Supporting Learning of an Unfamiliar Authentication Scheme

Alain Forget, Sonia Chiasson, Robert Biddle
School of Computer Science, Carleton University, Ottawa, Canada
aforget@scs.carleton.ca, chiasson@scs.carleton.ca, robert_biddle@carleton.ca

Abstract: Difficulties with plaintext passwords are well-documented. Many alternative authentication schemes have been proposed, but a key evaluation metric has typically been ignored: learnability. For wide deployment, some form of tutorial is the most assistance users would be provided. This paper presents the results of two user studies (one local and one on Amazon Mechanical Turk) of 134 total participants. Our studies compared four methods of teaching a new authentication scheme: a single page of instructions, a hypertext tutorial with images, an interactive demo, and a video tutorial. As one may expect, demo and video users invested more time in their tutorials than text and hypertext users. We found few differences in the learnability and security between the conditions, but to our surprise, the hypertext and video tutorials resulted in greater password memorability than the text and demo tutorials.

Introduction

Plaintext password systems have been the ubiquitous method of authentication for computer applications and online Internet services for many years. All the while, users have been coping with difficulties with plaintext passwords [1,7,16], such as a lack of user understanding and ability to create and remember secure passwords. In response, a multitude of alternative authentication schemes have been proposed [3,14]. Although there have been many developments in novel password schemes, including token-based and biometrics [14], the advantages of knowledge-based authentication schemes resembling passwords are still significant [13]. Novel authentication schemes are typically tested in experiments where users are given an in-person walkthrough by an expert [6] or asked to read a page of instructions [15,23]. However, if the scheme is widely deployed, traditional in-person training would impossible and ineffective [4]. Instead, some form of self-training material should accompany the deployed scheme.

Over one week, we tested four different tutorial formats with participants both local (i.e. in-person) and from Amazon Mechanical Turk[1]. They learnt to use the graphical password scheme Persuasive Cued Click-Points (PCCP) [6] to register with and log in to three websites to perform a task. We hypothesised that demo users would spend more time on their tutorial, but that demo users would spend less time registering and create more secure and memorable passwords, due to the richer interaction. We found only two clear differences between conditions. As expected, text and hypertext users spent less time on their respective tutorials than video and demo users. Curiously, hypertext and video users more easily recalled their passwords than text and demo users, despite demo being more engaging. Overall, users seemed to benefit most from our hypertext tutorial.

Authentication is a necessary process in order to protect users' assets. Given the difficulties users have with creating memorable and secure plaintext passwords, introducing a more secure novel authentication scheme is a likely possibility. Users will need a tutorial to teach them how to use the new scheme. This paper contributes to education of computer security by evaluating different tutorial methods, in the hopes of easing users' transition from plaintext passwords to a novel method of authentication. We compare tutorial methods along four criteria: Investment, Learnability, Security, and Memorability. We believe that high performance in all four areas is essential before deployment.

[1] A crowdsourcing website where users perform tasks remotely for compensation (https://www.mturk.com).

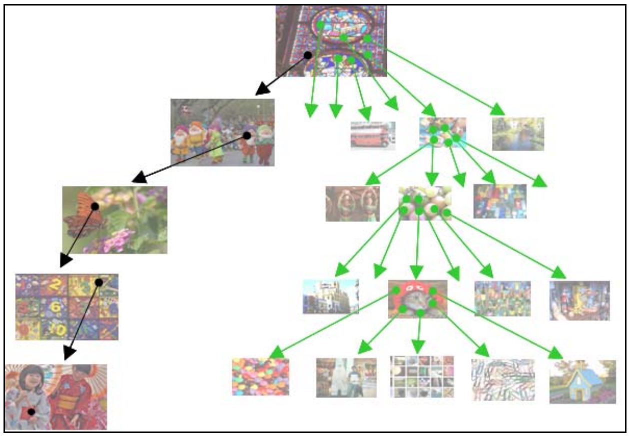


Figure 1. PCCP's image selection is based on users' previous click-points

Background

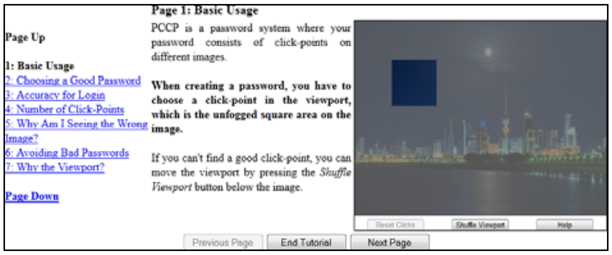
Adams and Sasse [1] documented the difficulties users face with plaintext password systems, They recommended that organisations "provide instruction and training on how to construct usable and secure passwords." Yet, creating and managing dozens of secure text passwords remains a challenge for users [7,22]. System administrators have been unsuccessful in teaching users to create and recall secure text passwords [10].

