indicate the experience of users *as they are resetting their passwords*. This then becomes another area of focus for subsequent interviews. It is also unclear from the summary log data how individuals reset passwords in response to reminder emails, especially whether reminders are knowingly ignored (a general behaviour noted elsewhere [18]). How end-users react to notification emails is another subject to explore in interviews. Without recording for example the duration of password reset or recovery events (online or in-person, for comparison), it is unclear whether the helpdesk is little-used because of a comparative convenience in using the PPM system, or because of a lack of awareness about its existence. This is an important point to clarify if system managers ever intend to retire or reduce helpdesk support (for instance to manage costs, a concern touched upon in Sect. 2).

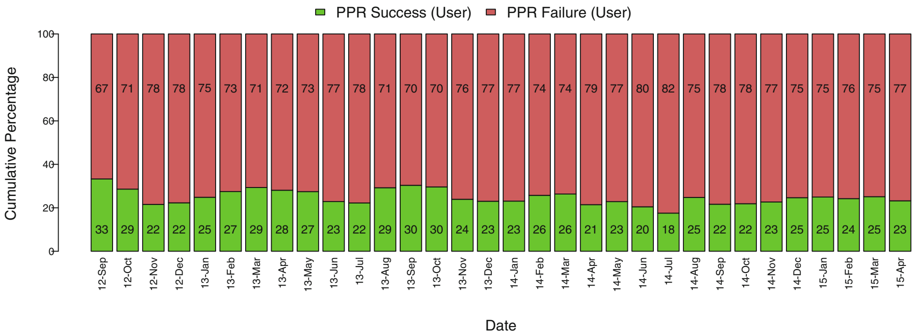

Fig. 3. Percentage of attempts that are failures/successes.

The relative proportions, expressed as percentages, between PPR failure and PPR success show that on average 75.2 % of attempts to authenticate to the PPR fail. Likewise, the mean success rate stands at 16 %. Figure 3 shows a time plot illustrating the relative percentages between PPR failure (user) and PPR success (user). If all such events were attributable to legitimate users the figures could be distressing, considering that based on the data available the lowest failure rate (September 2012) is 66.7 % and the highest 82.5 % (July 2014). The high failure rates may be due to the *demands* of the recovery process, or may reflect success in managing automated attempts to access the system – other work has for instance highlighted that it can be difficult to distinguish brute-force attacks from other attempts to access user accounts [1]. The PPR failure and success rates also raise the question of how many PPM password reset attempts fail (something which is not recorded in the dataset) – user interviews examine the personal impact and perception of failure to create a new, policy-compliant password.

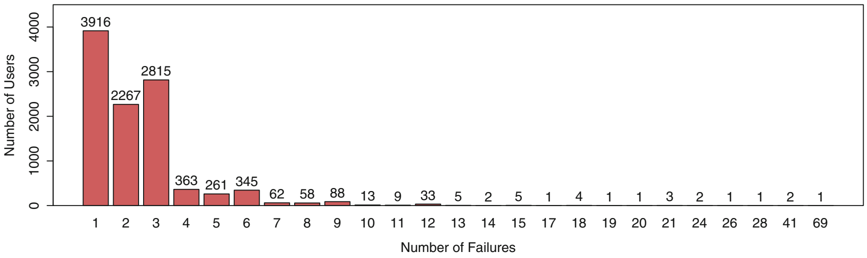

Fig. 4. Number of users failing PPR authentication x times within a 24 h period of using the system.

Going beyond the aggregated data, an event-by-event dataset was examined to explore PPR failure rates. This dataset associated anonymised user identifiers with specific PPR/PPM/helpdesk events, enabling further exploration of the PPR failure counts. Figure 4 shows a bar plot of the number of PPR failures for a user within a 24 h period (representing a unique session of interaction with the PPR). From Fig. 4, there are many users who are failing 1-2 times each time they use the system. However, the number of users who are failing 3 times also represents the second highest figure in the chart – this is important as after 3 successive failures access to the PPR is locked for a time. At this point a user may decide to go to the helpdesk or wait before trying to use the PPR again (where the latter may explain the lower occurrences of higher failure rates beyond 3 attempts). Given the number of users at the subject institution this may constitute a large drain on working time for both users and helpdesk staff. The effectiveness of interactions with the PPR system was then discussed in interviews. Understanding responses to failure would also rationalise the distinction between the "noise" of unusable security and genuine attempts to access the system without first-hand knowledge of the associated credentials. There is a long-tail distribution to the tallies of failed PPR use. It becomes unclear as tallies increase whether this represents determined users or automated authentication attempts, where the distinction (and perhaps then the response) is not clear. The high PPR failure rates were then of further interest in subsequent interviews.

