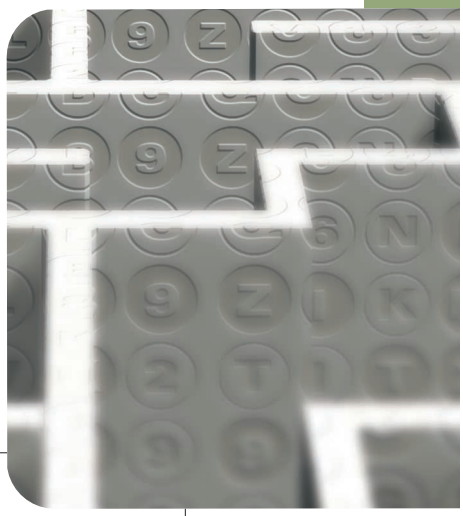
Password Memorability and Security: Empirical Results

Users rarely choose passwords that are both hard to guess and easy to remember. To determine how to help users choose good passwords, the authors performed a controlled trial of the effects of giving users different kinds of advice. Some of their results challenge the established wisdom.



Jeff Yan
Chinese University of Hong Kong

Alan Blackwell, Ross Anderson, and Alasdair Grant
Cambridge University

Many of the deficiencies of password authentication systems arise from human memory limitations. If humans didn't have to remember their passwords, a maximally secure password would have maximum entropy: it would consist of a string as long as the system allows with characters selected from those the system permits in a manner that provides no redundancy—that is, totally random selection. These requirements run contrary to the properties of human memory, however. Human memory for sequences of items is temporally limited,[1] with a short-term capacity of around seven plus or minus two items.[2] In addition, when humans do remember a sequence of items, those items must be familiar chunks such as words or familiar symbols.[2] Finally, human memory thrives on redundancy—we're much better at remembering information we can encode in multiple ways.[3]

Password authentication therefore involves a trade-off. Some passwords are easy to remember (for example, single words in a user's native language), but also easy to guess through dictionary searches. Other passwords are secure against guessing but difficult to remember. In this case, human limitations can compromise the password's security because the user might keep an insecure written record of it or resort to insecure backup authentication procedures after forgetting it. This doesn't mean we accept the common doctrine that writing passwords down is always wrong. If your machine isn't in a publicly accessible area, writing down a long, random boot password and taping it to the machine can be worthwhile because you can then have a strict policy against disclosing passwords over the phone. Preventing social engineering attacks is a separate research topic, however.

We conducted an empirical investigation of this trade-off in a population of password users. Research in cognitive psychology has defined many limits of human performance in laboratory settings where experimental subjects are required to memorize random and pseudorandom sequences of symbols. It's difficult to generalize from such research to password users, who can select the string, rehearse it while memorizing, and must recall it at regular intervals over a long period of time. We show that in this context, users can exploit mnemonic strategies for password memorization. Humans can use many mnemonic techniques to successfully memorize apparently random sequences. Password alternatives such as *Passfaces* exploit superior human memory for faces, for example.[4] Rather than change the password authentication procedure, however, we suggest changing the advice given to users for selecting passwords.

Password selection advice

Anne Adams and Martina Angela Sasse note that users aren't enemies of security, but collaborators who need appropriate information to help maintain system security.[5] They observe that users, when not told how to choose good passwords, make up rules for password generation, resulting in insecure passwords. They therefore recommend that organizations "provide instruction and training on how to construct usable and secure passwords."

Later research by Sasse, Sacha Brostoff, and Dirk Weirich based on a survey of system users found that 90 percent of them had difficulty with standard password mechanisms and that they welcomed advice on password

Password Experiment—Group A

This sheet offers some advice on how to choose a good computer password. We are giving you this sheet as part of the password security experiment that was described in your introductory lecture. Different people are receiving different advice (but all advice should result in passwords at least as secure as you would choose if not participating in the experiment). Please do not discuss the experiment, this advice, or your choice of password with your friends.

Please log on using the initial password you have been issued, and choose a new password not known to anybody else. The "Windows NT Tutor" tells you how to do this on pages 1.6–1.7.

Your password should be at least seven characters long and contain at least one nonletter.