There are many published proposals to help users create more memorable and secure passwords, such as alternative authentication mechanisms. In particular, graphical passwords [3] leverage humans' superior memory for visual stimuli over text [17]. One such system claiming to be usable and secure is Persuasive Cued Click-Points (PCCP) [6], where a password is a sequence of click-points on different images (Figure 1). Users' previous click-point location determines the next image shown. During registration, users must select their click-points within a randomly-positioned persuasive viewport (Figure 2 and Figure 3), which helps users choose a more random (and secure) password. If users cannot find a desirable click-point within the current viewport location, they may *shuffle* to move the viewport to another random location. At login, users must click within an invisible tolerance region around their password's click-point. We chose PCCP over other schemes because it offers reasonable security and usability [3], but seems complicated and challenging for users to learn.

Any novel authentication scheme or enhancement should be accompanied by some kind of primer on how users can create and login with memorable and secure passwords. Barton and Barton [2] may have been the first to advocate for the availability of material on "user-friendly password methods", both before and while using the password system. They advise such material contain descriptions and illustrations for creating memorable passwords, as well as more technical information, such as minimum password length and other restrictions. Carroll et al. [4] advocate a minimalist approach to training users for new computing tasks. They suggest structuring the training material to facilitate users' desire to begin the task as soon as possible, present only material that is essential to performing the task, and support error recognition and recovery. Grossman et al. [12] present 25 learnability metrics used throughout the learnability literature. They also found that online videos helped users complete seven times more tasks as participants learning from text-based information [11].

Some of our tutorials' designs used Persuasive Technology (PT) [8]. PT is a framework for building technological solutions to support users in learning new behaviours. The framework identifies numerous principles that are based upon established psychological research on human behaviour and motivation. PT has been successfully used in numerous domains to assist people in learning new or improving existing behaviours.

To our knowledge, there is no published research on teaching users new authentication methods. However, the commercial authentication product Passfaces [18] includes a demo that emphasises the memorization of users' assigned faces, and walks users through the process of confirming and logging in.



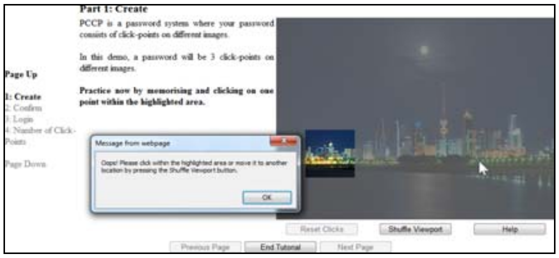
Page 1: Basic Usage

PCCP is a password system where your password consists of click-points on different images.

When creating a password, you have to choose a click-point in the viewport, which is the unfogged square area on the image.

If you can't find a good click-point, you can move the viewport by pressing the *Shuffle Viewport* button below the image.

Figure 2. Hypertext Tutorial



Part 1: Create

PCCP is a password system where your password consists of click-points on different images.

In this demo, a password will be 3 click-points on different images.

Practice now by memorising and clicking on one point within the highlighted area.

Oops! Please click within the highlighted area or move it to another location by pressing the Shuffle Viewport button.

Figure 3. Demo Tutorial

Tutorials

We drew upon several sources when designing the tutorials. The Minimal Manual [4] principles ensured the tutorials contained only the minimum material users would need to create passwords that were both memorable and secure. For example, users were informed that the first half of the tutorials contained material essential to use PCCP, while supplementary material was placed in the second half, which could be skipped if desired. Persuasive Technology (PT) [8] tools advised us how to deliver suggestions at the most opportune moments and to tunnel users through the tutorial. Finally, the tutorials' sections were "dynalinked" [19] to make it easier for users to move through them to form correct mental models of the password system's usage.

The content of the tutorials was based on our several years' experience tutoring participants in using novel authentication schemes. We attempted to make the tutorials as effective learning tools as we could within the constraints of the modality. For example, we tried to make the tutorials as easy to navigate as possible, whether participants were reading through the tutorial sequentially, or browsing for particular information. Users were always shown one of the tutorials before creating a password, and could navigate the tutorials freely, either by clicking on a particular section or viewing them in the prescribed order. The tutorial could be ended at any time. While using PCCP, users could press a *Help* button to open the tutorial in a new window alongside the password system. The tutorial material was divided into the following sections:

1. *Basic Usage*: PCCP, the persuasive viewport, and the shuffle button were introduced.
2. *Choosing a Good Password*: Users were encouraged to choose click-points that were "easy to click on precisely, but don't stand out very much", and were given an example.
3. Accuracy for Login: The tolerance region was described and illustrated with an example and image.
4. *Number of Click-Points*: Users were told their password would consist of 4 click-points. The tutorial also mentioned that users now knew all they needed to use PCCP. Although users could have exited the tutorial at any time, they were now explicitly told that they could either start using PCCP right away, or continue the tutorial.
5. *Why Am I Seeing the Wrong Image?*: This section discussed PCCP's implicit feedback property, whereby each image shown to the user is chosen based on their previous click-point. Thus, when logging in, if the user sees an image they do not recognise, they know immediately that they made a mistake and should re-enter their password.
6. *Avoiding Bad Passwords*: The tutorial explained that objects that are either very obvious or difficult to click on precisely are poor click-points selections, because they are either insecure or hard to remember. An example of each was given, and users were advised to shuffle the viewport when only poor click-points were in the viewport's current position.
7. *Why the Viewport?*: The tutorial described how the randomly-positioned persuasive viewport helps users select more random (i.e. more secure) click-points.

Based on this material, we designed four different tutorial types:

Text. This was the simplest tutorial; one page of instructions with an image. This low-tech tutorial omitted sections 5, 6, and 7 (see above), which were not absolutely necessary in order to get started with PCCP.

Hypertext. We split the tutorial material into seven sections of instructions and images (Figure 2). The material was easier to parse with links and buttons to move between page-like sections, each of which included an illustration to support the accompanying text. This condition is a richer use of web technology than the text tutorial.

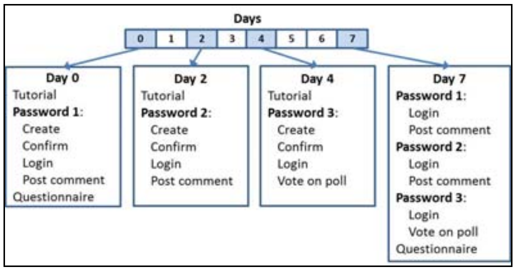


Figure 4. Study Timeline

	Local	MTurk
Text	12	15
Hypertext	9	24
Demo	10	20
Video	19	24
Male	14	62
Female	27	31
Age (min, median, max)	17, 20, 55	18, 26, 50

Table 1. Participant Demographics

Demo. We constructed an interactive demo with which users could practice creating, confirming, and logging in with a three-click password, supplemented with textual guidance (Figure 3). Links and buttons similar to the hypertext condition aided navigation through the tutorial. Users received immediate feedback whenever they interacted with the demo. We leveraged PT principles [8] in designing this tutorial, since PT has been successfully used to educate users in other domains. Principles [9] used in the demo include:

- *Tunnelling*: The tutorial stepped through creating, confirming, and logging in with PCCP.
- *Suggestion*: The tutorial delivered key advice at opportune moments, such as suggesting secure but memorable click-points when creating a demo password.
- *Self-monitoring*: Immediate feedback on users' progress was provided after each click-point selection.
- *Conditioning*: Users was informed of errors, and the tutorial advised how to avoid future mistakes.

Video. We recorded and uploaded a five-minute video of screen-captured interactions with PCCP, with narrated instructions and advice. Users could pause, play, and seek as with most online streaming video. The video began with an index of the sections with timestamps, providing an overview of the content and quick access to desired information. Numerous benefits have been ascribed to streaming media tutorials [21], which take advantage of humans' cognitive abilities described by dual coding theory, the modality effect, and others. Some research [11] suggests video may better assist users with unfamiliar tasks than text-based help.

Study Methodology and Hypotheses

To compare the four types of tutorials, we performed two between-subjects user studies[2]; one locally and one online using Amazon Mechanical Turk (MTurk). Our study used live websites with a custom authentication system for user testing purposes [5]. These blog-style websites allowed users to perform typical online tasks, such as commenting and voting on posts.

Participants were told the study was about overall website usability (including registering and logging in), in order to focus their goal on performing the specific website-related tasks, and make the password-related activities a secondary task. Over the course of one week, participants performed the following tasks (Figure 4):

Day 0. Users created their first password and account on the first website. They logged in, performed a task, and answered a questionnaire. MTurk participants were e-mailed the experiment's instructions, while local participants visited our lab. To minimise experimenter influence, local participants reviewed the tutorial and created their account alone, and were provided no information beyond the tutorial.

Day 2. Users were requested to visit a second website to create an account, login, and perform a task.

Day 4. Users were asked to create an account, login, and perform a task on a third website.

Day 7. Local participants returned to the lab to log in to the three websites, perform a task on each, and complete a final questionnaire. MTurk participants were e-mailed the same instructions.