5 Results: User Interviews and NASA-RTLX

The goal of the interviews was to explore the perceived impacts that the password reset policy is having on its users, in this study specifically students. Equally the study aimed to determine how user feedback could be used to explain findings identified in the EDA process, with interview questions targeted to explore specific outcomes from the log data analyses.

Interviews were transcribed, and transcripts re-read many times to become familiar with the data, and a priori notes and patterns were recorded to develop codes. An initial list of codes was generated by one of the authors following data reduction (e.g., subtracting and merging codes) and complication (e.g., adding and splitting codes). An inductive thematic analysis approach derived concepts and perspectives from the interview data. A code book was maintained throughout. This codebook was reviewed at regular intervals by three of the authors.

From the pre-screen responses, 20 participants were selected to be interviewed, all students at the university. 16 were female and 4 male, mean age was 25. Furthermore, 9 participants reported being registered on the PPR system whilst 11 stated they were unaware of it.

5.1 Results

Our analysis identified four overarching themes (with a total of 10 sub-themes) which we described in the following sections.

Impact of Stringent Policy. Many interviewees felt that the password reset policy was restrictive, with many using words "*annoying*" (10 interviewees) and "*inconvenient*" (8) to describe changing their password. Interviewees identified themselves as security-conscious and understanding of the need for secure behaviour, but nonetheless the majority felt burdened by the process.

(A) Difficulty Meeting Criteria. For the majority of interviewees (17), creating a new password that met the necessary criteria required multiple attempts, and was a process of trial and error. Some (8) felt that the password criteria were too strict, with the perception that no words or names could be used. For some (6), there was an expectation of having problems with every password change. Two interviewees implied that their initial experience informed their impression of how the process would be in future, P8 explained: "*...so like when I* [went to] *reset it, like my first password. It was a pain because I had to* [try] *like 10, 15 times in a row and then I went: "Ah okay, this is gonna be fun if I have to do it again.""* Several respondents (9) expressed frustration at having produced passwords, which they felt met the criteria, but which were then rejected by the system. The difficulties of failed or protracted attempts to change passwords were not represented directly in the helpdesk logs.

(B) Overwhelming Frequency of Password Change. More than half of participants (13) felt that password reset requests were too frequent with 6 considering the frequency was more of a problem than meeting the creation criteria. P8 pointed out that passwords are subject to such strict requirements that they should be kept for longer: "*...the password, like the requirements are so strict so I don't see the point of changing it like 2 or 3 times a year.*"

When contrasting the policy with other accounts they had used, a number of participants (7) indicated that the university's policy required the *most* frequent changes; they were puzzled as to why the frequency was so high, "[At] *my previous university where I was for 4 years, I didn't have to reset my password at all, during the 4 years!*" (P13).

Effectiveness of Support

(A) Lack of Information. Some (6) participants saw "*no rational explanation*" for the policy, querying why it was so stringent. P15 explained: "*I mean if there are any pertinent security reasons for having to change it so often, I think you better explain why for non-technical users. Um, cause it's like...I'm just feeling angry that I have to change so often if I haven't understood why.*". Several interviewees (7) felt that insufficient feedback during a password reset made it a

"*guessing game*" to satisfy the various rules, as P17 told us: "*It just tells me that it's just too similar or it has to have this and it... I think it tells you one thing at once.* [...] *Instead of explaining everything in one go... What a good password should look like, I have to* [use trial] *and error every time.*" This implies that unique attempts to create a password may take a long time or that there may be many failed attempts in quick succession before a successful password change.

(B) Efficacy of Email Reminders. 8 participants *knowingly* ignored reminder emails as they were either not near a computer to carry out the reset (a phone screen was regarded as inconvenient to type on), or did not feel a sense of urgency to do so (6); "*I read* [the reminder email] *and... I figured I could do that at some point in time and I didn't really see the urgency and then I got a second reminder and then I thought: "Okay, now I have to do it.""* (P17).

Nearly all interviewees (18) reset their passwords before the deadline, which aligns with the log data analysis (see Fig. 2), and implying that the reminder emails were having an effect. For some (5), this was because the email reminder prompted a *fear* of the unknown as to the consequences of missing the deadline. However, three participants did perceive reminders as an effective tool that helped to avoid expired passwords. This highlights that "successful" users of the system may or may not at the same time experience stress in complying with expectations.

