Tricerion®
identity integrity

The usability of picture passwords

Norman Fraser PhD
Tricerion Limited

Executive summary

- In the context of growing identify theft, and especially phishing and related scams, conventional passwords no longer provide adequate authentication security for online services.
- *Zero-footprint* strong authentication (i.e. requiring no software installs or hardware tokens) can be achieved with careful use of picture passwords, i.e. sequences of picture symbols.
- The *picture superiority effect* was established in experimental psychology almost 50 years ago, and has withstood frequent experimental testing since then. It shows that under normal conditions people remember information presented in picture form much better than information presented in textual form.
- The *picture superiority effect* holds up in experiments specifically comparing picture passwords with numeric PINs or alphabetic passwords. Picture passwords have been shown at worst to be no less usable than conventional passwords and at best are significantly more usable.
- Users typically expect picture passwords to be more difficult to use but change their minds after a short period of exposure.
- On average, users make more input errors when first presented with picture passwords than they do with text passwords or PINs. This result reverses as quickly as the second occasion of use.
- A significant percentage of users share their password or PIN with family or friends. They see shareability as a useful feature. But shareability presents a significant security threat and is one of the main enabling factors in phishing attacks. Picture passwords are very difficult to share and so are more resistant to self-disclosure.
- Tricerion's SafeLogin™ solution implements the usability advantages of picture passwords in a low cost, zero-footprint strong authentication solution.

Introduction

Conventional online authentication protocols, involving a combination of username and password (or PIN), are well-established but problematic. As Bill Gates observed at a recent conference, “There is no doubt that over time, people are going to rely less and less on passwords. People use the same password on different systems, they write them down and they just don't meet the challenge for anything you really want to secure” (RSA Conference, 2006).

This password vulnerability is related in part to the practices of users, and in part to inherent weaknesses in the passwords themselves, such as the ease with which passwords may be shared or stolen (Adams & Sasse, 1999; Besnard & Arief, 2004; Ensor *et al*., 2004).

In addition to the weaknesses of password properties and usage, regulatory pressures look set to render simple username plus password protocols obsolete. For example, the US banking regulator, the FFIEC, published guidelines making ‘strong authentication’ mandatory for identity verification on all regulated institutions’ web sites (FDIC 2005). Inertia may lead institutions and users to regret change, but the *status quo* with passwords is no longer an option; conventional username and password protocols **must** be changed or augmented to meet regulatory requirements.

There are three main strong authentication technologies available for organizations looking to upgrade their approach to identity verification:

Tokens/one-time passwords. Users are issued with a physical device (often called a ‘token’ or ‘keyfob’) which displays a frequently changing PIN. Users must carry the device with them and can only log in using the currently valid PIN. Alternatively, users may be issued with a list of passwords, each of which may be used once only, in the order given. Token devices can be expensive to roll out and administer, and may meet with user resistance when users are issued with different tokens by different service providers. Some financial institutions see tokens as a solution for high-net-worth accounts only, and not for the mass market. Both tokens and one-time passwords involve logistical overheads and require the user to be in physical possession of their ‘second factor’ in order to access the secure service. This latter fact may drive users away from the online service on occasions when they do not have the second factor with them.

Downloads/installs. Users may install software on their PC (e.g. a security toolbar or custom user interface) which ensures that only *bona fide* online services are accessed. Disadvantages with this approach include clashes with corporate software installation policies and the requirement for the user always to access online services from the same PC(s), i.e. those on which access software has previously been installed. In our view the realities of online service use in many applications (e.g. retail banking) do not support an installation-based authentication solution.

Strong mutual authentication. Strong mutual authentication authenticates the user to the organization and the organization to the user without the need to issue any physical devices or password lists to the user, and without making any installation on the user’s PC. Tricerion’s SafeLogin strong mutual authentication system offers just such a ‘zero-footprint’ authentication solution. Details of the Tricerion SafeLogin system can be found at www.tricerion.com.

Picture passwords in the Tricerion SafeLogin system

In its most secure configuration the Tricerion SafeLogin system supports picture passwords. Thus, the user will be issued with a password such as the one shown in Figure 1. (The number of pictures in the password is determined by the service provider; we use four here for illustrative purposes only.)