If you have already changed your initial password to one of your choice, and your new password meets this standard, then you do not need to change it again. However we strongly recommend that you change your password from time to time—at least once a term. As the experiment will run for the duration of this academic year, please keep this sheet and use this advice again when you choose your new passwords for Lent and Easter.

Figure 1. Control group instruction sheet. We asked participants to choose a seven-character password with at least one nonletter.

generation. (See the "Related work on password selection and memorability" sidebar.) The authors conclude that, "instructions for constructing and memorizing a strong password ... should be available when a password needs to be chosen or changed."

Many large organizations do give specific advice to new users about selecting good passwords. For example, a password should be reasonably long, use a reasonably large character set, and still be easy to remember. There are some subtleties about whether the attacker will try many passwords over a network or has obtained a copy of the password file and is cracking it offline, but we propose to ignore them in the present study.

We informally surveyed advice given to new users at large sites by searching the Web for the terms "choose," "good," and "password." Many sites didn't recognize memorability's importance; rather, they merely emphasized resistance to brute-force search. Some typical pieces of advice were:

> [A good] password should consist of mixed characters or special characters, and should not consist of words found in the dictionary. It should not be written down in an easily accessible place and especially not next to login. It may either be all in capital or small type letters.

> Use the output from a random password generator. Select a random string that can be pronounced and is easy to remember. For example, the random string "adazac" can be pronounced a-da-zac, and you can remember it by thinking of it as "A-to-Z." Add uppercase letters to create your own emphasis—for example, aDAzac.2.

> Good passwords appear to be random characters. The wider the variety of characters the better. Mixing letters with numbers is better than letters alone. Mixing special characters with number and letters is better still.

One recommendation that seems increasingly popular is the pass phrase approach to password generation. A typical description of this is as follows:

> A good technique for choosing a password is to use the first letters of a phrase. However, don't pick a well known phrase like "An apple a day keeps the doctor away" (Aaadktda). Instead, pick something like "My dog's first name is Rex" (MdfniR) or "My sister Peg is 24 years old" (MsPi24yo).

Of course, our informal survey doesn't include sites that don't give users advice on password selection. We believe that many simply tell new users the minimum requirement for a valid password (length and character set), with no further advice regarding security or memorability. Others, in our experience, enforce rules such as

> Passwords must be at least eight characters long and must contain at least two nonletter characters. They must also be changed at least once a month.

Users typically respond to such rules with a personal password generation system, such as "Juliet03" for March, "Juliet04" for April, and so on. This system is clearly weak. Other attempts to compel user behavior have backfired. For example, Bill Patterson reports that when users were forced to change their passwords and prevented from using their previous few choices, they changed passwords rapidly to exhaust the history list and get back to their favorite password.[6] Forbidding password changes until after 15 days prevented users from changing possibly compromised passwords without the system administrator's help.[6] If the restrictions placed on users are tedious, users will likely circumvent them.[5]

Designing advice to users and enforcing it on a system level are important problems involving subtle questions of applied psychology with no obvious answers.

Experimental study

To investigate the trade-off between security and memorability in a real-world context, we compared the ef-

fects of giving three alternative forms of password selection advice to different user groups. The experimental subjects were first-year students in the University of Cambridge's School of Natural Sciences, which includes physics, chemistry, geology, and materials science. All Natural Sciences students are given an account on a central computing facility with a user ID and a randomly generated initial passwords. They also have access to several other facilities. When students receive their account details, they are generally advised to select their own passwords. Some students receive this advice informally from their department or residence hall computer officer. Many students attend a central facility introductory lecture and tutorial session.

Methodology

We asked students attending the introductory lecture to participate in an experiment on password selection. Of the 300 students at the lecture, 288 consented to participate in the experiment. At the tutorial session, we randomly assigned students to one of three experimental groups and gave each group a sheet of advice. (Figures 1–3 show the text we used.)

