“Too taxing on the mind!” Authentication grids are not for everyone

Kat Krol[1], Constantinos Papanicolaou[1], Alexei Vernitski[2], M. Angela Sasse[1]

[1] University College London (UCL), Department of Computer Science, London, UK
{kat.krol.10,constantinos.papanicolaou.12,a.sasse}@ucl.ac.uk
[2] University of Essex, Department of Mathematical Sciences, Colchester, UK
asvern@essex.ac.uk

Abstract. The security and usability issues associated with passwords have encouraged the development of a plethora of alternative authentication schemes. These aim to provide stronger and/or more usable authentication, but it is hard for the developers to anticipate how users will perform with and react to such schemes. We present a case study of a one-time password entry method called the Vernitski Authentication Grid (VAG), which requires users to enter their password in pairs of characters by finding where the row and the column containing the characters intersect and entering the character from this intersection. We conducted a laboratory user evaluation (n=36) and found that authentication took 88.6s on average, with login times decreasing with practice. Participants were faster authenticating on a tablet than on a PC. Overall, participants found using the grid complex and time-consuming. Their stated willingness to use it depended on the context of use, with most participants considering it suitable for accessing infrequently used and high-stakes accounts and systems. While using the grid, 31 out of 36 participants pointed at the characters, rows and columns with their fingers or mouse, which undermines the shoulder-surfing protection that the VAG is meant to offer. Our results demonstrate there cannot be a one-size-fits-all replacement for passwords – usability and security can only be achieved through schemes designed to fit a specific context of use.

1 Introduction

The pressure to replace passwords with other authentication solutions has been growing. The number of passwords that the user has to create and remember has been steadily increasing leading users to create weaker but more memorable ones, hand their lives over to a password manager, write passwords down or simply reuse the same credential or a variation for multiple systems (Florêncio, Herley, & Van Oorschot, 2014). Many new authentication schemes have been created with the aim to make authentication easier and/or more secure. But with the exception of the Android Unlock Pattern, none of these replaced passwords – rather, widely adopted solutions have added to the password scheme to create 2-factor authentication, such as Google Authenticator or the 2-factor solutions deployed in online banking (Krol, Philippou, De Cristofaro, & Sasse, 2015). From the security point of view, a password scheme devised today would need

To access the published version of this paper, please visit: http://link.springer.com/chapter/10.1007/978-3-319-20376-8_7

to protect against passwords being collected through malware, phishing and shoulder-surfing. In this paper, we present our preliminary lab-based evaluation of a scheme that was designed to protect against these types of attacks – the Vernitski Authentication Grid (VAG) (Vernitski, 2015).

The VAG was created by one of the authors, the mathematician Alexei Vernitski. Users choose a password consisting of an even number of letters and/or digits. The entry of this password is through a 6x6 grid (the order of characters is random and reshuffled on each authentication) and the user needs to enter the characters of their password in pairs. They need to find the row that contains the first character and the column that contains the second character and enter the letter/digit that is in their intersection. Then this is repeated for each next pair of consecutive characters until the end of the password.

We do not consider the VAG a graphical password since there is no graphical element, such as a picture (Biddle, Chiasson, & Van Oorschot, 2012), that the user would need to remember. Instead they are asked to remember a password and then enter it using a grid. The mechanism is a form of challenge-response (C-R) authentication. The aim of the grid is to guard against shoulder-surfing and key-logging. It can protect the password at three stages: (1) at entry in a case when the interface cannot be trusted, for example, when using an unknown computer, (2) at entry when the physical environment is hostile, for example, in the presence of a potential shoulder-surfer and (3) in transit when communication can be eavesdropped on. Capture of the secret password shared between user and system is made more difficult by the characters in the grid being reshuffled each time. If the attacker has a key-logger installed on the users' computer, all they receive will be a set of random-looking characters. A statistical security analysis of the VAG was conducted by Papanicolaou (2013).

Another potential use case for the grid would be as back-up or very infrequent authentication – such as annual tax returns. Long periods of not recalling a password or backup credential lead to high failure rates. To increase memorability of the shared secret, users could pick a password with high personal entropy (Ellison, Hall, Milbert, & Schneier, 2000) – a word or phrase that is hard to guess, but meaningful to the user, and strongly embedded in biographical memory.