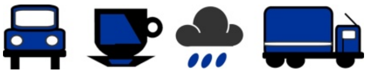

Figure 1: Picture password

The password is entered by clicking symbols on a password entry keypad, e.g. Figure 2.

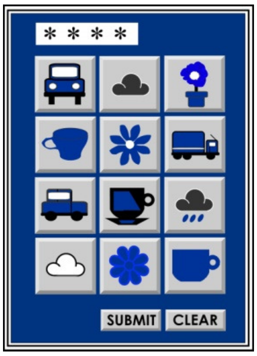

Figure 2: Password entry keypad, using pictures

A number of features of this keypad are personalized to the user, and selected from the user's stored account record by means of the user's username. For example, the background colour and border design are constant for this user, but may be different for other users. This provides a measure of protection against phishing attacks, by alerting the user to any changes from their familiar keypad 'look-and-feel', which is unknown to the phishing attacker. However useful this may be, any security 'solution' which relies solely on the alertness of the user to changes in the user interface should be treated with grave suspicion.

SafeLogin effectively renders it *impossible* for users to disclose their password on a randomly generated phisher's keypad. It achieves this by making the set of pictures to be presented to the user part of the user's personalization data. Each user always sees the same picture set, which is only a small sub-set of the total set of available pictures. Thus a randomly generated keypad has a vanishingly small chance of including all the pictures necessary for the user to enter their password.

Figure 3 shows a randomly generated keypad, which differs from the one shown in Figure 2 in look-and-feel but also, crucially, in that it does not contain all the pictures necessary to enter the password shown in Figure 1.

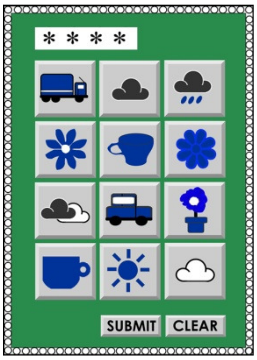

Figure 3: Randomly generated keypad

Difficulties posed to phishing attackers by the use of personalized keypads might lead them to attempt to elicit passwords from users simply by asking them to write their password down, e.g. in a reply email or on a custom phishing site. However, the use of picture passwords renders this extremely difficult, such that any act of self disclosure may be deemed wilful. For example, the picture password shown in Figure 1 can be described simply as “car, cup, cloud, truck”. The keypad in Figure 2 yields no less than 18 different sequences of pictures describable in these terms, only one of which is the correct password. Neither is it possible to describe the password by reference to the symbol position on the keypad, since the position of the symbols is scrambled on each presentation as in Figure 4. (This also serves an additional security function within SafeLogin, since values returned are simply the co-ordinates of each click on a one-time keypad, thus ensuring that passwords never exist in digital form outside the service provider’s firewall).

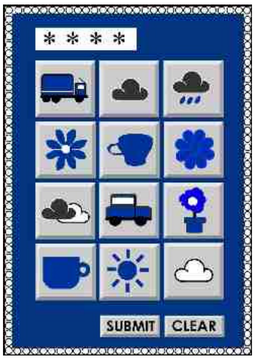

Figure 4: Scrambled version of keypad in Figure 2

There is no doubt that picture passwords as implemented in SafeLogin offer a significant improvement on the security of conventional passwords, and that they comply with regulatory requirements such as the strong authentication guidelines of the FFIEC. But are they usable by ordinary users?

Picture passwords and usability

It is often said that passwords involve a trade-off between usability and security. Conventional passwords were known to have security problems, but these were deemed reasonable in light of their economy and ease of use. But as online fraud grew and new identity theft phenomena such as phishing appeared, bringing the spectre of widespread account hijacking, so it became clear that security should take centre stage. Some organizations introduced password rules, defining minimum characters required, and specifying a mix of alphabetic and numeric characters – both innovations which lessened usability but improved security. Now, as regulators have indicated, minor adjustments like these are not enough; what is needed is a more significant change to strong authentication methods, such as that embodied in the SafeLogin system.

Changing from alpha-numeric to picture passwords cannot be done without users noticing. Viewed positively, this has the advantage of signalling to users that the service provider is taking their security seriously.