- We gave the *control group* (95 members) traditional advice—that is, "Your password should be at least seven characters long and contain at least one nonletter." (See Figure 1.)
- We gave the *random password group* (96 members) a sheet of paper with the letters A–Z and the numbers 1–9 printed repeatedly on it. We told them to select a password by closing their eyes and randomly picking eight characters. We advised them to keep a written record with them until they'd memorized the password. (See Figure 2.)
- We told the *pass phrase group* (97 members) to choose a password based on a mnemonic phrase. (See Figure 3.)

We expected that the random password group would have stronger passwords than the pass phrase group but would find them harder to remember and easier to forget; and that the pass phrase group would stand in the same relation to the control group. One month after the tutorial sessions, we took a snapshot of all password files and conducted four types of attack on the passwords:

- *Dictionary attack*: Simply use different dictionary files to crack passwords.
- *Permutation of words and numbers*: For each word from a dictionary file, permute with 0, 1, 2 and 3 digit(s) to construct possible password candidates. Also, make common number substitutions, such a 1 for I, 5 for S, and so on.
- *User information attack*: Exploit user data collected from password files (such as userid, user full name, and initial substring of name) to crack passwords.

Password Experiment—Group B

This sheet offers some advice on how to choose a good computer password. We are giving you this sheet as part of the password security experiment that was described in your introductory lecture. Different people are receiving different advice (but all advice should result in passwords at least as secure as you would choose if not participating in the experiment). Please do not discuss the experiment, this advice, or your choice of password with your friends.

A secure password is one that is very difficult to guess. Words that appear in a dictionary, or the names of people or places, are easy to guess. The most difficult passwords to guess are random sequences of letters. To help you choose a random sequence of letters for your password, we have printed a grid of random letters overleaf. Choose your password by closing your eyes and pointing at a random place on the grid. Choose eight characters this way and write them down on a scrap of paper.

Now log on using the initial password you have been issued, change your password to the new random password which you have chosen. The "Windows NT Tutor" tells you how to do this on pages 1.6–1.7.

You may find your password difficult to remember at first. Make sure that the scrap of paper on which you have written it is in a secure place, such as the back of your wallet or purse.

You should find that once you have entered it a dozen times or so, you will be able to remember it. Once you are sure you can remember it, destroy the scrap of paper where you wrote it down.

Finally, we strongly recommend that you change your password from time to time—at least once a term. As the experiment will run for the duration of this academic year, please keep this sheet and use this advice again when you choose your new passwords for Lent and Easter.

Figure 2. Random password group instruction sheet. Group members chose their passwords by closing their eyes and pointing randomly to a grid of numbers and letters.

- *Brute-force attack*: Try all possible combinations of keys.

We performed all but the fourth attack against all passwords, and attempted the fourth attack only on passwords that were six characters long.

We collected information on password length distribution and the number of cracked passwords in each group. We monitored the number of times users' passwords were reset, assuming users might forget the more difficult passwords and would either have to ask system administrators to reset their passwords or use other facilities. Four months after the tutorial, we surveyed experimental subjects by email, asking how difficult it was to remember their passwords and how many weeks had passed before they'd memorized it.

We also tested our experimental sample's validity by making the same attacks on the accounts of 100 first-year students who hadn't attended the introductory lecture or received experimental instructions.

Password Experiment—Group C

This sheet offers some advice on how to choose a good computer password. We are giving you this sheet as part of the password security experiment that was described in your introductory lecture. Different people are receiving different advice (but all advice should result in passwords at least as secure as you would choose if not participating in the experiment). Please do not discuss the experiment, this advice, or your choice of password with your friends.

To construct a good password, create a simple sentence of 8 words and choose letters from the words to make up a password. You might take the initial or final letters; you should put some letters in upper case to make the password harder to guess; and at least one number and/or special character should be inserted as well. Use this method to generate a password of seven or eight characters.

An example of such a composition might be using the phrase is "It's 12 noon I am hungry" to create the password "I's12&Iah" which is hard for anyone else to guess but easy for you to remember. By all means use a foreign language if you know one: the password "AwKdk.Md" from the phrase "Anata wa Kyuuketsuki desu ka ... Miyu desu" would be an example. You could even mix words from several languages. However, do not just use a word or a name from a foreign language. Try being creative!

Now log on using the initial password you have been issued, change your password to the new password which you have chosen. The "Windows NT Tutor" tells you how to do this on pages 1.6–1.7. Do not write your new password down.

