

# The Effectiveness and Usability of Passphrases for Authentication

**Mark Keith**

Arizona State University  
mark.keith@asu.edu

**Benjamin Shao**

Arizona State University  
benjamin.shao@asu.edu

**Paul John Steinbart**

Arizona State University  
[paul.steinbart@asu.edu](mailto:paul.steinbart@asu.edu)

## ABSTRACT

In developing password policies, IT managers must strike a balance between security and memorability. Rules that improve structural integrity against attacks (e.g., increasing length and multiple character types) may also result in passwords that are difficult to remember. Recent technologies have relaxed the 8-character password constraint – permitting the creation of longer pass-“phrases” consisting of multiple words. Psychology theories suggest users can remember passphrases at least as well as passwords. This paper reports an experiment currently in progress that tests the usability of passphrases. Subjects are randomly assigned to three different password creation techniques: a control group with no constraints, a secure group given strong password requirements, and a passphrase group. It is expected that the passphrases group will have fewer failed login attempts than the secure group and no more failed login attempts than the control group. Practical implications include stronger authentication with reduced help desk costs.

## Keywords

Passwords, security, authentication, memory

## INTRODUCTION

The effectiveness of user-generated passwords as a means of protection for sensitive data and resources has been a concern for IT managers and users for decades (Morris and Thompson, 1979). Consequently many organizations have created policies or guidelines for password development. While stringent password guidelines (e.g., length, character set, etc.) may significantly increase their effectiveness as an authentication mechanism, they can also create additional costs associated with resetting strong passwords that users forget. In addition, such policies may cause users to write down their passwords or use the same passwords across multiple domains, thereby weakening overall security (Ives et al., 2004). A smart hacker then only needs to compromise a less-secure website, for example, to obtain the user’s secure password which can then be used to access the real target. These problems suggest that IT managers must strike a delicate balance between security and memorability in developing password requirements to avoid various types of risks. Many studies have attempted to support this endeavor either by suggesting authentication schemes which may be easier for users to remember, by offering guidelines to improve password quality, or by both (Yan et al, 2004) (Zviran and Haga, 1999; 1993; 1990) (Gong et al, 1993).

One common assumption in both practice and previous research is that passwords can be no longer than 8 characters. Tools designed to discover a password by “brute force” or guessing methods are freely available on the World Wide Web (WWW) that can crack many 7-8 character passwords in only one or two hours. Recent technologies like Windows XP and web-based authentication schemes have removed the 8-character maximum for passwords, thus permitting the creation of longer pass-“phrases” consisting of multiple words. Analytically, it can be shown that increased password length directly improves resistance to brute-force attacks. There is no empirical evidence, however, concerning the usability of such passphrases.

This paper reports an experiment currently in progress designed to test how well users can remember longer passphrases. The next section identifies relevant literature and research in the area of password security and memorability. Subsequent sections describe the experimental design and status of the project. This paper concludes with a discussion of expectations and recommendations for future research and practice.

## LITERATURE REVIEW

The idea of passphrases is not new. Porter (1982) suggested the use of passphrases 30 to 80 characters long formed in sentence structure to improve memory performance. To maneuver around the length limit, he recommended using a hashing algorithm to reduce the passphrase to 8 characters before submission. More recently, researchers have suggested using the first letter of each word in a phrase to construct a seemingly random password restricted to 8 characters (Yan et al, 2004). Empirical results indicate that such phrase-based passwords provide both security equal to that of completely random passwords and memorability as good as that of user generated passwords. However, there are still two limitations of such passphrase-based passwords: (1) the password length is still limited to 8 characters, and (2) the step of translating the passphrase into a password creates an opportunity for mental errors. Replacing such phrase-based passwords with the passphrases *per se* eliminates both problems.

Psychology theories suggest an appropriate length and structure for passphrases. Miller (1956) proposed that the human mind can distinguish 7 plus or minus 2 “chunks” of information in short-term memory. A “chunk” of information need not be limited to a single character. To increase the amount of information stored in memory, one can simply increase the size of the “chunks” – transforming characters into words. It was later clarified that Miller’s magical number 7 is not necessarily the *optimal* size for short-term memory, but more of an *asymptotical limit* (Doumont, 2002). Human memory may actually be optimized around 3-5 chunks of information. Three- to five-word passphrases may vary in length from approximately 12 to 30 characters. From a security standpoint, 12 to 30 characters would significantly improve password integrity over that of 8 characters.

## EXPERIMENTAL DESIGN

This study examines the usability of passphrases. We compare user login performance (success) across three conditions: passphrases, passwords that conform to stringent requirements, and user-generated passwords created without specific requirements. Undergraduate students from a variety of disciplines enrolled in an elective course on web design and development participated in the experiment.