For tasks to be completed at home (days 2 and 4 for local participants and all days for MTurk participants), users were e-mailed the instructions. We label the start of the study as day 0 (rather than day 1) since conceptually "day *n*" is *n* days since beginning the study.

A total of 134 people participated in our study. Table 1 summarises the demographics for local and MTurk participants. The numbers of participants in each condition are slightly unbalanced due to unpredictable drop-outs. Participants in both groups rated their computer skills as quite strong (a median of 8), on a scale of 1 (novice) to 10 (expert). All local and ~40% of MTurk participants were university students from various

[2] Our user study was approved by the Carleton University Ethics Committee for Psychological Research.

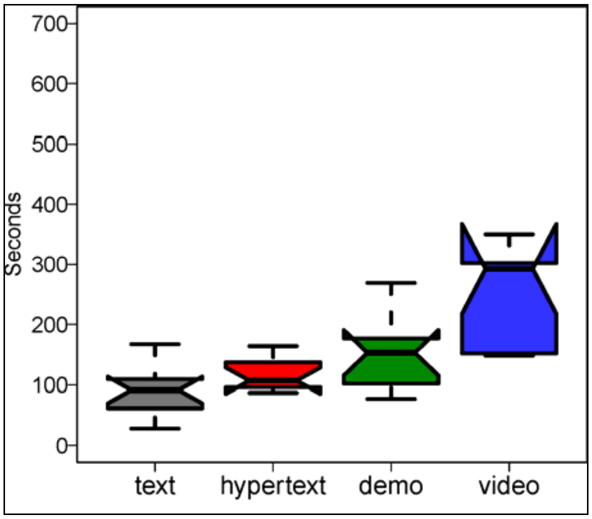


Figure 5. Time local users spent on the tutorial

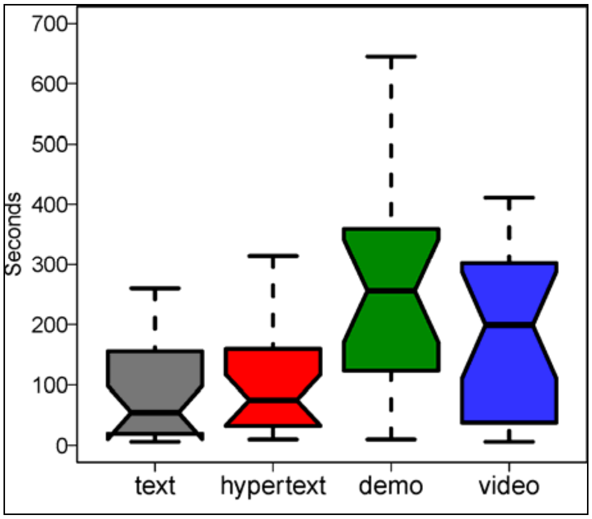


Figure 6. Time MTurk users spent on the tutorial

disciplines. The non-student MTurkers' occupations varied, from unemployed to a company's "Vice President – Materials". No participants reported studying or working in computer security.

We initially thought that the demo tutorial would be most successful, since people generally learn better with greater engagement [8]. Using Grossman et al.'s [12] categories of learnability metrics, we used their task and documentation metrics to formulate our hypotheses along four pillars:

- *Investment*: Demo users will spend more time on the tutorial than users in any other condition.
- *Learnability*: Demo users will spend less time registering passwords than users in other conditions.
- *Security*: Demo users will shuffle less than users in any other condition.
- *Memorability*: Demo users will successfully login more often than users in any other condition.

Results

We analyse four study dimensions: Investment measures the time invested in looking at tutorial materials. Learnability evaluates the effectiveness of the tutorials at imparting their material to users. Security evaluates password strength. Memorability examines how easily users could recall their passwords.

We use a variety of statistics to determine if two distributions are significantly different. For normal distributions, we use one-way ANOVAs and t-tests. When the data is not normally distributed, we use Kruskal-Wallis (KW) and Mann-Whitney U tests instead. For categorical data, we primarily use chi-squared or Fisher's Exact tests depending on category counts. The Bonferroni correction is applied as appropriate when doing multiple comparisons. In all tests, we accept $p < .05$ as statistically significant.

We analyse the local and MTurk data in parallel rather than together because factors such as Internet speed, language or cultural barriers, or other unknowns may have affected results in ways that our experimental design was not intended to measure. We informally identify differences between the two studies where appropriate, but direct comparisons would require further experiments specifically measuring such differences.