(C) Recovering Forgotten Passwords Is Not a Problem. Some (5) participants had at one time employed the PPR service to reset their forgotten password which all described as "*straightforward*" (what may be regarded as comparable to the log data in Fig. 2). Though no experiences were reported of being locked out of an account, 3 had experienced difficulties in recalling memorable answers before successfully authenticating. PPR registration was encouraged when joining the university, where memorable answers may be required some time after registration, challenging the capacity to correctly recall configured answers (or in the case of the study participants, recall whether they had registered for the system at all). Similarly, in a study on user-chosen challenge questions, Just and Aspinall [21] found that 18 % of their participants were unable to provide at least one of their answers correctly after 23 days from creation and this despite, as the authors emphasised, the participants being young and potentially having excellent memory.

Of the 20 participants, 9 had registered for the PPR service whilst 3 had heard of it but not registered and the remainder were not aware of its existence. Those who had visited the helpdesk for queries – a minority, as reflected in the log data represented in Fig. 2 – described it as a "*clear cut*" and "*quick*" procedure. These accounts raise the question of how well the support systems are publicised.

Development of Coping Strategies

(A) Coping with Frequency of Reset. Participants were divided between resetting upon receipt of the initial email (10) and relative to the expiry date of the password (10). This division is not reflected in the metrics of the helpdesk logs, where a reset anywhere between the initial email and the expiry date is logged as a "Password Change" either way (as in Fig. 2).

Two participants deliberately delayed resetting until the last few days preceding the expiry date, to put off the next time that they would have to complete the process. Four participants also felt that delaying the reset allowed them to plan ahead. It is possible that similar strategies can explain password resets during the expiry grace period (as seen in the log data, Fig. 2), in that for some users there may be a deliberate effort to delay interactions with the system as long as possible (and not just for instance representing users who experience difficulties in changing a password in time).

(B) Coping with Password Creation. More than half of the interviewees (12) developed a strategy for creating passwords that were both acceptable and memorable. Most common was a scheme of sequential changes, taking words from a category of items (e.g., colours, city names), and modifying them to meet the policy requirements. A new item was then derived from the category with each change. Such strategies are examined elsewhere [3], where the genuine change in composition (and guessability) across old and new passwords is put into question.

(C) Coping with Recalling Passwords. The most evident strategy for recalling passwords was for users (6) to write their passwords down [4] which either took the form of physical paper (4) or an electronic document (2). Most of these individuals (4) acknowledged the security risk in doing so yet felt compelled to do so given the complex criteria. With paper-based recall aids, some interviewees (3) admitted to having lost it. These "offline" recall aids would require a further investment of resources for security managers to monitor. Five participants stated that they only work from their personal computers, and hence reduce the burden of remembering by caching passwords in their browser (a behaviour that is not unique and has been noted elsewhere [10]).

Shaping Perceptions Towards Security

(A) Accepting of Security, But... Many (10) acknowledged the need for security despite feeling that it was a *pain*. P7 explained: "*Oh yeah, I understand it should be done like that... Because you know it can be something quite personal, with my degree and stuff. But um... It just gets annoying, it's like both sides anyways. I find it annoying but I understand it has to be done.*". A number of users (7) also felt a sense of immunity to security breaches and as such did not comprehend the level of security being mandated by the policy. P5 told us: "*I only have my emails which are literally just: "Come to the seminar, it's next week." and I have no sensitive information in my email. [...] I think my personal email might... require more of these mandatory password changes*".

(B) the Definition of Security. 9 participants had the impression that the university takes security “*very seriously*”, acting as a model of what constitutes “security”. P3 explained: “*I think* [it’s] *definitely one of the most secure from that point of view, I’ve never had to change my password so often.*”. For some, this influenced their decisions in managing passwords. 4 participants admitted setting other account passwords to mirror their university password, in part because the university rules seemed to be the most “*secure*”.

5.2 RTLX Data Analysis

All participants completed the RTLX ratings for the password reset task. RTLX for PPR registration and authentication applied to 9 and 5 of the participants respectively.

Figure 5 shows box plots of the subscales for each task (with the scale of Performance inverted so that higher values equating to better perceived task performance). With mental demand, the password reset task was comparatively more challenging than the PPR-related tasks, with a mean of 58 (on a scale of 0 to 100) for the password reset task whilst for PPR registration and authentication, these values were at 31.6 and 43 respectively. The distributions for the PPR-related tasks suggest that some users of the system would have difficulty creating/remembering memorable answers. Temporal demand was higher and more widely-distributed for the password reset task than for PPR tasks (which again would not have been reflected in the log data).

