PONTIFÍCIA UNIVERSIDADE CATÓLICA DO RIO GRANDE DO SUL
FACULDADE DE PSICOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM PSICOLOGIA
DOUTORADO EM PSICOLOGIA

A MEMÓRIA HUMANA NO USO DE SENHAS

Tese apresentada ao Programa de Pós-Graduação em Psicologia da Pontifícia Universidade Católica do Rio Grande do Sul como requisito parcial para a obtenção do título de Doutora em Psicologia.

Denise Ranghetti Pilar da Silva

Prof.ª Dr.ª Lilian Milnitsky Stein
Orientadora

Porto Alegre, agosto de 2007.

PONTIFÍCIA UNIVERSIDADE CATÓLICA DO RIO GRANDE DO SUL
FACULDADE DE PSICOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM PSICOLOGIA
DOUTORADO EM PSICOLOGIA

Denise Ranghetti Pilar da Silva

A MEMÓRIA HUMANA NO USO DE SENHAS

COMISSÃO EXAMINADORA

Prof.ª Dr.ª Lilian Milnitsky Stein
Presidente

Prof. Dr. Marcelo Soares Pimenta
Instituto de Informática, Departamento de Informática Aplicada - UFRGS

Prof.ª Dr.ª Lia Buarque de Macedo Guimarães
Escola de Engenharia, Depto de Engenharia de Produção e Transportes - UFRGS

Prof.ª Dr.ª Léa da Cruz Fagundes
Instituto de Psicologia, Laboratório de Estudos Cognitivos - UFRGS

Prof. Dr. Norberto Hoppen
Escola de Administração, Programa de Pós Graduação em Administração - UFRGS

Porto Alegre, agosto de 2007.

Dados Internacionais de Catalogação na Publicação (CIP)

S586m Silva, Denise Ranghetti Pilar da
A memória humana no uso de senhas / Denise Ranghetti Pilar da Silva. — Porto Alegre, 2007.
105 f.

Tese (Doutorado) – Faculdade de Psicologia. Programa de Pós-Graduação em Psicologia. PUCRS, 2007.

Orientador: Profa. Dra. Lilian Milnitsky Stein

1. Memória - Aspectos Psicológicos. 2. Senhas. 3. Repetições Elaborativa. 4. Segurança da Informação. I. Título.

CDD : 153.12

Bibliotecário Responsável
Ginamara Lima Jacques Pinto
CRB 10/1204

AGRADECIMENTOS

Gostaria de registrar minha profunda gratidão a todas as pessoas que, de uma forma ou de outra, contribuíram para a conclusão deste doutorado.

À minha orientadora, Prof.ª Dr.ª Lilian Milnitsky Stein, que sempre demonstrou acreditar no meu potencial, pela oportunidade de trabalhar em seu grupo de pesquisas, pelo incentivo constante, pela orientação no mais pleno sentido da palavra, por compartilhar e celebrar as boas conquistas e por procurar uma saída comigo quando as coisas não correspondiam às expectativas.

Ao Dr. Renato Stein, por me introduzir ao Grupo de Pesquisa em Processos Cognitivos e pelo "empréstimo" da Daniela. À Daniela Benzano, pela imensa ajuda com as análises estatísticas, sempre de bom humor, inclusive às segundas-feiras pela manhã.

Ao Grupo de Pesquisa em Processos Cognitivos, que poderia se chamar Grupo de Pesquisa Modelo, por proporcionar o equilíbrio perfeito entre trabalho sério e ambiente descontraído, tão significativamente produtivo. Obrigada a todos os colegas, trabalhar com vocês me ensinou muito.

Em especial, agradeço ao Carlos Gomes, auxiliar de pesquisa que partilhou comigo as agruras e as glórias deste projeto. À Priscila Brust, pelo ombro amigo, pelo apoio moral, pelos cafés e também pelas revisões, pelo suporte técnico com normas e análises, além da ajuda com a coleta de dados. Ao Gustavo Rohenkohl, Leandro Feix, Renato Santos e Luiza Feijó, se não fosse por vocês, a morte amostral teria se transformado em genocídio. Ao Ronie Silveira e à Rosa Busnello, pelas revisões de manuscritos. Ao Diego Dewes e à Juliana Pureza, ex-auxiliares de pesquisa, pela grande ajuda nas coletas do Levantamento.

À Prof.ª Dr.ª Marilene Zimmer, pela criteriosa revisão da Seção Teórica.

Às instituições e participantes voluntários que viabilizaram o Estudo de Levantamento. Aos professores André Duhá, Patrícia Franzoni e Christian Kristenssen e seus alunos que generosamente nos cederam seu tempo. Aos funcionários do LACE (Laboratório de Informática da FACE-Faculdade de Administração, Economia e Contabilidade da PUCRS), e seu gerente, Vergilio Ricardo Brito da Silva, pelo suporte no uso das salas e computadores.