The remainder of the paper is organised as follows. In Section 2, we review relevant existing work. In Section 3, we outline the set-up of the study followed by a presentation of both qualitative and quantitative results in Section 4. We discuss our findings in Section 5, and conclude and provide recommendations in Section 6.

2 Related work

There is a large body of research looking at graphical passwords, a survey by Biddle, Chiasson and Van Oorschot (2012) provides an overview of the schemes and their evaluations. Although the VAG is not a graphical password, some of the graphical passwords studied were grid-based which can make their findings relevant to our evaluation. GrIDsure is one such example, users are asked to memorise a pattern and then when authenticating they receive a grid filled with digits and need to enter the digits

that correspond to their pattern. Brostoff, Inglesant and Sasse (2010) looked at the usability of GrIDsure and found that in nearly 18% of usages, participants were trying to enter the PIN on the grid directly instead of typing it. This undermines the security property offered by the grid, namely resistance to shoulder-surfing.

New authentication schemes were also shown to alter user behaviour when accessing systems. Brostoff and Sasse (2000) found that participants logged in with one third of the frequency when they authenticated using a grid of Passfaces® rather than passwords, because the former took significantly longer. In a study by Steves et al. (2014), participants reported they did not follow their natural workflow but batched multiple activities on the same system together to save on authentication time and workload.

With the emergence of touchscreen devices, we are increasingly moving away from a traditional screen and keyboard set-up to virtual keyboards. Previous studies have demonstrated that password entry on touchscreens can be significantly more difficult and time-consuming (Greene, Gallagher, Stanton, & Lee, 2014). Schaub, Deyhle and Weber (2012) investigated the usability and security of six types of virtual keyboards. They showed that the keyboards differed in the usability of password entry (entry time, accuracy) and susceptibility to shoulder-surfing. They found that keyboard designs with poor usability were more resistant to shoulder-surfing. Moreover, research has demonstrated that the entry method affects users' password choice and security as users choose passwords that are easier to enter on touchscreens (Yang, Lindqvist, & Oulasvirta, 2014).

3 Study set-up

3.1 Design

The study was conducted in a laboratory with one experimenter and one participant at a time. There were two groups of participants: first one with 31 and the second one with 5 participants. The first stage was the same for both. After being introduced to the scheme, participants were asked to perform six logins. While the first group authenticated six times during one lab session, the second group was asked to return for another session a week later where they were asked to perform another six logins. Each participant in the first group ended their session with a brief interview where the experimenter asked them about their experience.

3.2 Study goals and hypotheses

The purpose of the study was to evaluate the user experience of the VAG both quantitatively and qualitatively. The study was meant to be a preliminary evaluation, looking at what users think of the grid, what the learning curve is to use the system and generally explore users' experiences of authentication, their expectations and preferences.

We devised the following hypotheses.

H1: There will be a difference in the time that participants take to enter a password on a PC and a tablet.

H2: The time of entry will decrease the more practice participants have.
H3: Authentication speed will depend on participants' individual characteristics such as age, computer literacy and experience with touchscreens.

3.3 Procedure

Upon arrival, participants received an explanation of what the study will involve, they were asked to read an information sheet and sign a consent form. The workings of the grid were then explained to them using a laminated sheet of paper with a grid on it and a marker pen. Participants were then asked to enrol by setting up a username and password. The password had to be of an even number of characters. Once they were finished, they were asked to perform six logins overall: three on a PC and three on a tablet (order counterbalanced). After having performed these logins, participants were briefly interviewed about their experience and then asked to fill out a brief questionnaire about demographics as well as their computer literacy and cyber-threat exposure.

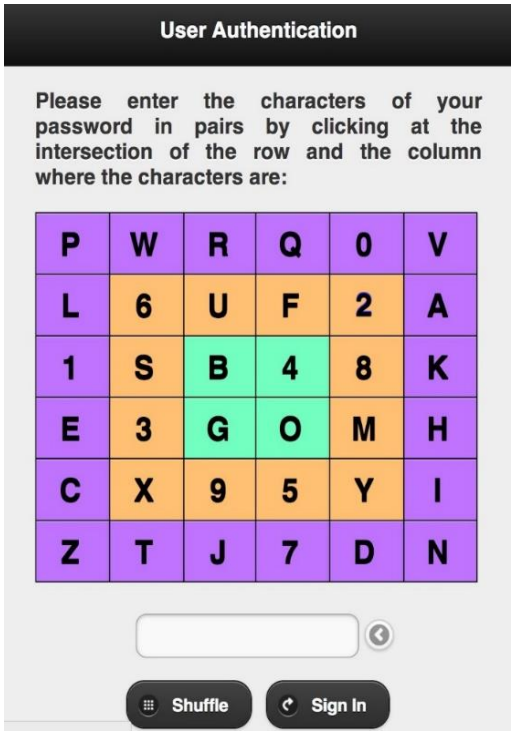

Figure 1. A screenshot of the login page used in the study. The "Shuffle" option generates a new grid.