For password changes/resets the performance varied, further highlighting that participants found it difficult to generate a valid password. With PPR-related tasks, perceived performance was high, with users feeling that they successfully managed these tasks – this is in contrast to the EDA findings that apply to the larger base of users.

Frustration levels were the highest measure for password reset tasks, with a mean of 61.8. Frustration was also notable for PPR authentication, perhaps owing to recalling memorable answers – this and the related mental demand may help to explain some of the high PPR failure rate described in the log data

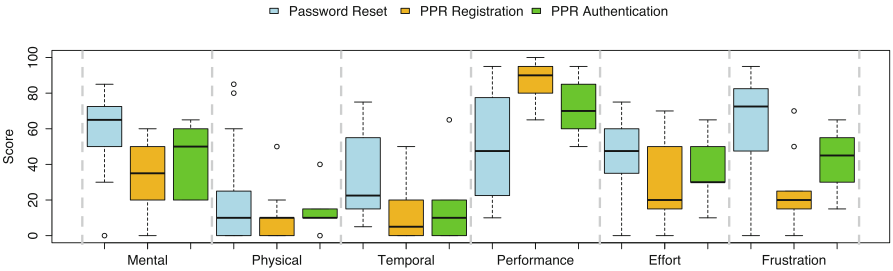

Fig. 5. Box plot of average RTLX subscale results for password reset (20 responses) and PPR registration (9 responses) and PPR authentication (5 responses) tasks.

analysis results. The low frustration for the PPR registration process complements the perception of the task being "*straightforward*". Regarding the frustration scale for the password reset task, dialogue with interviewees implied that reminder emails and the actual task of creating a new password were all part of an interlinked process – this may create some uncertainty as to what it was that participants assigned scores to, but reinforces that the perceptions of reminder emails and resetting passwords can impact (and potentially reinforce) each other.

The mental demand of PPR authentication was more varied (exhibited a wider span of responses on the TLX subscale) compared to password reset and PPR registration. This could be attributed to needing to recall "memorable" answers which were set potentially months (if not years) beforehand, where some participants used what they regarded as knowingly obvious "memorable answers" and others had difficulty in recalling their answers later on when they needed them.

6 Discussion

Interviewees devised coping strategies to both produce passwords that the system would not deem similar to previous ones, and avoid the effort of resetting (e.g., postpone the next reset for as long as possible). The latter may contribute to – and be masked by – the peaks in resets seen in log data (Sect. 4.1), warranting further investigation. Coping strategies have been noted during other studies with university participants [3] (especially modification of a previous password), suggesting that this is far from being an isolated case.

Interviewees regarded the stringent policy as the "*most secure*" they had seen, thereby confounding number of rules with security. Participants felt that replicating their university password across their other accounts was appropriate to protect their (personal) data. It may be that students especially may take these habits and understanding with them after their time at university and into workplace environments.

The combination of log analysis and user interviews indicated that the Personal Password Recovery (PPR) facility may have usability issues, as a great proportion of attempts to authenticate to the system resulted in failures (where the significance of authentication failures of genuine users within the data warrants further analysis, as hinted in Fig. 4). Use of the PPR system is then counter to the expectation that self-service mechanisms minimise helpdesk calls [13], although the log data did show that visits to the helpdesk were minimal (as in Fig. 2). To the contrary, interviewees who had used the PPR expressed a general consensus that it was "*straightforward*" to use, with relatively low TLX values for both registration and authentication tasks. These results raise the possibility that there is a *subset* of users, under-represented in our interviewee cohort, who consistently have difficulties authenticating to the system. Consideration should be given to how to engage with these "forgotten users".

The log data did not indicate the effectiveness of reminder emails; resets after expiry were low, though non-negligible (Fig. 2). Some interviewees felt that reminder emails were too frequent and too early, where they ignored them or developed a distracting sense of *pressure* to change their password. This contributed to the wide-ranging responses for the RTLX "temporal demand" measure for password resets.

For interviewees who replicated their university password across other accounts, the workload of resetting their university password would then be compounded by the number of accounts they have in and outside the university. Not having a coping strategy is also a drain on productivity, see P13: "*And so you spent time thinking of something, you put that in and then it [tells you it's not valid] and then you go: "Alright, I have to think of something else"*". Logging both the *number of attempts* to create a password and the *amount of time* to devise a valid password could indicate the usability of the composition rules and whether predefined coping strategies are being used.