### Hypotheses

Based on previous studies (Yan et al, 2004), we believe users of phrase-based passwords will demonstrate better memory performance than users of highly secure 8-character random passwords. In addition, research by Millar (1956) and subsequent authors leads us to believe passphrases constructed from complete words rather than just acronyms will be as easy to remember as simple passwords.

In laboratory settings, memory performance of passwords has been operationalized as the true/false value of a subject’s ability to reproduce a correct password match (Zviran and Haga, 1993). A field study thus can measure memory performance from the unsuccessful login attempts made by users who actually use their passwords to perform some tasks. However, this measure is complicated by the fact that not every failed login attempt is the direct result of a memory failure. Because passwords are typically represented with ‘\*’ during entry to prevent discovery by onlookers, users are prone to making typographical errors. This being the case, we formulate the following hypotheses:

H1: Users of passphrases will experience a rate of unsuccessful logins per attempt no different from that of users of self-generated simple passwords.

H2: Users of passphrases will experience a rate of unsuccessful logins per attempt lower than that of users of highly secure 8-character passwords.

### Methodology

Students make use of a website from which they download assignments and various class materials. Each student must register on the class website in order to access these materials. Subjects are informed that their activities on the site will be monitored for research purposes, but are not explicitly aware of which activities are being monitored or why. Every time a student logs in to the class website, the system records their username, password, timestamp, and IP address. We believe this testing environment increases the external validity of our results because it examines the actual use of passwords and passphrases as a step in accomplishing a meaningful task, rather than their ability to recall passwords from memory.

At the beginning of the semester, every student’s school username was obtained from a class registration list and entered into an online database used by the class testing site. Students were then assigned randomly to one of three groups by associating a group identification number with their username in the class website database. Registering on the website was a simple 3-

step process: 1) after following the “Register” link on the class website, students were asked for their usernames; 2) upon verification of the username, each student then had the opportunity to offer voluntary demographic information about themselves besides what was already known from the class registration list; and 3) based on their randomly assigned group number, each student is then redirected to one of three separate web pages, each with different criteria for generating a website password.

*Control Group*

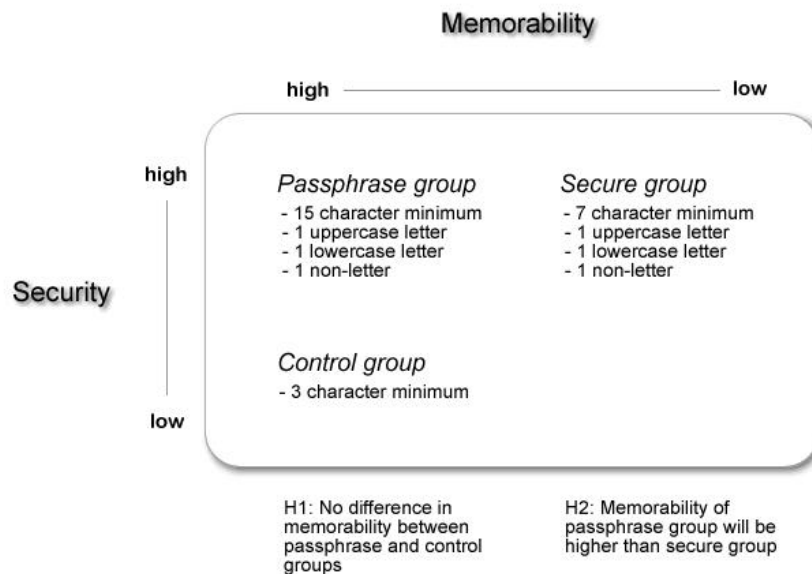
Some organizations allow their users to generate passwords with little or no restrictions. Students in the control group received no explicit directions for generating a password. The website simply ensured that users enter a password at least three characters long. We assume that such passwords were created with an emphasis on the ease of memorability with less concern about the level of security.

*Secure Group*

Many organizations enforce some minimum level of security in user’s passwords. To replicate a similar situation, the secure group was required to generate passwords that were at least seven characters in length and contained at least one upper-case letter, one lower-case letter, and one non-letter. Examples were provided to help users see how they could generate such a password. We expect that the passwords of this group to be more architecturally secure than those of the control group, but more difficult to remember.

*Passphrase Group*

Because passphrases are less common in research and practice than the passwords of groups one and two, careful thought was given on how these passphrases should be constructed. Users were required to create a passphrase at least 15 characters in length including one non-letter, one lower-case letter, and one upper-case letter. Fifteen characters represent approximately 3 to 5 words, which is the optimal range for short-term memory (Miller, 1956). We expect these passphrases to significantly improve security and be as easy to remember as the passwords of the control group.