À Dr.ª Gislaine Baroni, por me ajudar a compreender que não é preciso acertar sempre.

À minha família, meus pais por terem me ensinado a perseverar, e pelo apoio sempre que foi preciso recomeçar; à minha irmã, por ouvir todos os desabafos sem reclamar; aos meus sogros, por, entre outras coisas, todas as horas de baby-sitting que me permitiram coletar dados.

Ao meu marido Julio e aos meus filhos, Lucas e Sabrina, por fazerem a vida valer a pena.

SUMÁRIO

LISTA DE TABELAS

Seção Empírica I

LISTA DE FIGURAS

Seção Empírica I

Seção Empírica II

LISTA DE SIGLAS

ATM	Automated Teller Machine
ANOVA	Analysis of Variance
CI	Confidence Interval
DoD	American Department of Defense (Departamento de Defesa Americano)
FTT	Fuzzy Trace Theory
GC	Grupo Controle
GEP	Grupo Experimental Pista
GER	Grupo Experimental Repetição
ID	Identificação
INSS	Instituto Nacional do Seguro Social
LACE	Laboratório de Administração, Contabilidade e Economia
M	Mean / Média
Mdn	Median / Mediana
Mo	Mode / Moda
MySQL	My Structured Query Language
OR	Odds Ratio
PHP	PHP Hypertext Preprocessor
PUCRS	Pontifícia Universidade Católica do Rio Grande do Sul
RS	Rio Grande do Sul
SD	Standard Deviation / Desvio Padrão
SI	Segurança da Informação
SPSS	Statistical Package for Social Sciences
TTD	Teoria do Traço Difuso

RESUMO

Quantas senhas você precisa lembrar apenas para gerenciar suas informações na Internet? Isso sem falar nas contas bancárias, ou naquelas referentes à vida profissional. De fato, muitas das deficiências dos sistemas de autenticação por senhas se originam das condições de funcionamento da memória humana. Se não fosse necessário lembrar de senhas, elas poderiam, com certeza, ser muito seguras.

A abordagem da Segurança da Informação ao enfrentar os problemas relacionados ao uso de senhas parece se concentrar nos aspectos tecnológicos. Da mesma forma, o usuário tem sido considerado o elo mais fraco na cadeia de segurança, indicando que os fatores humanos acabam comprometendo a segurança que a tecnologia pretende aumentar. Por isso, verificou-se a necessidade de melhor compreender os fatores humanos e buscar alternativas de incluí-los em sistemas de autenticação por senhas.

Dessa forma, a presente tese relata dois estudos empíricos. O primeiro estudo descreve um levantamento realizado com o objetivo de identificar os principais fatores que comprometem a recordação de senhas. Neste estudo foram entrevistados 263 participantes de ambos os sexos, com idades entre 18 e 93 anos, e com escolaridade variando de baixa a superior. Os dados indicaram que, independente da idade e da escolaridade, o número de usos de senhas (em média 5,38 senhas por usuário) é o fator que mais influencia o desempenho da memória para senhas. Assim, usuários com escolaridade mais alta, por possuírem várias senhas, mostraram uma maior tendência ao à confusão no uso de senhas. Ao contrário das expectativas, não foi observado efeito do declínio cognitivo (devido ao envelhecimento) na memória para senhas. Em suma, a necessidade de memorizar senhas seguras, por ignorar as condições naturais de funcionamento da memória humana, gera efeitos de vingança:hábitos que comprometem a própria razão pela qual as senhas são usadas.

Com o segundo estudo buscou-se explorar idéias baseadas na Psicologia Cognitiva visando a melhorar o desempenho da memória no uso de senhas. Através de dois experimentos investigou-se o efeito da repetição elaborativa e do uso de pista como auxílio à recordação de senhas, visando a promover a geração de senhas fortes e recordáveis. O Experimento 1 avaliou o efeito da repetição elaborativa e do auxílio de pista na composição, tamanho e potencial de segurança das senhas geradas, comparadas com um grupo controle. A recordação das senhas foi testada em dois momentos, após 5 minutos e uma semana depois. Os resultados do experimento 1 indicaram que as pessoas tendem a observar somente os requisitos obrigatórios, apesar da instrução. O Experimento 2, buscou avaliar o desempenho da memória após um intervalo maior de tempo decorrido, ou seja, cinco semanas depois da geração da senha. Não foi observado efeito de grupo (experimentais vs. controle) na recordação da senha. Entretanto, foi identificada uma possível variável confundidora, o efeito de espaçamento causado pelo *login* depois de um intervalo de 5 minutos, o que de acordo com os estudos de memória favorece a codificação. Em ambos os experimentos os níveis de recordação foram altos. Ainda, os erros cometidos em tentativas de *login* mal-sucedidas foram criteriosamente examinados e categorizados. Os tipos de erros observados sugerem que muitas vezes os usuários lembram da essência da senha, mas esquecem de detalhes do formato no qual a senha foi criada.