Also on a positive note, SafeLogin's use of a clickable cellphone-style keypad for password entry is so simple that it does not have to be explained to users – its purpose and operation are obvious. It embodies the principles of intuitive design celebrated by design guru Donald Norman in his ground-breaking study *The psychology of everyday things* (Norman 1988). He pointed out, for example, how the design of a cup is so intuitive that nobody needs training to figure out that the handle is there for holding onto when the cup is full of hot liquid. Similarly, we have found that nobody needs lessons to figure out how to enter their password using a Tricerion keypad, even using a scrambled keypad (though this only remains true so long as the keypad is kept fairly small, i.e. not more than 20 picture items).

But how will users respond to picture passwords? Will they be able to remember them? Will they prove easier or more difficult to use than conventional passwords? Will users like them or resist them?

In the sections below we examine the following sources of data:

- Common sense knowledge
- The findings of psychology about human capacities
- Experiments which have specifically explored picture passwords versus conventional passwords

Common sense knowledge

Picture passwords may initially be predicted to be harder to use than conventional passwords because they are unfamiliar, and unfamiliar processes always appear harder than familiar ones. But some common sense reflection should undo the initial impression.

Consider the way in which infants acquire their understanding of the world. First they experience real-world objects, then they encounter pictorial representations of objects (e.g. in infant story books), and only some time – typically a couple of years – after mastering picture recognition do they move on to be able to handle more abstract alphabetic and numeric text.

Some adults with normal intelligence never learn to handle text, i.e. they never learn to read or write, because it is a non-trivial learned skill which needs to be both taught and learned. But picture perception and pictorial reasoning are innate human accomplishments. Even 'wild children' starved of normal human input learn to process and handle the objects they see.

Some adults with learning disabilities never learn to use language at all, not even spoken language, but they can still recognize, discriminate and remember pictures. A number of picture-based communications programs exist for such non-linguistic or delayed linguistic individuals, many of them based on the pioneering work of Woodcock *et al* (1969).

These observations suggest that human processing of pictures is more basic, rooted deeper in human cognitive processes than processing of text. This also helps explains why we find icons and graphics more readily accessible than words, as, for example, in the case of road signage, where meaningful graphics can communicate more information faster than text.

The findings of psychology

There is a large body of experimental evidence to back up these common sense observations. Amongst other things these experiments demonstrate what is known as the *picture superiority effect,* namely that concepts are much more likely to be remembered if they are presented as **pictures** rather than in textual form. Many studies show that people's recognition memory for pictures is extraordinary. Shepard (1967) reports an experiment he conducted in 1959 in which he showed 600 pictures to subjects, who were later required to identify the previously seen pictures from a set including previously unseen pictures. Similar tests were carried out for words and for sentences. The successful recall rates achieved were as follows:

Table 1: Recall success rates (Shepard 1967)

	Recall success
Pictures	98.5%
Words	90.0%
Sentences	88.2%

Even a week after first presentation subjects successfully recalled more than 85% of the pictures, a result far superior to the word and sentence recall rates. Similar results are reported by Nickerson (1965), Standing *et al* (1970), Nelson (1976) and a host of other experimental psychologists. To date, all studies in recall memory have confirmed that picture information is recalled better than verbal information.

Specific properties of the pictures also affect the rate of recall. Ritchey (1982) demonstrated that recall is better for outline drawings than detailed drawings. Jesky (1984) showed that recall of colour drawings is superior to black and white drawings, and recall of both colour and black and white drawings are superior to line-only drawings. Alfahad (1990) and Berry (1991) amongst others confirmed these results.

A number of studies have been conducted into the effects of age on picture memory. Keitz & Gounard (1972), Winograd *et al* (1982) and Park *et al* (1986) demonstrated comparable picture superiority in both young and old subjects. Gadzella (1991) found that all ages recalled picture information better than textual information, though the older age group of subjects recalled picture information slightly less well than the younger age group.

The experimentally established picture superiority effect has been applied practically in a number of fields, such as advertising, where the importance of visual branding, product placement and association of product messages with positive images are well understood (Childers & Houston 1984). The use of pictures to improve learning effects in education is standard practice and uncontroversial, and builds directly on the picture superiority effect.