Investment. *Investment* is measured by the time users spent reviewing the tutorial and by user perception of the tutorial. The less time users need to spend on the tutorial, the sooner they can register and resume their primary tasks. All other factors being equal, we feel that a shorter time invested into a tutorial would be a positive result. Furthermore, the time spent viewing the tutorial before and during password creation are equivalent to Grossman et al.'s [12] documentation metrics D2 (time taken to review documentation until starting a task) and D1 (help commands used over certain time interval) respectively.

Tutorial Times. We examined the time users spent with the tutorial. Users spent the most time with the tutorial when first seeing it, before ever using PCCP, but seldom return to the tutorial thereafter for the duration of the study. The boxplots in Figure 5 and Figure 6 illustrate the times each condition's participants initially spent on their tutorial. Boxplots show the median at the centre, the box showing the central quartiles, the whiskers showing the outer quartiles, and the notch showing the 95% confidence interval. Table 2 shows that Kruskal-Wallis tests found differences between conditions. Six Bonferroni-corrected Mann-Whitney post-hoc

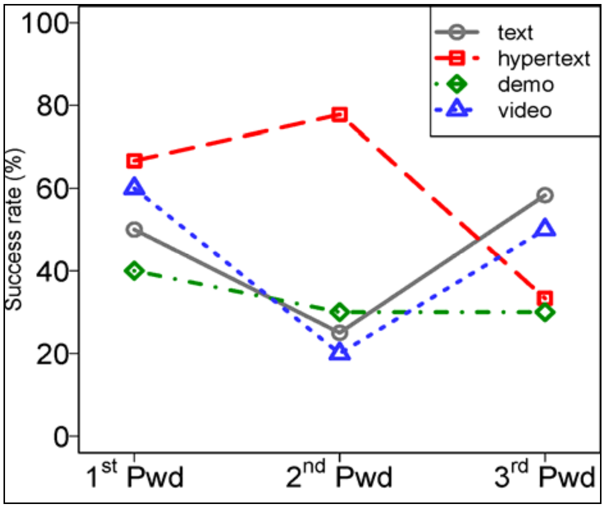


Figure 7. In-person day 7 login success rates.

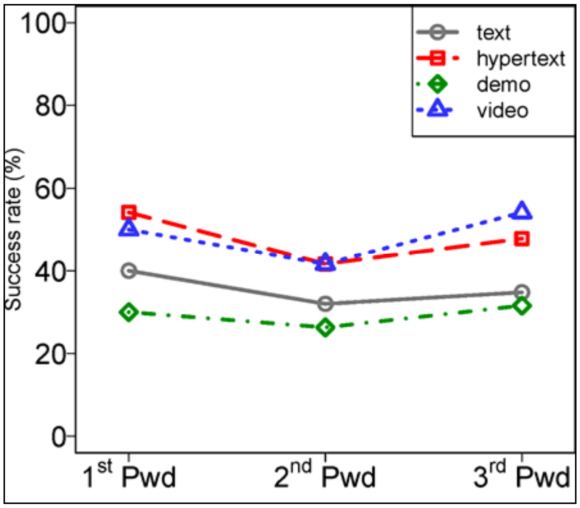


Figure 8. MTurk day 7 login success rates.

tests reported that video participants initially spent the longest on the tutorial in the local study, while demo participants spent the longest in the MTurk study[3].

User Perception of the Tutorial. On day 0, users completed a questionnaire after registering their first password. Among other topics, the questionnaire included 10-point Likert scale questions about the tutorial, where 1 represents *strongly disagree* and 10 represents *strongly agree*. All user perception Likert-scale topics had both a positively- and negatively-phrased question for counter-balancing. However, in the analysis, we eliminated questions whose results showed evidence that users may not have understood the wording.

No significant differences were found between conditions regarding how easy participants found their tutorial to understand or how quickly they felt they could complete the tutorial. This suggests that participants in any one condition felt their tutorial was no harder to understand nor slower to navigate than participants in any other condition. Local responses show high variance. This is possibly a result of a small sample size, since the largest local condition (text) only had 12 participants. MTurk responses appeared more consistent in rating their respective tutorials very easy to understand and quick to navigate across all conditions.

Learnability. We define *learnability* as the tutorial's effectiveness in enabling users to create and login with the password system efficiently. If participants understood the tutorial and learnt to use the scheme, then they should be able to quickly create and login with few to no errors. Thus, we measure learnability by the time taken to successfully register and log in to a new account and by the number of times users restarted the registration process. Respectively, these measurements are the same as Grossman et al.'s [12] task metrics T5 (time until user completes a certain task successfully) and T4 (task errors made over a certain time interval). We found no significant differences across conditions in registration time or number of registration restarts.