3.4 Apparatus

The prototype of the VAG used in the study (Fig. 1) was programmed in Java (using JSP/Servlet technology) and was linked to a MySQL database. The JQuery framework was also used to enhance the interactivity of the prototype and provide a better user experience. The study was performed on a PC running Windows XP using a 22″ monitor with a resolution of 1920×1080 pixels and a 9.7″ iPad 2 with a resolution of 1024×768.

3.5 Participants

The study was conducted in August 2013. The research was exempted from an ethics committee review. We recruited our participants through a participant pool at University College London. Anyone was welcome to participate as we did not set any requirements. Participants were paid £6 for their participation which took around 30 minutes. Overall, there were 36 participants in the study, 19 male and 17 female. Mean age was 26.9 years (range: 20–49, *SD*=7.2). In terms of education, 28 participants had completed a university degree and 7 had A-levels (UK school leaving certification).

4 Results

Overall, across all participants, devices and trials, a login attempt (regardless of if successful or not) took 63.7s. However, in real life failed attempts to authenticate add to the time needed to access a system, therefore in our analyses we consider the cumulative time needed for a successful login, which means we counted the times of failed attempts too. The cumulative average time for a login using the VAG across all participants, devices and trials was therefore 88.6s (median=32.5). This average is skewed by some trials requiring several attempts. Therefore, to illustrate it better, we can say that 45% of logins took under half a minute, 70% under 1 minute, 79% under 1.5 minutes and 85% under 2 minutes. Figure 2 shows login times across different trials and treatments. There were six logins and the number of attempts participants needed to successfully log in ranged from the required 6 to 14 (*M*=8.16, *SD*=2.11). On average, 1.4 attempts were needed for a successful login (*SD*=0.8). Passwords chosen by participants were on average 6.7 characters long (*SD*=1.24, range: 4–10). Out of 36 participants, 22 chose a 6-digit password, presumably influenced by the example password (`zebra1`) given by the experimenter which had 6 characters.

4.1 Hypotheses

We conducted a 2 (Device: PC, Tablet) × 6 (Trials) × 2 (Order: PC first, Tablet first) repeated measures ANOVA on the time needed to log in. There were significant main effects of Device ($F(1,34)=12.6$, $p=0.001$), Trial ($F(2,33)=9.74$, $p=0.001$) and Order ($F(1,34)=9.8$, $p=0.004$). There was a significant Device × Order interaction ($F(1,34)=5.24$, $p=0.028$) and Device × Trial interaction ($F(2,33)=3.35$, $p=0.041$). There

was a marginally significant Trial × Order interaction ($F(1,34)=2.49$, $p=0.091$). The three-way Device × Trial × Order interaction was non-significant ($F<1$).

H1. Device type. We hypothesised that there would be a difference between how long participants would take on a PC and on a tablet. On average, our participants needed 118.8s to log in on a PC and 58.4s on a tablet (see Figure 2). The ANOVA test described earlier showed a significant main effect of Device indicating that participants were faster to log in on a tablet than on a PC. Post-hoc effects indicated this difference is statistically significant ($U=319.5$, $p=0.02$). H1 is therefore supported.

H2. Practice. We hypothesised that login times would decrease the more practice our participants had. We could clearly see the learning curve in that participants' login times were long at the first trial but decreased by the third trial. Upon switching to the other device, the login time was longer at first trial but decreased again with practice (see Figure 2). Participants who started on a PC were slower in their first trial than those who started on a tablet ($U=39$, $p=0.008$). After switching to the other device, participants who used a tablet first and switched to a PC were marginally faster than those who switched from a PC to a tablet but we did not find this difference to be statistically significant. The group that started with a tablet was on average authenticating faster than the group that started on a PC ($U=3931$, $p<0.001$). Assignment to these groups was random and we could not find any significant differences between the two groups. For our smaller sample of five participants who authenticated on two occasions, we hypothesised that participants would be faster in their second authentication session. For the first session, the average authentication time was 54.4s and for the second 39.9s. Each participant authenticated faster in the second session by an average of 14.6s. Despite this trend, we found this difference not be statistically significant ($p=0.4$). A larger sample of participants would be needed to be able to decisively prove or disprove this hypothesis.