What are the psychological mechanisms that make picture superiority work? Many psychologists conclude that different cognitive capabilities are applied. Whereas previously encountered textual information is *remembered*, picture information is *recognized*. It is interesting that picture information is consistently recalled better than verbal information, but within verbal recollection concrete words are recalled better than abstract words. The 'dual coding hypothesis' offers an explanation for this based on the 'picturability' of concrete words, which allows certain visual recognition effects to influence the memory process, even when the picture is not present (Paivio 1971).

All this is bad news for conventional passwords. Being textual, experimental psychology predicts that conventional passwords will be harder to remember than picture passwords – for all user ages. It also predicts that meaningful, concrete textual passwords (which are easiest for hackers to guess) will be easier for users to remember than more secure abstract passwords. Users' preference for meaningful rather than abstract textual passwords is well-grounded in human psychology – even though it goes against the requirements of higher security.

Picture passwords, on the other hand, seem well-suited to human abilities. But how do they actually perform in realistic conditions of password use?

Experiments with picture passwords

A considerable body of experimental research specifically on picture passwords has been built up over more than ten years. These experiments have explored picture passwords of many kinds, including recognizing target pictures from a selection of 'distractors', selecting pre-arranged areas within a single picture, recognizing human faces from the password equivalent of police lines, and asking the user to draw a previously agreed picture.

Here we will examine two independent studies, which bear directly on the SafeLogin picture password methodology.

The first experiment, carried out by two Berkeley scientists Dhamija and Perrig (2000), involved a series of experiments comparing picture passwords with conventional passwords.

At the start of the procedure Dhamija and Perrig interviewed their subjects to learn about their experience of passwords. They found that on average subjects had 10-50 different instances in their present lives where passwords were needed, but they relied on only 1-7 different text strings to cover all of them. When the researchers probed subjects' experience of password use they found that:

> "All participants expressed strong feelings of dislike and frustration with their experiences remembering, using and losing passwords. Yet surprisingly, most people preferred them to alternatives. For example many disliked hardware tokens because of experiences losing or

> misplacing them. A couple of participants who had experience with biometrics (fingerprint readers) felt that these systems were unreliable and performed poorly compared to passwords. Others disliked biometrics because of perceived privacy threats.”

Users have a love-hate relationship with passwords: they have some frustrating features, but they are perceived to be the least of the currently available evils. However, these observations were collected before users were exposed to picture passwords.

Next, the researchers conducted two experimental sessions. During the first session, subjects were asked to create a four digit PIN, and a password with a minimum of six characters. Subjects were also asked to create two picture passwords, one consisting of five abstract art images and another consisting of five photographs. Subjects selected each picture password from a set of one hundred images. From subject to subject, the order in which passwords, PINs, abstract art and photo passwords were created was varied to ensure there was no bias due to task sequence.

Subjects were next asked to authenticate themselves using all four techniques, in the same order that they had created them. This ensured that several minutes and tasks had elapsed between each creation and authentication event. To authenticate using art and photo passwords, subjects had to select their five images from amongst twenty ‘distractor’ images.

The second session occurred a week later and subjects were asked to repeat the authentication task for each of the four techniques.

The login failure rates under all conditions are shown in Table 2.

Table 2: Login failure rates

	PIN	**Password**	**Art**	**Photo**
Failed logins	5%	5%	0%	0%
Failed logins (after 1 week)	35%	30%	10%	5%

The results show dramatically better recall rates for both picture conditions over both textual conditions.

It is interesting to note the subjective responses of subjects before and after exposure to picture passwords. (Dhamija and Perrig refer to picture passwords as ‘portfolios’.)

> “Although some users remarked that they would never be able to remember the portfolios they created, all were surprised that they could recognize their images and at how quickly the selection took place. It is interesting to note that after the first week, more users forgot their usernames than their portfolios. The majority of users reported that photo portfolios were easier to remember than

> PINs and passwords, especially after 1 week, and that they would use such a system if they were confident that it was secure and if [initial] image selection times were improved."

In interviewing subjects, Dhamija and Perrig noted that many subjects viewed the ability to share passwords as a desirable 'feature', and that almost all confessed to sharing bank PINs with family or friends. Unfortunately, the 'shareability' feature of PINs and conventional passwords, though popular with users, seriously compromises security. In a 'lighthearted' and 'unscientific' experiment, security vendor VeriSign found that 65% of people stopped at random on a street in San Francisco were willing to give up their passwords in return for a Starbucks coffee (Claburn 2005). This confirmed an earlier survey by PentaSafe Security Technologies, who found that four out of five commuters in London's Victoria Station were willing to disclose their passwords in return for a cheap pen (Lemos 2002).