Security. The tutorials included advice on how to choose a secure PCCP password. We define the *security* dimension as the randomness in users' chosen passwords. In PCCP, the less users shuffle the persuasive viewport, the more random their password (see Background). PCCP is designed to more evenly distribute users' click-points across the image. Thus, we measure security by the number of shuffles as well as users' perceptions of whether the tutorial material helped them choose a more secure password. Shuffles could be considered analogous to Grossman et al.'s [12] task metric T7 (quality of work performed during a task).

There were no significant differences in shuffling between passwords or conditions for local or MTurk participants. There was a large variance in the number of shuffles, but on average users shuffled 10 times per image. This shuffling rate per image is similar to those in the online study by Chiasson et al. [6]. We are less concerned that our shuffling rates seem higher than Chiasson et al.'s in-lab studies [6], since it is possible their participants may have been unintentionally influenced to behave more securely by the laboratory setting.

[3] Statistically significant differences were found in all pairs except demo-video.

Users	Conditions	Test Results
Local	All	KW: $X^2(3) = 21.37$, $p < .0001$
	Video vs Text	U = 117, p < .0005
	Video vs Hypertext	U = 86, p < .005
	Video vs Demo	U = 85, p < .05
MTurk	All	KW: $X^2(3) = 14.93$, $p < .005$
	Demo vs Text	U = 361, p < .01
	Demo vs Hypertext	U = 338.5, p < .05

Table 2. Differences between conditions on time initially spent on the tutorial during the first visit. Only statistically significant results are reported.

Users	Condition	Attempted logins	Immediate resets
Local	Text	8 (67%)	4 (33%)
	Hypertext	9 (100%)	0 (0%)
	Demo	4 (40%)	6 (60%)
	Video	9 (90%)	1 (10%)
MTurk	Text	13 (81%)	3 (19%)
	Hypertext	18 (86%)	3 (14%)
	Demo	8 (73%)	3 (27%)
	Video	17 (85%)	3 (15%)

Table 3. Number of day 7 attempted logins and immediate resets.

Memorability. We evaluated *memorability* by looking at whether users were able to remember their passwords on the last day of the study (day 7). Thus, we measure memorability by the proportion of users who successfully logged in at the study's end. Successful logins are closest to Grossman et al.'s [12] task metric T2 (percentage of users who complete a task without any help), considering password reset requests as ``help".

Success Rates. Figure 7 and Figure 8 show the login success rates on day 7. Local participants returned to the lab for this final login to each website while MTurk participants received an email requesting that they log in to each site. We considered the login successful if the participant was able to login on day 7 without having to reset their password registered on day 0. We found no significant differences in success rates across conditions. We speculate that the local study's small sample size accounts for its higher variance.

End-of-Study No-Shows and Immediate Resets. The success rates in Figure 7 and Figure 8 exclude 22.5% of MTurk participants who did not return to complete day 7's tasks (*no-shows*). Table 4 shows the distribution of MTurk no-shows across the different conditions. We see significantly more no-shows in the text and demo conditions ($X^2(3) = 8.20, p < .05$).

We further examined the number of local and MTurk participants who chose to reset their password immediately on day 7, without first attempting to login. Table 3 suggests that, as with the no-shows, text and demo participants immediately reset their password more frequently than those in other conditions. However, this difference is only significant for local participants (*Fisher, p < .05*), not MTurk ($X^2(3) = 7.44, p = .06$).

We believe that the most likely explanation is that these users had forgotten their passwords, and decided not to try at all. This may provide evidence that text and demo participants' created less memorable passwords than hypertext and video participants.

User Perception of Memorability. Users were asked to rate on how much they felt their tutorial helped them create more memorable passwords. Statistical tests show no significant differences in responses between conditions. Responses from local participants were more neutral while MTurkers overall felt that their tutorial was helpful in creating more memorable passwords. These questions were posed at the end of day 0, and reflect users' initial perception of memorability.

Users	Condition	Returned	No-shows
MTurk	Text	16 (64%)	9 (36%)
	Hypertext	21 (87%)	3 (13%)
	Demo	11 (55%)	9 (45%)
	Video	20 (83%)	4 (17%)

Table 4. Number of day 7 MTurk no-shows. All local participants returned on day 7.

	Text	Hyper	Demo	Video
Investment	2.5	2.5	1.0	0.0
Learnability	1.5	1.5	1.5	1.5
Security	1.5	1.5	1.5	1.5
Memorability	0.5	2.5	0.5	2.5
Total	**6.0**	**8.0**	**4.5**	**5.5**

Table 5. Comparative scores for each condition along our evaluation dimensions. Scores range from 0 (worst) to 3 (best).

Interpretation

We revisit our hypotheses in light of the results of the studies, which demonstrate that the demo tutorial did not have the effect we predicted. We propose a simple framework based on our four evaluation dimensions and rank the tutorials accordingly. Finally, we discuss differences between the two studies.