Interviewees lacked comprehension as to why such stringent rules were necessary, which contributed to perceived inconvenience. A perceived lack of reasoning for the password policy may also support the sense of immunity to threats voiced by some participants. It is then important to relate policy to the context in which an account is used, P10 stressed: "*So I'm guessing it's the same policy for all...From undergrad to like PhD and bachelors and all. So I can see why higher up the chain they probably* [need that level of security] *at a high level but at undergrad it might not.*".

6.1 Recommendations for Practitioners

1. Provide users with *real-time* feedback during password creation, which may reduce number of attempts, frustration levels, and time expended on creating a new password.
2. Improve the communication and justification of password policies to users. That is, *explicitly* relate potential/existing threats to what they do with their accounts.
3. Do not *assume* that policies are highly usable even if there is evidence that a majority do not require helpdesk support, as users may devise coping strategies which although compliant with policy may be counter-productive to the security of both the individual and the organisation.

7 Conclusions

Password support log data at a university was analysed. Taken by itself, analysis suggested that most users comply with the university's password policy with few problems. However, there may be hidden costs not reflected in typical helpdesk logs – many interviewees saw the policy as too stringent and frustrating, in some cases developing (insecure) coping strategies, tailored specifically towards complying with the policy. Several interviewees delayed password resets and the effort they associated with the process. A number of interviewees interpreted the strictness of the password policy to mean it was a *secure* policy.

Interviewees who had used the Personal Password Recovery (PPR) system indicated through NASA-RTLX workload scores that it was a *straightforward* process. Analysis of both aggregated data and event-by-event log data indicated that most attempts to provide memorable answers for a specific account failed 1-2 times, and that a minority of users may be experiencing trouble using the system (or otherwise that illegitimate attempts to access accounts are difficult to distinguish from genuine attempts, exacerbating the matching of resources to the support of users). The lack of feedback and justification of policy that users saw when creating passwords contributed to frustration.

Future work will continue analysis of the event-by-event log data obtained after initial analysis and reporting, and will engage with students and staff equally. Initial focus will be on responses to reminder emails and individuals' interactions with the PPR facility. It was noted that interviewees treat the entire password reset process as a single task (from initial email reminder to devising a new password), and so workload assessment techniques which acknowledge secondary tasks (such as security) may be necessary.

Findings highlighted that the log data offered limited insight into the user experience of password management, such as habits and points of frustration. Measures of effort and consequences are then important to consider, as these were shown to have potential hidden costs for security managers. NASA-RTLX assessments that quantify perceived workload were a tentative first step. In a similar vein, findings may inform repeatable surveys deployed on a larger scale to reach both students and staff (particularly long-term users).

Acknowledgements. Simon Parkin and Angela Sasse's research is funded in part by EPSRC, grant number: EP/K006517/1 ("Productive Security"). The authors would like to thank the participating university and especially their IT department for providing the data that informed this publication. The authors would like to thank Ingolf Becker for his help with the editing of this paper.

References

1. Florêncio, D., Herley, C.: Where do security policies come from?. In: Symposium on Usable Privacy and Security, p. 10. ACM (2010)
2. Kirlappos, I., Parkin, S., Sasse, M.A.: Learning from "Shadow Security": Why understanding non-compliance provides the basis for effective security. In: NDSS Workshop on Usable Security (USEC) (2014)
3. Shay, R., Komanduri, S., Kelley, P.G., Leon, P.G., Mazurek, M.L., Bauer, L., Cranor, L.F.: Encountering stronger password requirements: user attitudes and behaviors. In: Symposium on Usable Privacy and Security (SOUPS). ACM (2010)
4. Adams, A., Sasse, M.A.: Users are not the enemy. Commun. ACM **42**(12), 40–46 (1999)
5. Adams, A., Sasse, M.A., Lunt, P.: Making passwords secure and usable. In: Thimbleby, H., O'Conaill, B., Thomas, P.J. (eds.) People and Computers XII, pp. 1–19. Springer, London (1997)
6. Albers, M., Patton, J.T.: Measuring cognitive load to test the usability of web sites. In: Annual Conference-Society for Technical Communication, vol. 53 (2006)