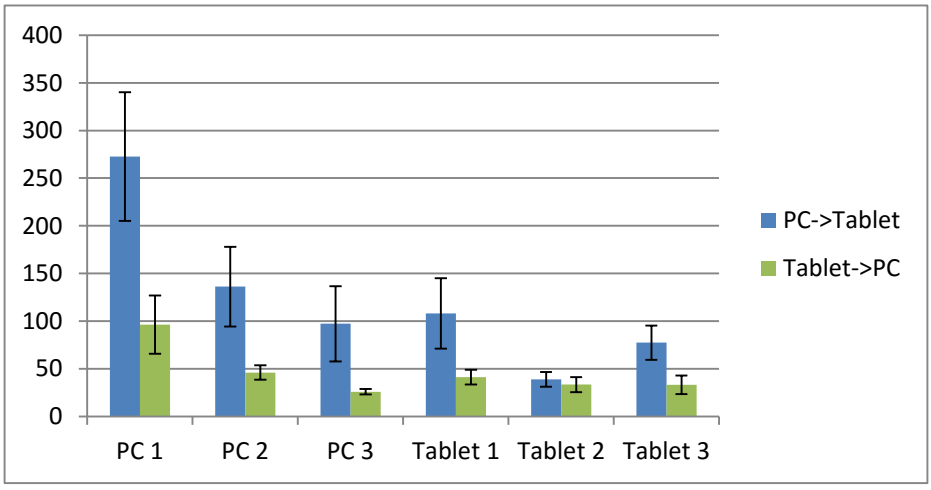

Figure 2. Login times in seconds presented across different trials.

H3. Personal characteristics. We hypothesised that age, computer literacy and experience with touchscreens would influence the speed with which participants authenticated. We found a strong positive correlation between age and authentication time (r=0.42, p=0.01), the older our participants were the slower they authenticated. We found a moderate negative correlation between computer literacy and authentication time (r=-0.36, p=0.03), the more computer literate a participant, the faster they authenticated. Finally, we did not find a statistically significant correlation between authentication speed on a tablet and experience with touchscreen devices (p=0.095).

Apart from taking objective performance measures, we also conducted structured observations of participant behaviour while they authenticated. Out of 36 participants, 31 pointed at the grid with their fingers or mouse when they were trying to find the row and the column where the characters of their passwords were. Three participants wanted to write their password down to facilitate breaking it down into pairs of characters.

4.2 Interview results

After the login tasks, participants were asked to share their experiences with us. The brief interviews were audio-recorded and later transcribed. The transcripts were analysed by one researcher using thematic analysis (Braun & Clarke, 2006). In what follows, we describe the themes that emerged from the analysis.

Effort. Out of 31 participants, 16 emphasised the authentication scheme was complicated. Eight participants said it required mental effort with P2 saying it was *"too taxing on the mind"*. P31 explained: *"You do have to give full attention to it, so you can't be doing any other stuff. You can't be on the phone and be like: "Wait a minute I will just check my email. Oh, I have to login, hold for 5 minutes"*. As this quote suggests, participants also found the scheme time-consuming to use with overall 13 interviewees stressing it required more time than traditional password entry. P7 emphasised that authentication should be fast especially if it guards access to a critical task: *"If someone is going into cardiac arrest and those seconds matter, you need a procedure that's going to be as quick as possible, not something that's going to complicate things and, potentially, lose a life. I mean I realise logging on a system isn't life and death, but sometimes it is that crucial that you get in as quickly as possible."* Both effort and time needed contributed to the feeling of frustration in some participants. Six of them stressed they found the grid frustrating, P7 explained: *"It just seems like just the password itself is just far easier to remember and quicker to type in than having to find all these different letters and matching up and stuff. Like, if this was a real login for a site that I was on, and especially, you know, a business site or a professional site, it would drive me absolutely up the wall. I mean it would waste a lot of valuable time. So, like I don't understand the purpose of it actually. It's a big time waster."*