**Figure 1. Summary of Groups and Hypotheses**

If students in any group were to construct passwords that are either identical or similar to those they have used in other situations, their ability to remember those passwords is likely to be enhanced. This potential bias would make it difficult to accurately compare the passphrase group with the other groups. To help prevent this potential bias, the students in both the control and secure groups received a verbal warning against creating a previously used password. Group 3, the passphrase group, was not given this warning because the 15 character minimum requirement is likely to prevent reuse. Subjects were again reminded of this in the password instructions on the website and by a validation message when they attempted to

complete the registration process. The warning provided was similar to the one employees might receive from their network administrator and provided the option to go back and change the password. While these procedures still do not completely prevent any student from violating the recommendation, we feel it will help to remove some bias from our results.

Students who forget their password must contact the instructor by either phone or email. The instructor then verbally relays the password back to the student and records the time, date, and student name. Password and passphrases are never changed during this experiment.

Of the 60 students originally registered for the class, 58 still remain actively enrolled and participating. There are 8 assignments in total spanned across 10 weeks which require the students to login using their passwords or passphrases. Details about every attempt to login to the class website are recorded in an online database. After data collection is complete, Analysis of Variance (ANOVA) will be used to determine the significant differences in login failures between groups. To separate memory-related login failures from those due to typographical errors, two grading assistants who are unaware of the study's hypotheses will code their interpretation of the cause of each failure. Also, subjects will indicate the cause of each of their login failures from a list provided at the conclusion of the experiment. In addition, we will administer a short anonymous survey about the techniques students used in constructing their passwords and passphrases. It will also ask them if their passwords are based on anything they have previously used.

## DISCUSSION

While this experiment is designed to replicate a more realistic environment in which to study the use of passwords and passphrases, some limitations still exist. Most notable is the use of student subjects who may not accurately reflect the actual behaviors and practices of the population. Some researchers have suspected that the perceived risk associated with the potential loss of data due to compromised passwords may influence password creation patterns (Highland, 1997). While risk is a variable controlled for by this design (all subjects are protecting the same information), results may still vary across perceptions of risk.

Since the experiment is still ongoing, it is difficult to draw any conclusions. At least one student from each group had forgotten the password and had to ask the experimenter to resend it. Coding of the cause of login failures has not yet been completed.

We expect to find that passphrases are easier to remember than are cryptic passwords that meet stringent security requirements. Such a finding has important practical implications. It can be shown analytically that long passphrases are more resistant to brute-force cracking techniques than strong 8-character passwords. There is also abundant evidence that strong password creation policies either create help desk costs when users forget their passwords or result in other vulnerabilities (e.g., writing down the password). Therefore, our expected findings would suggest that adopting the use of passphrases may increase overall security and reduce password maintenance costs.

## REFERENCES

1. Doumont, J. (2002) Magical Numbers: The Seven-Plus-or-Minus-Two Myth, *IEEE Transactions on Professional Communication*, 45, 2, 123-127.
2. Gong, L., Lomas, M. A., Needham, R. M. and Saltzer, J. H. (1993) Protecting Poorly Chosen Secrets from Guessing Attacks, *IEEE Journal on Selected Areas in Communications*, 11, 5, 648-656.
3. Highland, J.H. (1997) Changing Passwords. *Computers & Security*, 16, 3, 183-184.
4. Ives, B., Walsh K. R. and Schneider H. (2004) The Domino Effect of Password Reuse, *Communications of the ACM*, 47, 4, 75-78.
5. Miller, G. A. (1956) The magical number seven, plus or minus two: some limits on our capacity for processing information, *Psychology Review*, 63, 81-97.
6. Morris, R. and Thompson, K. (1979) Password Security: A Case History, *Communications of the ACM*, 22, 11, 594-597.
7. Porter, S. N. (1982) A Password Extension for Improved Human Factors, *Computers & Security*, 1, 1, 54-56.
8. Yan, J., Ross, A. B. and Grant, A. (2004) Password Memorability and Security: Empirical Results, *IEEE Security & Privacy*, 2, 5, 32-39.
9. Zviran, M. and Haga, W. J. (1990) Cognitive Password: The Key to Easy Access Control, *Computers & Security*, 9, 8, 723-736.
10. Zviran, M. and Haga, W. J. (1993) A Comparison of Password Techniques for Multilevel Authentication Mechanisms, *The Computer Journal*, 36, 3, 227-237.
11. Zviran, M and Haga, W. J., (1999) Password Security: An Empirical Study, *Journal of Management Information Systems*, 15, 4, 161-185.