A presente tese conclui considerando alguns dos principais achados relatados nos estudos e suas possíveis implicações. Além disso, visando reduzir a distância entre o mundo da tecnologia e de seus usuários humanos, cuja interação tem sido freqüentemente negligenciada, são apontadas sugestões para futuras investigações, bem como limitações dos trabalhos realizados.

Palavras-chave: senhas, memória, segurança da informação, recordação com pista, repetição elaborativa

ABSTRACT

How many passwords and PINs do you have to remember just to manage your affairs on the Internet? In fact, many deficiencies of password authentication systems arise in consequence of regular working conditions of human memory. If one needed not to remember passwords, they could be very secure indeed.

Traditionally, Information Security approaches to the password problems seem to have focused on the technological aspects. Besides, the user has been considered the weakest link in the security chain, indicating that the human factors end up jeopardizing the security that the technology is supposed to enhance. Hence, it is necessary to better understand the human factors involved in authentication, so that these factors can be accommodated into both new and existing systems.

Therefore, this dissertation reports two empirical studies. The first study describes a survey conducted in order to identify the main factors that hinder password recall. For this study, 263 male and female participants were interviewed. Participants' ages ranged between 18 and 93, and education level ranged from grade school to graduate degree. The results indicated that, regardless of age or educational level, the number of password uses (in average 5.38 passwords per user) is the factor that influences memory performance the most. Thus, better-educated users, for owning more passwords, were more prone to password forgetting and mix-ups. Contrary to the expectations, the effect of cognitive decline (due to the aging process) on password memory was not observed. In sum, the necessity of memorizing strong passwords, by ignoring the natural working conditions of human memory, generates revenge effects: habits that jeopardize the very reason for using passwords.

In the second study, we intended to explore ideas based on Cognitive Psychology aiming at enhancing memory performance in password usage. By means of two experiments, the study investigated the effect of elaborative rehearsal and of cue support as an aid to password recall, with the goal of encouraging the generation of strong and memorable passwords. Experiment 1 evaluated the effect of elaborative rehearsal and of cue support on password composition, length, and security potential, as compared to a control group. Password recall was tested in two occasions, 5 minutes after password generation and after a week interval. The results from Experiment 1 indicate that people tend to observe only the requirements that are somehow enforced, in spite of the instructions. With Experiment 2, we sought to evaluate the memory performance after a longer delay, that is, five weeks after password generation. Group effects (experimental vs. control) on password recall were not observed. However, a possible confound was identified, the spacing effect, caused by the 5 minute login which, according to memory studies, favors the encoding process. In both experiments, recall levels were higher than expected. In addition, the errors from the unsuccessful login attempts were carefully analyzed and categorized. The observed error types suggest that, oftentimes, the users do remember the gist of the password, but forget the details of the format in which the password was coded.

This work concludes with considerations about the main findings described in the two studies, as well as their potential implications. Moreover, in an attempt to bridge the gap between the world of technology and the world of its human users, whose interaction has often been overlooked, we point out suggestions for future investigations and limitations of the reported studies.

Keywords: passwords, memory, information security, cued recall, elaborative rehearsal

NÚMERO DA ÁREA DO CNPq:
7.07.00.00-1 : Psicologia
7.07.02.00-4 : Psicologia Experimental
7.07.06.00-0 : Psicologia Cognitiva

INTRODUÇÃO
O DILEMA DAS SENHAS NA LITERATURA

A memória desempenha um papel tão onipresente na nossa vida diária, mas freqüentemente não a valorizamos. Até que um importante incidente de esquecimento ou distorção de lembranças demanda nossa atenção. Erros de memória vêm fascinando cientistas há bastante tempo. Encontros esquecidos, óculos extraviados e falhas em lembrar-se de nomes ou números se tornaram ocorrências comuns, para adultos que tentam equilibrar as demandas de trabalho e família tendo de, ainda, administrar novas e assustadoras tecnologias (Schacter, 2001). Quantas senhas você precisa lembrar apenas para gerenciar suas informações na Internet? Isso sem falar nas contas bancárias, ou naquelas referentes à vida profissional. Quem sabe dizer, num piscar de olhos, onde seu passaporte está nesse momento? E sua certidão de nascimento? As baterias extras para sua câmera? Esses são apenas alguns exemplos de pequenos incômodos que podem ocorrer no dia-a-dia em conseqüência de lapsos de memória. Devido a essas questões, desde o fim do século XX, a preocupação com problemas decorrentes de falhas de memória cresceu, ocupando, hoje, um lugar importante na nossa sociedade.