Out of 31 participants, 10 emphasised the learning curve for the practice of how to use the grid. They mentioned using the grid became easier with time which we also see in the quantitative data. Nevertheless, P31 stressed that using the grid will never be as easy as entering a password since it cannot be automated: "*Because you have to look*

at, I mean it's not something you can just memorise. If you just want to check email you memorise a password, or you check in some password, like I memorise the key strokes on my keyboard, that's fine. I hate logging in using tablets and stuff but you know like, it's usually like if I was checking email it's usually just yeah, the pass code and a mobile and everything is there. But there is only your computer that can memorise your keyboard strokes, so you are like whatever, it's muscle memory, you know your password. So that is like, oh no, you actually have to be awake and you have to do like puzzle type thing and then it's annoying to do on an iPad, like say if you've just woken up or are really tired, then you can't log in." The difficulty P31 described is that a login using the grid requires thinking and focus whereas standard password entry can be done out of muscle memory and the user does not have to concentrate as much.

High-value accounts. When asked if they would use this authentication scheme to log in in real life, participants generally gave varied responses. Out of 31 participants, 15 stressed it depended on the context of use. In ten cases, participants emphasised their decision whether to use a grid would depend on what kind of an account they were trying to protect. Participants mentioned that better security would be needed for systems that hold sensitive or confidential information or that could lead to money loss as with access to online banking. P18 explained they would be more likely to use the VAG *"for websites for banks and stock exchanges and all, where the real money is involved and your interest is at stake"*. However, participants were not unanimous on this, four of them stressed their banking was secure enough already, P31 emphasised: *"I'm fine with my bank account I think because I have a key and I have to press one number and then it gives you a number and I do that, so it's not like I have additional mental stress. So, I wouldn't use it for my bank. And anyway my bank is really good because I have had fraud twice and they return it, they block it when I have had fraud, so I am fine."* Nevertheless, P24 stressed they would prefer to use the grid since it does not require them to carry an additional device with them: *"Definitely this one is easier for my bank account than my token because I have to have my token everywhere I go. This way I can do it anywhere in the world, any platform, yeah. So in terms of easier sign-in, it's easier for bank account but in terms of security, I don't know."*

Experience of fraud. We also saw that prior experience of fraud influenced participants' stated willingness to adopt the VAG. P23 stressed: *"If you could choose, so I would probably want something like this for like my bank account or actually just recently my iTunes password was hacked, someone in Canada got my iTunes password so I had to close my account and cancel my credit card."* Conversely, P24 stressed they did not feel the need to strengthen their authentication through the use of a grid since they had not experienced fraud: *"I might feel more receptive to taking this up if I had ever been a victim of password fraud before, but I haven't. So, let's say if I'd had my account hacked into before I would probably be much more receptive to using this one."*

Frequency of use. Overall, nine participants made their willingness to adopt the VAG dependent on the frequency of use of the account it would protect. In four cases, participants stressed there was also a link between frequency of use and account importance. They said they logged in to important accounts like banking less often than to their email, thus they would be more willing to put the effort into logging in for something that is high-value and infrequent. P23 explained: "*Logging in to like my*

email, because I do it so frequently, I probably wouldn't want to go through the hassle of that. But if it's something more secure, I probably would." P25 expressed a similar view saying: *"I think that would take too long given how often I log into my email account."*

Security. Overall, ten participants stressed the authentication grid was more secure than passwords. Five of them stressed the complexity made it more secure, P26 explained it is difficult to use for the user, so it must be for the attacker too: *"It sounds like it's safer yes, because it's so complex, even for the user themselves."*

5 Discussion

We conducted a preliminary laboratory evaluation of the Vernitski Authentication Grid. A login took 88.6s on average. We found that the more practice our participants had, the faster they authenticated. Younger participants and those who had higher computer literacy authenticated faster too. We did not find a statistically significant relationship between experience with touchscreen devices and authentication speed on a tablet. Interestingly, we saw that participants authenticated faster on a tablet and subsequently faster on a PC. As mentioned earlier, previous research has shown that authentication on touchscreens poses many usability challenges and it is surprising to see that a login to the VAG was faster on a tablet than a PC.

We also saw that in theory this authentication scheme was meant to guard against shoulder-surfing but participant behaviour undermined this as participants pointed at the screen which might reveal to an attacker what the characters of their password are.