Finally, we strongly recommend that you change your password from time to time—at least once a term. As the experiment will run for the duration of this academic year, please keep this sheet and use this advice again when you choose your new passwords for Lent and Easter.

Figure 3. Pass phrase group instruction sheet. Group members chose passwords based on mnemonic phrases.

Results

The selected passwords were on average between 7 and 8 characters long (7.6, 8.0, and 7.9 for the three groups, respectively). All three experimental groups chose slightly longer passwords than the sample group (mean length 7.3, difference statistically significant at $t = 4.53$, $p < .001$), who hadn't attended the introductory lecture.

The permuted dictionary attack was the most successful. Cracking passwords based on user information wasn't successful in any case, probably because of the limited amount of user information available in the password files (they don't include first names, for example). A brute-force attack successfully cracked all six-character passwords. Table 1 summarizes these results (treating brute-force attacks separately).

Password selection advice didn't affect the six-character passwords' susceptibility to brute-force attacks. In each experimental group, as well as in the comparison sample, a few users ignored the advice regarding password length and chose an insecure password.

We successfully cracked far more of the passwords that were longer than six characters in the control group than in either the random character or pass phrase group (significant at $\chi^2 = 24.8$, $p < .001$). The proportion of passwords cracked in the control group was lower than in the comparison sample (for example, 13 percent of the comparison sample used six-character passwords versus 5 percent in the control group; 13 passwords in the comparison sample were verbatim dictionary words versus three in the control group).

All passwords that were cracked successfully in the random character and pass phrase groups were dictionary words or permutations of dictionary words and numbers, contrary to advice given to the student. These results, together with the number of six-character passwords, provide a reasonable estimate of the level of user noncompliance with password-selection advice.

No one used special characters (neither letters nor numbers) except in the pass phrase group, whose instructions had given examples of passwords containing punctuation. This suggests that users should be explicitly encouraged to select passwords combining alpha, numeric, and special characters.

Very few users asked the system administrator to reset their passwords. Within a three-month period following the tutorial session, two members of the control group, one member of the random password group, and three members of the pass phrase group requested administrator resets.

Our email survey, which we sent to participants four months after beginning the experiment, asked two questions:

- How hard did you find it to memorize your password, on a scale from 1 (trivial) to 5 (impossible)?
- For how long did you have to carry around a written copy of the password to refer to? Please estimate the length of time in weeks.

Of the 242 replies we received, 13 indicated that the student had not used their accounts or had dropped the course. Valid responses from the groups clearly differed, as Table 2 shows.

Users in the random password group reported having more difficulty remembering their passwords (significant at $t = 8.25$, $p < .001$) and keeping a written copy for far longer (significant at $t = 6.41$, $p < .001$) than the other groups. This confirms Moshe Zviran and William Haga's results in an operational setting (see the "Related work" sidebar).

Because the differences in response rates weren't significant, we don't believe our results were significantly skewed by students in the random password group finding our advice so difficult that they stopped using the computer facilities.

Table 1. Results of password attacks, by test group.

GROUP	PASSWORDS CRACKED USING FIRST THREE ATTACKS		PASSWORDS CRACKED USING BRUTE-FORCE ATTACKS
	NUMBER	PERCENT OF TOTAL	
Control group	30	32	3
Random password group	8	8	3
Pass phrase group	6	6	3
Comparison sample	33	33	2

Many members of the random character group still carried written copies of their passwords at the time of the survey, indicating that they hadn't been able to memorize them.

Folk beliefs: True or false?

Our study confirms a number of widely held folk beliefs about passwords and debunks some others.

We've confirmed two folk beliefs—that users have difficulty remembering random passwords and that passwords based on mnemonic phrases are harder to guess than naïvely selected passwords.

However, we've debunked another folk belief—that random passwords are better than passwords based on mnemonic phrases. In our study, each appeared to be as strong as the other.

We've likewise debunked the belief that passwords based on mnemonic phrases are harder to remember than naively selected passwords. In fact, each type is as easy to remember as the other.