O uso de senhas envolve preocupações similares. Donald Norman (1990), em seu livro "*Design of Everyday Things*", chama a atenção para a dificuldade que a maioria das pessoas encontra, ao precisar lembrar de códigos secretos ou senhas. De fato, como mostra a figura 1, sempre que um usuário esquece ou confunde suas senhas, a autenticação falha e, para evitar isso, as pessoas se utilizam de estratagemas, tais como guardar cópias de suas senhas em papel, usar a mesma senha para vários sistemas, ou escolher senhas muito fáceis de serem adivinhadas.

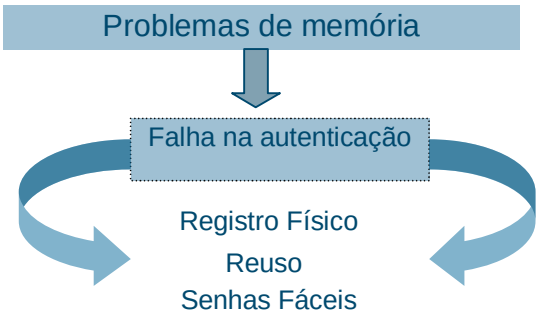

Figura 1

Apesar disso, na “era da Informação” em que vivemos, manipulamos uma quantidade de informações cada vez maior. Por essa razão, a demanda por segurança digital tem crescido, em função das sérias preocupações quanto às conseqüências do uso não-autorizado de informações sigilosas e e senhas vêm sendo usadas para proteger o acesso a quase tudo, de contas bancárias a e-mail. Na verdade, o uso de senhas como forma mais comum de se controlar o acesso a informações privilegiadas, envolve dois mundos com características diferentes e por vezes conflitantes: o mundo da tecnologia e o mundo dos seres humanos, cuja interação tem gerado inúmeros problemas e desafios.

A Segurança da Informação e a sociedade

Segredos e códigos secretos existem desde os primórdios da humanidade. Há registros de escrita codificada já no Egito Antigo, datando de aproximadamente 1900 a.C. (Aranha, 2005). Da mesma forma, as tentativas de decifrar tais códigos são provavelmente tão antigas quanto eles. Pode-se então dizer que, de certa forma, a Segurança da Informação (SI) sempre existiu, embora sua relevância tenha crescido ao longo do tempo, especialmente nos últimos anos. Hoje a SI se tornou um problema importante da sociedade moderna. Desde grandes empresas a indivíduos comuns, todos têm o direito de esperar que seus dados privados sejam preservados e disponibilizados apenas a pessoas autorizadas.

As organizações estão cada vez mais cientes dos riscos de ataques a suas informações privilegiadas, porém os indivíduos comuns, e em alguns casos, até mesmo órgãos do governo, tendem a acreditar que é improvável que eles sejam alvo de ataque. Em 29 de Novembro de

2005, em São Paulo, um estagiário do INSS de apenas 18 anos foi preso e acusado de inserir dados falsos nos sistemas da previdência usando senhas de colegas. Em dois anos o jovem acumulou três milhões de reais. Ele adquiriu seis carros de luxo, equipamentos eletrônicos de alto custo e mobiliou sua casa com móveis de alta qualidade. A polícia conseguiu reaver em torno de dois milhões de reais, mas continua investigando o caso (GloboOnline, 2005). Em 2004, outro estudante brasileiro foi condenado a seis anos de prisão por invadir contas bancárias pessoais do Banco do Brasil, Bradesco, Caixa Federal e Itaú. Ele já tinha sido preso outras duas vezes anteriormente, mas foi liberado por falta de provas (FolhaOnline, 2004).

Assim, para que se possa compreender melhor os riscos de ataque a informações privilegiadas, faz-se necessário examinar alguns conceitos próprios da SI que se referem aos potenciais invasores, assim como seus métodos e ferramentas.

Invasões

Hackers são intrusos não-autorizados que tentam burlar a segurança de uma empresa ou indivíduo, por qualquer que seja a razão. Ataques de *hackers* vêm sendo noticiados com uma freqüência cada vez maior. Atualmente, as vítimas de tais ataques já não se restringem a grandes companhias ou departamentos governamentais. Ao contrário, hoje em dia qualquer indivíduo pode ser alvo desse tipo de ataque. É importante lembrar que existe uma distinção entre *hackers*, também chamados de *white hat hackers* (invasores), que apenas procuram acessar território proibido, sem intenções explícitas de causar prejuízos, e os *crackers* ou *black hat hackers* (invasores criminosos), que são motivados por objetivos criminosos.