In terms of qualitative feedback, half of our participants thought the authentication grid was complex and over one third described it as time-consuming. Interestingly, when asked about their willingness to use it in the future, 19 participants made their decision dependent on the context where it would be used. They thought the grid could add extra security for systems holding important and sensitive information like banking and systems that they do not access as frequently. They emphasised the notion that for frequent accounts the password is in their muscle memory and they can enter it fast without much thinking. In such a situation, the use of the grid would not be suitable as it requires focus and time. Especially since activities like email or Facebook are quick, users log in for a few minutes just to check if there has been anything new they need to attend to. This is in line with findings from previous studies. Brostoff and Sasse (2000) found that if a login procedure was elaborate and taking longer than password entry, participants logged in less often and once they logged in, they worked on the system for longer than those who logged in just using passwords.

The differentiation that participants made in terms of account importance and frequency of use is very interesting. Passwords were invented for a certain purpose (administering a shared computer), then expanded to all systems as a one-size-fits-all solution and nowadays virtually any Website offering some service requires users to register with a username and password. Context of use is a fundamental HCI concept, however it is often forgotten by security researchers who do not account for differences

between individuals and contextual factors such as account type or frequency of use (Bonneau, Herley, Van Oorschot, & Stajano, 2012).

We also saw that the stated willingness to use the VAG depended on participants' risk perception. A participant who had experienced fraud stated they would be more likely to use it than a participant who had not. This is in line with previous studies where participants' exposure to cyber-threats made them more cautious in subsequent online interactions (Krol, Moroz, & Sasse, 2012).

Moreover, participants frequently compared using the VAG with entering traditional passwords. With an authentication time of 88.6s, the grid performed poorly in this comparison. To put this number into perspective, Roth, Richter and Freidinger (2004) tested several types of cognitive trapdoor games, that is PIN-entry methods offering resilience to shoulder-surfing. The longest average entry time for these was around 25s. To give another example, a standard login where the user has to enter a username and a password (both 8 characters long) was analytically predicted to take approx. 14.8s (Steves, et al., 2014). Additionally, our participants emphasised that using the grid required their undivided attention and they could not enter their password from muscle memory. It is a general problem with one-time credentials that their entry cannot be automated.

5.1 Limitations

The study was a preliminary laboratory evaluation of a new authentication mechanism. Such evaluations are multi-stage processes starting with a lab study, through a real-life deployment to an assessment post-adoption. Being the first stage in a long process our study had a range of limitations. We had a convenience sample of participants who knew what we were studying what might have made them behave in an unnatural way. Also, in real life a login is a gateway to some primary task and users' focus is not on security but on that primary task. This is something we could not recreate in this study since the explanation of the workings of the VAG needed to be quite elaborate and hiding the fact that our study was looking at the grid was not possible. Finally, due to our recruitment through a university participant pool, our study suffers from a volunteer bias and we have a sample of relatively young and well-educated individuals.

5.2 Future work

Future work could continue with further stages of a usability evaluation of the VAG. Participant responses showed us in what kind of situations and for what types of systems the grid could be used and any future evaluations could focus on testing its deployment in these real-life contexts.

6 Conclusions

As more and more services go online, reliable and efficient authentication will become even more important in the years to come. Our study shows that users are unwilling to use long and elaborate authentication procedures, such as the VAG studied here,

unless it is for infrequent and/or high-value accounts. In line with the fundamental security principle saying that a security measure should be proportional to the value of the assets it is protecting, we believe that the strength of authentication should be proportional to the importance/value of the accesses it is protecting. But the need for stronger authentication does not mean increasing the burden on users. The old myth that there is 'usability-security tradeoff' leads security experts to assume that it is OK for stronger security to require more effort. The myth even affects some users: in a recent study (Krol, Philippou, De Cristofaro, & Sasse, 2015), some of our participants consoled themselves that if the mechanism is demanding, it is secure. But in this current study, the majority of our participants were frank that they found the mechanism too demanding for regular authentication, and research to date has shown that authentication mechanisms that create too high a burden are circumvented, avoided or abandoned altogether by users (Steves, et al., 2014). Performing security tasks can give users the rewarding feeling that they have contributed to making their online interactions secure, but the effort has to be proportionate. The challenge is to be able to strike the right balance between providing users sufficient reassurance and demanding their attention, time and effort.