7. Anderson, J.: Why we need a new definition of information security. Comput. Secur. **22**(4), 308–313 (2003)
8. Arnell, S., Beautement, A., Inglesant, P., Monahan, B., Pym, D., Sasse, M.A.: Systematic decision making in security management modelling password usage and support. In: International Workshop on Quantitative Aspects in Security Assurance, Pisa, Italy (2012)
9. Beautement, A., Sasse, M.A., Wonham, M.: The compliance budget: managing security behaviour in organisations. In: Proceedings of the 2008 Workshop on New Security Paradigms. ACM (2009)
10. Besnard, D., Arief, B.: Computer security impaired by legitimate users. Comput. Secur. **23**(3), 253–264 (2004)
11. Braun, V., Clarke, V.: Using thematic analysis in psychology. Qual. Res. Psychol. **3**, 77–101 (2006)
12. Brostoff, S., Sasse, M.A.: Ten strikes and you're out: Increasing the number of login attempts can improve password usability. In: Proceedings of CHI 2003 Workshop on HCI and Security Systems (2003)
13. Broome, C., Streitwieser, J.: What is E-support. Service and Support Handbook. Help Desk Institute, pp. 31–40 (2002)
14. Byers, J.C., Bittner, A.C., Hill, S.G.: Traditional and raw task load index (TLX) correlations: are paired comparisons necessary. Adv. Ind. Ergon. Saf. **I**, 481–485 (1989)
15. Coles, R.: Keynote address. In: Eighth Workshop on the Economics of Information Security (WEIS 2009), pp. 24–25. University College London, England (2009)
16. Charoen, D., Raman, M., Olfman, L.: Improving end user behaviour in password utilization: An action research initiative. Syst. Pract. Action Res. **21**(1), 55–72 (2008)
17. Hart, S., Staveland, L.: Development of NASA-TLX (Task Load Index): results of empirical and theoretical research. Adv. Psychol. **52**, 139–183 (1988)
18. Herley, C.: 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. ACM (2009)
19. Inglesant, P., Sasse, M.A.: The true cost of unusable password policies: Password use in the wild. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM (2010)
20. Jakobsson, M., Myers, S. (eds.): Phishing and Countermeasures: Understanding the Increasing Problem of Electronic Identity Theft. Wiley, New Jersey (2006)
21. Just, M., Aspinall, D.: Personal choice, challenge questions: a security and usability assessment. In: Symposium on Usable Privacy and Security (SOUPS). ACM (2009)
22. Parkin, S., Inglesant, P., Sasse, M.A., van Moorsel, A.: A stealth approach to usable security: helping IT security managers to identify workable security solutions. In: Proceedings of the 2010 Workshop on New Security Paradigms. ACM (2010)
23. Post, G., Kagan, A.: Evaluating information security tradeoffs: Restricting access can interfere with user tasks. Comput. Secur. **26**(3), 229–237 (2007)
24. Reeder, R., Schechter, S.: When the password doesn't work: secondary authentication for websites. IEEE Secur. Priv. **9**(2), 43–49 (2011)
25. Sasse, M.A.: Computer security: Anatomy of a usability disaster, and a plan for recovery. In: Workshop on Human-Computer Interaction and Security Systems, CHI (2003)
26. Sasse, M.A., Brostoff, S., Weirich, D.: Transforming the 'weakest link' a human/computer interaction approach to usable and effective security. BT Technol. J. **19**(3), 122–131 (2001)

27. Sasse, M.A., Fléchais, I.: Usable security: Why do we need it? How do we get it? In: Cranor, L.F., Garfinkel, S. (eds.) Security and Usability: Designing Secure Systems that People can use, pp. 13–30. O'Reilly (2005)
28. Skaff, G.: An alternative to passwords? Biometric Technol. Today **15**(5), 10–11 (2007)
29. Tari, F., Ozok, A.A., Holden, S.: A comparison of perceived and real shoulder-surfing risks between alphanumeric and graphical passwords. In: Symposium On Usable Privacy and Security (SOUPS) (2006)
30. Tukey, J.: Exploratory Data Analysis. Addison-Wesley, Reading (1977)
31. Whitten, A., Tygar, D.: Why Johnny can't encrypt: a usability evaluation of PGP 5.0. In: Proceedings of the USENIX Security Symposium (1999)
32. Zviran, M., Haga, W.J.: A comparison of password techniques for multilevel authentication mechanisms. Comput. J. **36**(3), 227–237 (1993)

http://www.springer.com/978-3-319-29937-2

Technology and Practice of Passwords
9th International Conference, PASSWORDS 2015,
Cambridge, UK, December 7-9, 2015, Proceedings
Stajano, F.; Mjølsnes, S.F.; Jenkinson, G.; Thorsheim, P.
(Eds.)

2016, XV, 151 p. 19 illus. in color., Softcover
ISBN: 978-3-319-29937-2