Um ataque de *hackers* pode ser descrito, de forma ampla, como adivinhação de senhas em alta velocidade. Assim, a força de uma senha poderia ser medida em termos do tempo necessário para decifrá-la, o que, teoricamente, pode variar de segundos a milênios. Há, basicamente, quatro tipos de ataques usados para descobrir senhas e, surpreendentemente, muitos não se utilizam de tecnologia, mas exploram o comportamento humano (Garancis, 2004; Kuo *et al.*, 2006). O primeiro desses ataques, também conhecido como Engenharia

Social, é um processo no qual o invasor obtém a confiança dos usuários e faz com que estes revelem suas senhas, ou mesmo pistas que facilitem sua descoberta. Em outro tipo de ataque, o *hacker* obtém informações suficientes sobre o usuário, de modo a conseguir descobrir suas senhas. Num terceiro tipo de ataque, o invasor obtém acesso físico ao ambiente onde o usuário digitará a senha e tenta capturá-la por observação, no momento do *login*, ou utilizando tecnologias tais como programas espiões, que gravam as teclas digitadas. No último tipo de ataque, o *hacker* obtém acesso ao sistema a ser atacado e usa programas especiais, muitos deles disponíveis gratuitamente na *Internet*, para decifrar as senhas dos usuários.

A maioria dos ataques realizados com o auxílio da tecnologia se utiliza de dicionários de senhas – também disponíveis na *Internet* – que são, na verdade, listas de palavras ou expressões, usadas no cotidiano. Há dicionários disponíveis em vários idiomas, além de dicionários específicos por domínio, como, por exemplo, os de termos médicos, ou dicionários híbridos, contendo versões ligeiramente modificadas de palavras que os usuários poderiam criar, no intuito de tornar uma senha mais difícil de ser decifrada. Nesse caso, por exemplo, algumas letras são substituídas por números visualmente similares, dígitos são adicionados ao fim da senha, a palavra é digitada de trás para frente, e assim por diante. No outro extremo se encontram os ataques de "força-bruta", nos quais são tentadas exaustivamente todas as combinações possíveis de gerar a senha.

Alguns *hackers* buscam lucro financeiro, outros procuram segredos corporativos, outros ainda estão atrás do fascinante desafio de encontrar a chave para o território proibido das informações confidenciais de outros. Com a capacidade de processamento dos computadores modernos e com programas especiais para decifrar senhas, uma senha composta de seis letras minúsculas - o que significa 308 milhões de combinações - pode ser decifrada por um *hacker*, em média, em dez segundos (Garancis, 2004). O mais surpreendente é que tais programas estão disponíveis gratuitamente na *Internet*.

O mundo tecnológico: segurança da informação

A SI é hoje um grande e oneroso problema para empresas e indivíduos. Atualmente, ao contrário do que acontecia no início da popularização do uso de tecnologia, cidadãos comuns, além de grandes empresas e agências governamentais, também correm o risco de ter suas informações privilegiadas atacadas por intrusos mal-intencionados. A maioria das definições de SI (S. Brostoff, 2004; Morris & Thompson, 1979; Sieberg, 2005; R. E. Smith, 2002; Wikipedia, 2005) pode ser sumarizada como a proteção contra o uso ou acesso não-autorizado à informação, assim como a garantia do acesso a usuários autorizados, enquanto a integridade e a confidencialidade dessa informação são preservadas. O nível de proteção deveria, em qualquer situação, corresponder ao valor dessa informação e aos prejuízos que poderiam decorrer de seu uso impróprio.

Uma das áreas-chave em SI, e talvez um de seus maiores desafios, é a autenticação, ou o processo pelo qual os sistemas distinguem usuários autorizados de outros não-autorizados. A autenticação de usuário é um componente vital de sistemas que possuem informações críticas ou serviços personalizados (Renaud & De Angeli, 2004). Assim, à medida que a oferta de serviços *online*, tais como *online banking* ou comércio eletrônico, cresce exponencialmente, a demanda por proteção de informações críticas vem aumentando na mesma proporção. Além disso, a autenticação via Internet envolve fatores que estão além do controle das equipes de segurança, tais como o equipamento ou o sistema operacional do usuário.

Apesar de sua vulnerabilidade e de serem suscetíveis a problemas de esquecimento e confusão pelos usuários, sistemas de senhas textuais ainda constituem a abordagem mais utilizada para autenticação. Em sistemas desse tipo, em primeiro lugar a pessoa declara sua identidade, por exemplo, com um nome de usuário. Após, revela ao sistema um código secreto ou uma palavra-chave (senha), conhecida, preferencialmente, somente por ela. A autenticação por meio de senhas é popular por ser amplamente aceita pelos usuários. Além

disso, tanto a criação da senha como a autenticação propriamente dita, levam apenas alguns segundos. Ainda, a implementação do sistema é simples e de custo baixo. As vantagens de sistemas de autenticação por senhas decorrem do fato de que estes não requerem equipamento especial, como por exemplo, leitores de impressões digitais ou *scanners* de retina. Além disso, se comprometidos por uma invasão, os objetos de identificação (isto é, nome de usuário e senha), podem ser alterados facilmente e a um custo muito baixo. Por outro lado, enquanto os usuários continuarem esquecendo ou confundindo senhas, a própria finalidade para a qual elas servem será colocada em risco. Assim, vê-se que o papel que as pessoas desempenham na segurança da informação é, pois, um papel importante, e a literatura sobre SI apenas começa a abordar esse tema.