Hypothesis Testing. Our hypotheses generally favoured the demo tutorial to better educate users about PCCP, at the cost of requiring more time to complete the tutorial. We will now review each hypothesis.

Investment. *Demo participants will spend more time on the tutorial than participants in any other condition.* We found little support for this hypothesis since it holds true for only the MTurk text and hypertext conditions. Local video participants spent significantly more time than demo participants with the tutorial when first seeing it. MTurk demo participants spent significantly more time initially on the tutorial than text and hypertext, but not video.

Learnability. *Demo participants will spend less time registering their password than participants in any other condition.* We found no support for this hypothesis because there were no significant differences in registration times between conditions in either study.

Security. Demo participants will shuffle less than participants in any other condition. We found no support for this hypothesis, as there were no significant differences in the number of shuffles between conditions for either study.

Memorability. *Demo participants will have higher login success rates than participants in any other condition.* We found no support for this hypothesis because demo participants avoided re-entering their password because they had forgotten them more often than hypertext and video participants. We also found no significant differences in success rates between conditions in either study. Contrary to our expectations, the richer interaction and immediate feedback provided by the demo tutorial showed no greater benefit over the other tutorials.

Evaluation. Overall, participants were able to learn how to register and login with PCCP with any one of the tutorials. Although no tutorial was consistently more effective than the others, we wished to derive some insights on which modality may be most promising to pursue in future work. Based on our hypotheses' topics, we ranked the tutorials along the four dimensions of Investment, Learnability, Security, and Memorability. Table 5 shows relative scores we gave to each condition for each dimension, ranging from 0 (worst) to 3 (best). When no differences were found between conditions, we divided the points evenly.

Investment. The video tutorial received the lowest score (0) because local video participants spent the most time on their tutorial compared to participants in any other local condition. Demo received the second lowest score (1) because MTurk demo users spend more time on the tutorial than text or hypertext participants. We found no differences between text and hypertext, so we awarded each condition the average of the top two scores (2 & 3 = 2.5).

Learnability. All conditions were given the average of all possible scores (0, 1, 2, & 3 = 1.5), as no differences in learnability were found between the participants in different tutorial conditions.

Security. We awarded all conditions the average of all possible scores (0, 1, 2, & 3 = 1.5), because we found no difference in the security of the passwords chosen by users in different conditions.

Memorability. We averaged the two lowest scores (0 & 1 = 0.5) for the text and demo tutorial, since participants in those conditions more often avoided attempting to login at the end of the study. Hypertext and video shared the two highest scores (2 & 3 = 2.5), as there were no differences in success rates or user opinion.

Summing the scores for each condition in Table 5, the hypertext tutorial ranks the best overall modality of the four we tested, as hypertext always tied for the highest possible score for each dimension.

We emphasise that this is only an informal ranking and that our study results found few statistically significant differences between conditions. However, Stobert and Biddle [20] ran several experiments where over 300 participants successfully learnt to use an novel authentication scheme with nothing more than tutorial similar to our text tutorials. Together, our studies suggest that simpler text or hypertext tutorials are sufficient for teaching users novel authentication schemes.

Conclusions

Computers and networks facilitate learning in many different areas. One important area for everyone to learn is to use these systems safely. As computing technology advances, new techniques need to be learned at home, the office, or school. This paper addresses one such technique: learning to use new password schemes.

Before creating a password, our participants were presented with one of four tutorials; text instructions with an image, smaller hyperlinked sections of text with several images, an interactive demo with immediate feedback, or a narrated screen-captured video of interactions with the password scheme. Over one week, users were asked to create and recall three different passwords to perform primary tasks on three websites. Among the alternative authentication schemes, we chose Persuasive Cued Click-Points (PCCP) for our study since it seems to offer reasonable security and usability, but is not straightforward to learn.

We hypothesised that the demo tutorial would help users quickly create more memorable and secure passwords than the other conditions, at the cost of more time spent on the tutorial. Although MTurk demo users did spend the most time on the demo tutorial, we did not find evidence to support our four hypotheses.

Users were able to independently learn to create passwords and login to their accounts. In fact, users surprisingly did not need an interactive demo; a hypertext or video tutorial was sufficient, and even preferred. This result was confirmed in work performed in parallel by Stobert and Biddle [20]. This is good news for authentication researchers and developers, since they may not need to spend extra time and effort building elaborate demos to get users started with new password systems. For researchers, assessing the learnability and user investment in their novel authentication schemes should be part of the design and proposal.