The last folk belief is that we can significantly improve security by educating users to select random or mnemonic passwords. In fact, both types of passwords suffered from a noncompliance rate of about 10 percent (including too-short passwords and passwords chosen contrary to the instructions). Although this is better than the approximately 35 percent of users who choose bad passwords with only cursory instruction, it's not a huge improvement. The attacker might have to work three times harder, but without password policy enforcement mechanisms, we can't make the attacker work a thousand times harder. In fact, our experimental group might be the most compliant a systems administrator can expect. Thus, this belief also appears to be debunked. In applications where one user can be harmed by another user's negligence, compliance monitoring and enforcement may be just as important as education.

Previous work suggests that the noncompliance rate could be even higher when users are required to remember multiple passwords, which usually increases the user's cognitive overhead and decreases memorability.[5] However, the issue of multiple passwords is beyond the scope of our experimental study.

Table 2. Responses to the email memorability survey.

GROUP	RESPONSES	DIFFICULTY LEVEL (1–5)	WEEKS
Control group	80	1.52	0.7
Random password group	71	3.15	4.8
Pass phrase group	78	1.67	0.6

The work we report here is merely a first step toward a better understanding of the applied psychology aspects of computer security. Many questions remain unanswered, and we plan to continue our experiments with future cohorts of students. In the meantime, we have some tentative recommendations for system administrators:

- Instruct users to choose mnemonic-based passwords, which are as memorable as naively selected passwords but as hard to guess as randomly chosen passwords.
- Size matters. With systems like Unix, which limit effective password length to eight characters, users should choose passwords of exactly eight characters. With systems such as Netware, which allows 14 characters but doesn't recognize case, users could be advised to choose passwords with 10 or more characters, which might further encourage the use of mnemonics. (This is a topic for our future work, as is enforcement generally.)
- Entropy per character also matters. Instruct users to choose passwords containing numbers and special characters as well as letters. If you don't, most users will choose passwords from a very small subset of the total password space.
- Compliance is the most critical issue. In systems where users can only put themselves at risk, it might be prudent to leave them to their own devices. In that case, expect that about 10 percent will choose weak passwords despite their instructions. In systems where a user's negligence can affect other users (for example, in systems where an intruder who gains access to a single user account can rapidly become root—that is, illicitly

Related work on password selection and memorability

The literature on password selection and memorability is surprisingly sparse. In their classic paper on Unix security, Fred Grampp and Robert Morris describe a test in which users were forced to create passwords of at least six characters with at least one nonletter.[1] The authors made a file of the 20 most common female names, each followed by a single digit. Of these 200 passwords, at least one was in use on each of several dozen machines they examined. Daniel Klein collected 13,797 password file entries from Unix systems and attacked them by exhaustive search, cracking about a quarter of them.[2] Password management guidelines from the US Department of Defense recommend using machine-generated random passwords.[3]

Moshe Zviran and William Haga asked 106 students to choose passwords and write them on a questionnaire.[4] The questionnaires also gave each student a random password and asked them to remember both. Three months later, 35 percent of the students could recall their self-selected passwords but only 23 percent recalled their assigned random passwords. In addition, 14 percent wrote the self-selected passwords down whereas 66 percent wrote down the random password. The students weren't actually using the passwords during the intervening three months, however. So, although these results provide a quantitative point of reference for the difficulty of remembering random passwords, the test doesn't model a real operational environment.

Anne Adams and Martina Angela Sasse, and Sasse, Sacha Brostoff, and Dirk Weirich surveyed system users about their experiences with passwords.[5,6] Although they discussed memorability issues and concluded that users should be instructed to construct secure and memorable passwords, neither article put much effort into identifying "specific, positive advice" on how to "compose passwords that are both easy to remember and difficult to crack."[7]