Do ponto de vista da Segurança da Informação, uma boa senha deveria ser, obviamente, segura, o que foi definido por algumas diretrizes publicadas pelo Departamento de Defesa Americano (DoD), em 1985. Além de várias recomendações técnicas para a implementação e gerenciamento de senhas, o documento do DoD forneceu recomendações sobre como os indivíduos deveriam selecionar e administrar suas senhas. Essas recomendações deram origem às seguintes regras (R. E. Smith, 2002):

1. Cada senha escolhida deve ser nova e diferente, já que o uso de uma única senha para vários sistemas pode dar aos invasores uma grande vantagem ao interceptá-la;
2. Senhas devem ser memorizadas. Se uma senha é registrada em papel, este deve ser armazenado em local seguro*;*
3. Senhas devem ser compostas de pelo menos seis caracteres, dependendo do tamanho do conjunto de caracteres usado, i.e., se contêm apenas números, para ser segura a senha precisa ser mais longa, enquanto que se contêm uma combinação de números, letras e outros caracteres do teclado, pode atingir o mesmo nível de segurança com menos caracteres;
4. Senhas devem conter uma mistura de letras (tanto maiúsculas quanto minúsculas), dígitos e caracteres de pontuação;

5. Senhas devem ser substituídas periodicamente.

Do ponto de vista do usuário, entretanto, a autenticação é apenas uma tarefa obrigatória, exercida para obter acesso aos recursos necessários à realização da atividade desejada. Sob essa ótica, uma boa senha deveria, então, ser facilmente disponível, não requerer equipamento especial nem conhecimento técnico, ser conveniente (isto é, não consumir muito tempo) e, acima de tudo, ser fácil de lembrar. Essas motivações, associadas às condições de funcionamento da memória dos seres humanos conflitam diretamente com as recomendações do DoD.

O componente humano em um sistema de segurança de informação

A ironia maior no uso de senhas é que uma senha deveria ser fácil de gerar e lembrar para seu proprietário, mas difícil de ser adivinhada ou decifrada por outras pessoas. Os critérios para gerar senhas fortes fazem com que seja difícil para seres humanos mantê-las na memória, especialmente quando se tem várias senhas para lembrar. Como conseqüência, vários "maus hábitos" já foram identificados e amplamente difundidos (A. S. Brown *et al.*, 2004; Yan *et al.*, 2004). Tais maus hábitos incluem escrever as senhas em papel e armazená-los em locais óbvios, como o monitor do computador ou sob o *mouse pad*, utilizar a mesma senha repetidamente, ou ainda, escolher palavras simples ou nomes que são muito fáceis de adivinhar. Maus hábitos no uso de senhas significam que políticas de segurança que foram cuidadosamente elaboradas, não estão sendo observadas. Na verdade, esses maus hábitos se materializam em vulnerabilidades de sistemas de informação, tais como senhas fracas (e.g., palavras ou datas), senhas repetidas ou senhas visíveis (em papel, armazenadas em local visível).

A comunidade de segurança da informação recentemente deu-se conta de que o comportamento do usuário desempenha um papel importante em incidentes de segurança. Sistemas de segurança da informação são freqüentemente comparados a uma corrente com muitos elos representando os componentes envolvidos, tais como equipamento, *software*,

protocolos de comunicação de dados e outros, incluindo o usuário. Na literatura sobre segurança da informação, o usuário humano é freqüentemente referenciado como o elo mais fraco (Sasse *et al.*, 2002). Entretanto, além de culpar o usuário pouco tem sido feito para identificar-se os fatores que o levam a comportamentos potencialmente inseguros, e menos ainda para tentar resolver tais problemas. Corporações já gastaram milhões de dólares em tecnologias e dispositivos de acesso seguro, recursos que talvez tenham sido desperdiçados, uma vez que os usuários desses sistemas ainda são humanos, com todas as suas limitações humanas e, portanto, ainda o elo mais fraco.

Ao longo dos últimos vinte anos, a abordagem tradicional à segurança da informação tem sido a de tentar solucionar o problema desenvolvendo tecnologias cada vez mais complexas, a fim de proteger informações, tais como protocolos de encriptação ou certificados de segurança. Considerando o crescente número de ataques, pode-se dizer que esse tipo de medidas não parece ser suficiente para garantir que a informação esteja protegida.

Entretanto, fora do mundo tecnológico, pouca atenção tem sido dada a problemas especificamente relacionados ao uso de senhas. Embora periódicos de tecnologia e administração (Ives *et al.*, 2004; Sasse et al., 2002; Sieberg, 2005; R. E. Smith, 2002) tenham tratado de alguns aspectos pragmáticos da segurança de senhas, tais como maus hábitos e perdas de produtividade associadas ao esquecimento de senhas, na literatura psicológica ou da área de Interação Humano-Computador pouco tem sido estudado sobre os aspectos cognitivos da criação, uso e esquecimento de senhas.