The combination of passwords' inherent shareability and the willingness of a large percentage of users to share their passwords inappropriately, is what makes social engineering-based phishing attacks possible. The importance of Dhamija and Perrig's results is that they show improved authentication usability for picture passwords, with dramatically reduced shareability. As one journalist reporting Dhamija and Perrig's results noted:

> "The advantage of all this is that while you can recognise an image, you can't really describe it to others. That means the 'password' stays safely in your memory and nowhere else. It's every financial institution's dream come true." (Dreyfus 2000).

The second set of experiments we shall consider was carried out by De Angeli *et al* (2005). In the first experiment, PINs were compared with three different types of picture password, as summarized in Table 3.

Table 3: Systems tested by Angeli *et al* (2000)

System	Password	Key Location	Input Order
PIN	Sequence of 4 numbers from 10	Constant	Fixed
VIP1	Sequence of 4 pictures from 10	Constant	Fixed
VIP2	Sequence of 4 pictures from 10	Random	Fixed
VIP3	4 pictures from 12	Random	Random

Before the experimental procedure was initiated subjects were asked to evaluate their own abilities to remember PINs versus picture passwords over a one week period. On average, subjects stated their belief that numbers are significantly easier to remember than pictures.

In fact, the experiment revealed no significant difference in outcome between the authentication conditions, since no users forgot their code. However, some 5% of authentications produced an input error at first attempt. Of these, condition VIP3 accounted for more errors than all three other conditions put together. VIP2 produced fractionally more errors than VIP1, but there was no significant difference between VIP1 and PIN. Clearly, inviting the user to select their password pictures in any order (VIP3) increased the error rate.

SafeLogin is a VIP2-type system, with a fixed sequence picture password, and an entry keypad that scrambles keys on each presentation. This method performed marginally less well than PINs and VIP1, though the difference was barely statistically significant. One explanation could be that ‘muscle memory’ contributes to success in fixed keypad conditions like PIN and VIP1. On the other hand, fixed position keypads also increase the likelihood of password disclosure through ‘shoulder surfing’, so the marginally higher error rate in VIP2 may be justified through security gains.

Differences between PINs and picture passwords emerge starkly when the time distribution of errors is analysed. Most of the picture password errors occurred at the start during the ‘training’ phase, whereas the opposite was true of PIN errors. For example, though PIN and VIP1 have identical overall error rates, all but one of the VIP1 errors occurred during the training phase and all but one of the PIN errors occurred after one week. What this actually reveals is that subjects who are not used to picture passwords make some errors during their first exposure to a new authentication methodology, but that once learned, picture passwords are more reliably recalled and entered than PINs, as predicted by the picture superiority effect.

Subjects were asked to rate their satisfaction with the user experience under the four different conditions. Before either of the experimental phases they were asked to rate their expectations of the usability of PIN entry by means of a clickable keypad. After the second experimental phase (at the end of the week) they were asked to rate their experience. Under all four conditions subjects revealed that they found the experience of using a clickable keypad better than they had expected before the procedure began. When the different conditions were compared, subjects rated the picture password conditions VIP1 and VIP2 more usable than conditions VIP3 and PIN.

Conclusion

Common sense knowledge, the general findings of experimental psychology, and specific experiments comparing picture passwords with conventional passwords and PINs all point to the comparative recallability, usability, non-shareability, and user acceptance of pictures. Pictures are *recognized* and this is a deeply rooted human capability, whereas textual passwords are *remembered* and this is inherently more difficult for humans to accomplish. The *picture superiority effect* is real and lends picture passwords ‘unfair advantage’ when compared to conventional passwords and PINs. Some of these facts are mildly counter-intuitive, so users need to be encouraged to start using picture passwords in order to discover that pictures turn out to be easier and better to use than was feared. Users can usually be convinced of the usability of picture passwords after a very short period of exposure.