6.1 Recommendations

In light of the findings of our study, we suggest that grids like the VAG are too complex and time-consuming to use for frequent authentication. Having said that, there are specific contexts of use where the effort is seen to be proportionate. The VAG was faster on touchscreens, so it is more usable there. Most users struggle with infrequently used passwords, and there the use of a more memorable password can offset the longer input times. We also note that the scheme offers better security if used infrequently, since the attacker has to capture many authentication events to increase the likelihood of guessing the password.

There is no usable one-size-fits-all replacement for passwords – rather, mechanisms need to be selected to fit the devices, context of use (primary task, physical and social context), security requirements, and – where possible – preferences of individual users.

7 Acknowledgements

We would like to thank Brian Glass, Ingolf Becker and Granville Moore for their help in data analysis. Kat Krol's research was supported by an EPSRC grant to the UCL Security Science Doctoral Training Centre (SECReT) (grant number: EP/G037264/1).

8 References

Biddle, R., Chiasson, S., & Van Oorschot, P. C. (2012). Graphical passwords: Learning from the first twelve years. *ACM Computing Surveys (CSUR), 44*(4).

Bonneau, J., Herley, C., Van Oorschot, P. C., & Stajano, F. (2012). The quest to replace passwords: A framework for comparative evaluation of web authentication schemes. *IEEE Symposium on Security and Privacy (SP)*, (pp. 553-567).

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative research in psychology, 3*(2), 77-101.

Brostoff, S., & Sasse, M. A. (2000). Are Passfaces more usable than passwords? A field trial investigation. *People and Computers XIV — Usability or Else!*, pp. 405-424.

Brostoff, S., Inglesant, P., & Sasse, M. A. (2010). Evaluating the usability and security of a graphical one-time PIN system. *24th BCS Interaction Specialist Group Conference*, (pp. 88-97).

Ellison, C., Hall, C., Milbert, R., & Schneier, B. (2000). Protecting secret keys with personal entropy. *Future Generation Computer Systems, 16*(4), 311-318.

Florêncio, D., Herley, C., & Van Oorschot, P. C. (2014). Password portfolios and the finite-effort user: Sustainably managing large numbers of accounts. *Proc. USENIX Security*, (pp. 575-590).

Greene, K. K., Gallagher, M. A., Stanton, B. C., & Lee, P. Y. (2014). I Can't Type That! P@$$w0rd Entry on Mobile Devices. *HCI International 2014, Human Aspects of Information Security, Privacy, and Trust (HAS)* (pp. 160-171). Heraklion, Crete, Greece: Springer.

Krol, K., Moroz, M., & Sasse, M. A. (2012). Don't work. Can't work? Why it's time to rethink security warnings. *7th International Conference on Risk and Security of Internet and Systems (CRiSIS)*, (pp. 1-8). Cork, Ireland.

Krol, K., Philippou, E., De Cristofaro, E., & Sasse, M. A. (2015). "They brought in the horrible key ring thing!" Analysing the Usability of Two-Factor Authentication in UK Online Banking. *USEC 2015: NDSS Workshop on Usable Security.* San Diego, CA, USA.

Papanicolaou, C. (2013). *Novel Authentication Solution.* Department of Computer Science. London: University College London.

Roth, V., Richter, K., & Freidinger, R. (2004). A PIN-entry method resilient against shoulder surfing. *11th ACM Conference on Computer and Communications Security (CCS)* (pp. 236-245). Washington, DC, USA: ACM.

Schaub, F., Deyhle, R., & Weber, M. (2012). Password entry usability and shoulder surfing susceptibility on different smartphone platforms. *11th International Conference on Mobile and Ubiquitous Multimedia*, (p. 13).

Steves, M., Chisnell, D., Sasse, A., Krol, K., Theofanos, M., & Wald, H. (2014). *Report: Authentication Diary Study.* National Institute of Standards and Technology (NIST).

Vernitski, A. (2015). *Authentication Grid.* University of Essex, Technical Report. Retrieved from http://repository.essex.ac.uk/13231/

Yang, Y., Lindqvist, J., & Oulasvirta, A. (2014). Text Entry Method Affects Password Security. *Learning from Authoritative Security Experiment Results (LASER 2014)* (pp. 11-20). Arlington, VA, USA: USENIX Association.