O Problema das Senhas

É consenso na literatura que grande parte das deficiências de sistemas de autenticação baseados em senhas está relacionada às condições de funcionamento da memória humana (Ives *et al.*, 2004; Renaud & De Angeli, 2004; Sasse et al., 2002; Yan et al., 2004). No mundo informatizado de hoje em dia, as senhas são usadas para proteger o acesso a todo tipo de

informações. Criar senhas fortes, portanto, é necessário, mas como fazer isso de modo que possam ser lembradas pelo usuário, ainda não está suficientemente estudado.

Dessa forma, partindo do pressuposto de que os usuários simplesmente carecem de motivação para se comportarem de maneira mais segura quanto ao uso de senhas, a abordagem mais comum ao enfrentar tais problemas tem sido a de concentrar esforços no aprimoramento das tecnologias envolvidas na proteção de informações privilegiadas. Para Adams e Sasse (1999), no entanto, já que os mecanismos de segurança são projetados, implementados, aplicados e violados por pessoas, então os fatores humanos deveriam ser considerados seriamente ao se projetar tais sistemas. As autoras ainda apontam que, atualmente, os *hackers* e *crackers* dedicam maior atenção ao papel do usuário humano na segurança do que os próprios profissionais que projetam os sistemas de segurança.

Além disso, quando o número de códigos secretos que uma pessoa precisa armazenar e ser capaz de lembrar aumenta muito, a memória pode falhar. Quando a memória fica sobrecarregada, pode ser muito difícil lidar com a variedade de dados necessários diariamente. Para as atividades cotidianas, necessitamos ter disponíveis, em nossa memória, desde números de telefone, números de contas bancárias, e números de documentos e senhas; sem falar em informações mais pessoais tais como endereços, datas de aniversário, tamanhos de roupas, e assim por diante. Especificamente no caso de senhas, é importante que sejam mantidos em segredo, uma vez que protegem informações confidenciais. Algumas senhas ainda devem ser periodicamente alteradas. Como pode alguém, lembrar-se de tantas senhas? Ao que tudo indica, não é possível.

Na tentativa de administrar as dificuldades de memorização de informações sem significado, muitas pessoas criam algum tipo de registro físico do código secreto, seja eletronicamente, seja em papel. Tal registro físico é, às vezes, disfarçado ou escondido, o que, por sua vez, cria outros problemas, como por exemplo, o de lembrar qual o disfarce usado ou onde o registro físico foi armazenado. Sabe-se que pessoas escondem coisas em lugares

improváveis, mas, em geral, existe uma lógica envolvida na escolha desses lugares. Há evidências que amparam a hipótese de que o esquecimento do disfarce ou do esconderijo pode ocorrer quando há tanto um julgamento de que o local é altamente memorável, quanto um julgamento de que o local é improvável para o objeto (Winograd & Soloway, 1986).

A maioria dos profissionais de tecnologia já se deparou, em algum momento, com problemas de usabilidade que não foram considerados na fase de projeto de seus sistemas , ou que foram, ao menos, subestimados pelos projetistas. Uma vez que o comportamento humano é complexo e envolve variáveis que não podem ser controladas, torna-se difícil, para muitos profissionais da informação, pensar no usuário humano como um componente dos sistemas com que trabalham. Sistemas estes que abrangem mais do que máquinas e métodos para coletar, processar, transmitir e disseminar dados que representam informação para o usuário. Assim, parece bem mais confortável aderir às variáveis que podem, de fato, ser controladas, tais como *hardware* e *software.*

Como conseqüência dos ataques e das inerentes dificuldades em administrar senhas, há uma demanda crescente de recursos das organizações encarregadas de proteger informações. Esses recursos envolvem o trabalho de administradores de sistema e de *help desk* (suporte a usuários), na forma de bloqueio de contas, solicitações de novas senhas ou para sanar os prejuízos com um sistema violado devido ao comprometimento de uma senha.

O Problema das Senhas na Literatura Científica

Há muito poucos estudos na literatura tratando do esquecimento de senhas e dos problemas causados por falhas mnemônicas (A. S. Brown et al., 2004; Dhamija & Perrig, 2000; Yan et al., 2004). A maioria dos estudos relacionados se concentra em outros aspectos que não a memória, como, por exemplo, a efetividade do sistema de senhas (S. Brostoff, 2004), proposta de sistemas alternativos textuais (Jeyaraman & Topkara, 2005; Spector & Ginzberg, 1994; Van Vleck, 1997; Zviran & Haga, 1990), ou, ainda, em outras alternativas de autenticação, tais como senhas gráficas ou baseadas em técnicas biométricas.