This paper presents the first research work on teaching people to use unfamiliar authentication schemes. Our user study demonstrates that users can learn to authenticate with a novel scheme with a tutorial, preferably a simple text or hypertext tutorial. An important area for future work is testing these tutorial modalities with multiple authentication schemes. Different tutorial modalities or interaction methods may also be worth studying. One example could be a differently designed demo that uses more passive forms of feedback instead of dialog boxes. It may also be interesting to compare the performance of novel scheme users who are not given any tutorial, since an authentication system that is intuitive and self-explanatory may be ideal. Finally, a common framework or scorecard for experiment authentication system evaluation would assist researchers and professional in comparing different schemes. We recommend such a framework include dimensions representing investment, learnability, security, and memorability, since high performance in all four is essential before a scheme be deployed.

References

[1] A. Adams and M. Sasse. Users are not the enemy. Communications of the ACM, 42(12):41–46, 1999.

[2] B. Barton and M. Barton. User-friendly password methods for computer-mediated information systems. Computers & Security, 3(3):186–195, 1984.

[3] R. Biddle, S. Chiasson, and P.C. van Oorschot. Graphical passwords: Learning from the first twelve years. ACM Computing Surveys, 44(4), 2012. in press.

[4] J. Carroll, P. Smith-Kerker, J. Ford, and S. Mazur-Rimetz. The minimal manual. Human-Computer Interaction, 3(2):123–153, 1987.

[5] S. Chiasson, C. Deschamps, E. Stobert, M. Hlywa, B. F. Machado, A. Forget, N. Wright, G. Chan, and R. Biddle. The MVP web-based authentication framework. Financial Cryptography (FC). Springer, 2012.

[6] S. Chiasson, E. Stobert, A. Forget, R. Biddle, and P.C. van Oorschot. Persuasive Cued Click-Points: design, implementation, and evaluation of a knowledge-based authentication mechanism. IEEE Transactions on Dependable and Secure Computing (TDSC), March/April 2012.

[7] D. Florêncio and C. Herley. A large-scale study of WWW password habits. World Wide Web Conference (WWW). ACM, May 2007.

[8] B. Fogg. Persuasive Technologies: Using Computers to Change What We Think and Do. Morgan Kaufmann Publishers, 2003.

[9] A. Forget, S. Chiasson, R. Biddle, and P.C. van Oorschot. Persuasion as education for computer security. AACE E-Learn, 2007.

[10] S. Furnell. An assessment of website password practices. Computers & Security, 26(7-8):445–451, 2007.

[11] T. Grossman and G. Fitzmaurice. Toolclips: an investigation of contextual video assistance for functionality understanding. CHI. ACM, 2010.

[12] T. Grossman, G. Fitzmaurice, and R. Attar. A survey of software learnability: metrics, methodologies and guidelines. CHI. ACM, 2009.

[13] C. Herley and P.C. van Oorschot. A research agenda acknowledging the persistence of passwords. IEEE Security & Privacy, January-February 2012.

[14] A. Jain, L. Hong, and S. Pankanti. Biometric identification. Communication of the ACM, 43(2):91–98, February 2000.

[15] C. Kuo, S. Romanosky, and L. Cranor. Human selection of mnemonic phrase-based passwords. SOUPS. ACM, July 2006.

[16] R. Morris and K. Thompson. Password security: A case history. Communications of the ACM, 22:594–597, 1979.

[17] A. Paivio and T.B. Rogers and P.C. Smythe. Why are pictures easier to recall than words? Psychonomic Science, 11(4), 1968.

[18] Passfaces Corporation. Passfaces online demo, December 2011.

[19] Y. Rogers and M. Scaife. How can interactive multimedia facilitate learning? In J. Lee, editor, Intelligence and Multimodality in Multimedia Interfaces: Research and Applications. AAAI Press, 1998.

[20] E. Stobert and R. Biddle. Memorability and usability of graphical password forms: A graphical password bake-off. Annual Computer Security Applications Conference (ACSAC). IEEE, December 2012, in submission

[21] N. Tempelman-Kluit. Multimedia learning theories and online instruction. College & Research Libraries, 67:364–369, July 2006.

[22] M. Weir, S. Aggarwal, M. Collins, and H. Stern. Testing metrics for password creation policies by attacking large sets of revealed passwords. Conference on Computer and Communications Security (CCS). ACM, 2010.

[23] J. Yan, A. Blackwell, R. Anderson, and A. Grant. Password memorability and security: Empirical results. Security & Privacy, 2(5):25–31, 2004.

Acknowledgements

This work was supported by the Natural Science and Engineering Research Council of Canada (NSERC), as well as partial funding from the NSERC Internetworked Systems Security Network (ISSNet). The second author acknowledges NSERC funding for her Canada Research Chair in Human-Oriented Computer Security.