References

1. F.T. Grampp and R.H. Morris, "Unix Operating System Security," *AT&T Bell Laboratories Technical J.*, vol. 63, no. 8, 1984, pp. 1649–1672.
2. D.V. Klein, "Foiling the Cracker: A Survey of, and Improvements to Unix Password Security," *Proc. Usenix Security Workshop*, Usenix Assoc., 1990; www.deter.com/unix/.
3. US Dept. of Defense, *Password Management Guideline*, CSC-STD-002-85, 1985.
4. M. Zviran and W.J. Haga, "A Comparison of Password Techniques for Multilevel Authentication Mechanisms," *Computer J.*, vol. 36, no. 3, 1993, pp. 227–237.
5. A. Adams and M.A. Sasse, "Users Are Not The Enemy," *Comm. ACM*, vol. 42, no. 12, Dec. 1999, pp. 40–46.
6. M.A. Sasse, S. Brostoff, and D. Weirich, "Transforming the 'Weakest Link': A Human–Computer Interaction Approach to Usable and Effective Security," *BT Technical J.*, vol. 19, no. 3, July 2001, pp. 122–131.
7. P. Abrahams, letter to *Comm. ACM*, vol. 43, no. 4, 2000, pp. 11.

get a system administrator's privileges—using well-known and widely available techniques), consider enforcing password quality by system mechanisms.
- If a benefit is to be had from the use of centrally assigned random passwords, it appears to come from the fact of central assignment (which enforces compliance) rather than randomness (which can be achieved with mnemonic phrases).

An interesting and important challenge is finding compliance enforcement mechanisms that work well with mnemonic password choice. Proactive password checkers,[7] which verify that a password is not part of a known weak subset of the password space, might be an effective tool. But as this article has shown, what engineers expect to work and what users actually make to work are two different things. Rigorous experimental testing of interface usability is in our view a necessary ingredient for robust secure systems. □

References

1. G.J. Johnson, "A Distinctiveness Model of Serial Learning," *Psychological Rev.*, vol. 98, no. 2, 1991, pp. 204–217.
2. G.A. Miller, "The Magical Number Seven, Plus or Minus Two: Limits on Our Capacity for Processing Information," *Psychological Rev.*, vol. 63, 1956, pp. 81–87.
3. A. Paivio, "The Empirical Case for Dual Coding," *Imagery, Memory and Cognition: Essays in Honor of Allan Paivio*, J.C. Yuille, ed., Erlbaum, 1983, pp. 307–322.
4. H. Davies, "Physiognomic Access Control," *Information Security Monitor*, vol. 10, no. 3, 1995, pp. 5–8.
5. A. Adams and M.A. Sasse, "Users Are Not The Enemy," *Comm. ACM*, vol. 42, no. 12, 1999, pp. 40–46.
6. B. Patterson, letter to *Comm. ACM*, vol. 43, no. 4, 2000, pp. 11–12.
7. J. Yan, "A Note on Proactive Password Checking," *Proc. 2001 ACM New Security Paradigms Workshop*, ACM Press, 2001, pp. 127–135.

Jeff Yan *is assistant professor at the department of computer science and engineering at the Chinese University of Hong Kong. He is interested in most aspects of information security, both theoretical and practical, and his recent research is largely about systems security (in particular security in online games), applied crypto, and human aspects of security. He has a PhD in computer science from Cambridge University. Contact him at jyan@cantab.net.*

Alan Blackwell *is a lecturer in the Cambridge University Computer Laboratory, with qualifications in professional engineering and experimental psychology. His research interests include design practice and human-computer interaction. He is a Fellow of Darwin College, Cambridge. Contact him at Alan.Blackwell@cl.cam.ac.uk*

Ross Anderson *leads the security group at the Computer Laboratory, Cambridge University, where he is a professor of security engineering. He was one of the founders of the study of information security economics. A Fellow of the IEE, he was also one of the pioneers of peer-to-peer systems, of API attacks on cryptographic processors, and the study of hardware tamper-resistance. He was one of the inventors of Serpent, a finalist in the competition to find an Advanced Encryption Standard. He wrote the standard textbook* Security Engineering—A Guide to Building Dependable Distributed Systems. *Contact him at Ross.Anderson@cl.cam.ac.uk.*

Alasdair Grant *is a technical lead on code generation in the compilation tools group at a leading CPU vendor. His research interests include program correctness, static analysis, and binary translation. He has an MA in mathematics and computer science from Cambridge University and is a member of the ACM and SIGPLAN. Contact him at algrant@acm.org.*