While the scientific evidence in support of picture passwords is strong, there is no substitute for testing the evidence in a realistic context of use. Tricerion welcomes inquiries from potential clients and will be happy to enable simple trials with such organizations to prove the usability of the SafeLogin solution in the client’s service context.

About the author. *Norman Fraser is one of the founders of Tricerion. He was previously held management positions in a number of high-growth technology companies, including Endava and Vocalis. He holds Masters degrees from Edinburgh and London Universities, and a doctorate in Computational Linguistics from University College London. He has published two books and around 50 scientific papers, many of them focused on human-computer interaction and interface usability.*

Web: www.tricerion.com
www.safelogin.co.uk

References

Adams, A & MA Sasse (1999) Users are not the enemy. *Communications of the ACM* 42(12): 40-46.

Alfahad, FN (1990) The interactive effects of cognitive functioning and visual realism on visual memory recall task. Doctoral dissertation, University of Pittsburgh.

Berry, LH (1991) The interaction of color realism and pictorial recall memory. Paper presented at the Annual Convention of the Association for Educational Communications and Technology.

Besnard, D & B Arief (2004) Computer security impaired by legitimate users. *Computers and Security* 23: 253-264.

Childers, TL & MJ Houston (1984. Conditions for a Picture-Superiority Effect on Consumer Memory. *Journal of Consumer Research* 11: 643-54.

Claburn, T (2005) Workers trade password security for Starbucks. Information Week, May 6.

De Agneli, A, L Coventry, G Johnson & K Renaud (2005) Is a picture really worth a thousand words? Exploring the feasibility of graphical authentication systems. *International Journal of Human-Computer Studies* 63(1-2): 128-152.

Dhamija, R and A Perrig (2000) Déjà Vu: A User Study. Using Images for Authentication, *Proceedings of the 9th USENIX Security Symposium*, August: Denver, Colorado.

Dreyfus, S (2000) Forget your password? Picture this. *The Independent,* October 9.

Ensor, B, M Bennett & E. Giovannini (2004) How consumers remember passwords. Forrester Research.

FDIC (2005). FFIEC guidance: Authentication in an online banking environment. Http:// www.tricerion.com/files/76_FFIEC_strong_authentication_guidance.pdf.

Gadzella, B, H Fullwood, D Ginther & L Kneipp (1991). Differences in Recall of Pictures and Words as a Function of Hemisphericity. Texas Psychological Association Convention. Conference proceedings:1-8.

Jesky, RR (1984) The interactive effects of pictorial representation and cognitive style on a visual recall memory task. Doctoral dissertation, University of Pittsburgh.

Keitz, SM & BR Gounard (1976). Age differences in adults' free recall of pictorial and work stimuli. *Educational Gerontology* 1: 237-241.

Lemos R (2002) Passwords: the weakest link? Hackers can crack most in less than a minute. CNet News.com, May 22.

Nelson, DL, US Reed & JR Walling (1976). Picture superiority effect. *Journal of Experimental Psychology: Human Learning & Memory* 2: 523-528.

Nickerson, RS (1965) Short term memory for complex meaningful visual configurations: a demonstration of capacity. *Canadian Journal of Psychology* 19: 155-160.

Norman, DA (1988) *The psychology of everyday things.* Basic Books.

Paivio, A (1971) *Imagery and verbal processes.* Holt, Rinehart and Winston: New York.

Park DC, JT Puglisi & AD Smith (1986) Memory for pictures: does an age-related decline exist? *Psychology and Aging* 1:11-17.

Ritchey, GH (1982) Pictorial details and recall in adults and children. *Journal of Experimental Psychology* 8(2): 139-141.

Shepard, RN (1967) Recognition memory for words, sentences and pictures. *Journal of Verbal Learning and Verbal Behavior* 5: 201-204.

Standing, L, J Conezio & RN Haber (1970) Perception and memory for pictures: single trial learning of 2560 visual stimuli. *Psychonomic Science* 18: 89-90.

Winograd, ES, AD Smith & EW Simon (1982) Aging and picture superiority effect in recall. *Journal of Gerontology* 37: 70-75.

Woodcock, RW, CR Clark & CO Davies (1969). *Teacher's Guide: Peabody Rebus Reading Program.* American Guidance Service: Circle Pines, MN.