Dentre os raros estudos que têm investigado a criação e o uso de senhas, Brown et al. (2004) encontraram a evidência empírica de que dois terços das senhas observadas tinham sido geradas em torno de características pessoais dos usuários, enquanto que a maioria das senhas restantes se relacionava à família, amigos ou relacionamentos amorosos. Nomes próprios e aniversários compunham aproximadamente metade de todas as senhas levantadas. O estudo ainda encontrou suporte empírico para os maus hábitos como escrever as senhas ou reusá-las. Esse estudo de Brown e colegas corrobora achados de estudos anteriores, menos abrangentes, mas que também detectaram alguns maus hábitos. Nesses estudos ainda foi relatado que apenas um pequeno percentual de senhas foi criado de acordo com as diretrizes de segurança. Carstens et al. (2004), por seu lado, encontraram evidências de que indivíduos com oito a onze senhas corriam maior risco de não conseguir lembrá-las. Com a proliferação de *websites* que requerem autenticação, *e-mails* pessoais e profissionais, contas bancárias, etc., possuir múltiplas senhas não é incomum nos dias de hoje.

Na tentativa de superar as falhas das senhas, pesquisadores de segurança têm tentado abordar o problema buscando métodos de autenticação alternativos, tais como técnicas biométricas (e.g., impressões digitais, identificação pela voz, *scan* de retina), senhas descartáveis (para um único uso), ou senhas gráficas (com figuras, fotografias ou faces). Todos visam a reduzir ou eliminar os problemas de memória humana, enquanto preservam ou melhoram a segurança da informação.

Alguns estudos tentaram abordar as deficiências dos sistemas de autenticação por senhas de várias maneiras. Zviran e Haga (1990) inventaram um sistema chamado Senhas Cognitivas, o qual envolve o usuário e o sistema em um diálogo onde o usuário responde a uma série rotativa de cinco questões sobre fatos e opiniões pessoais, escolhidos a partir de uma lista pré-definida de 20 pares de pergunta-resposta. Um conjunto de breves respostas substitui uma única senha. Outro estudo recente (Just, 2004) também explorou a concepção e avaliação de sistemas de perguntas-desafio, embora seu foco fosse a recuperação de

credenciais perdidas em vez da autenticação primária. Senhas cognitivas são uma variação das senhas associativas (S. L. Smith, 1987), uma técnica em que o usuário armazena pares de palavras, de forma que, no momento do *login*, para toda palavra apresentada, ele(a) precisa digitar a palavra associada.

Outra alternativa às senhas tradicionais, são as chamadas *passphrases*. Tratam-se de seqüências de palavras usadas para controlar o acesso a recursos computadorizados (Porter, 1982). Uma *passphrase* é similar a uma senha quanto ao uso, mas é geralmente mais longa para oferecer uma segurança adicional. Uma alternativa a longas *passphrases* poderia ser então, a de utilizar-se de técnicas mnemônicas, criando senhas baseadas em frases, contendo a primeira letra de cada palavra, ou, ainda, substituindo uma palavra por um número ou caractere especial. Foi demonstrado que senhas mnemônicas podem ser tão seguras quanto senhas aleatórias e tão memoráveis quanto simples palavras (Yan *et al.*, 2004). Contudo, Kuo e associados (2006) sugeriram que eficientes dicionários mnemônicos poderiam ser criados, uma vez que a maioria dos usuários escolhe frases oriundas de fontes pré-existentes, tais como filmes, livros ou músicas, em vez de criar suas próprias frases. Assim, dada a crescente capacidade dos computadores, as senhas mnemônicas poderiam ser facilmente decifradas.

Ao mesmo tempo, no intuito de explorar as características naturais da memória visual, e sabendo que o desempenho da memória é superior no reconhecimento que na recordação (Schwartz & Reisberg, 1991), foi criado o conceito de senhas gráficas. Além disso, a capacidade da memória visual humana é imensa, e esta parece menos vulnerável ao esquecimento do que os outros tipos de memória (Madigan, 1983). Um dos principais problemas com sistemas de senhas gráficas é o tempo necessário para a criação de senhas, assim como o tempo de autenticação propriamente dita (*log in*), em usos subseqüentes. Outro problema diz respeito aos casos onde o usuário esquece sua senha, já que não é possível enviar um “lembrete” por *e-mail*, assim como acontece com as senhas textuais. Nesses casos é necessário cadastrar uma nova senha. Ainda, pessoas com problemas visuais, como, por

exemplo, o daltonismo, ou com problemas motores que dificultem o uso preciso do *mouse*, teriam dificuldades para usar senhas gráficas (Wiedenbeck *et al.*, 2004). Finalmente, há que se considerar os possíveis efeitos de interferência, no caso do usuário possuir várias senhas desse tipo, onde novas senhas podem ser confundidas com as já existentes, e vice-versa, uma vez que.a evidência consistente de interferência entre imagens similares já foi observada (Radvansky & Copeland, 2006) .

Outra alternativa de autenticação, e que vem conquistando muitos adeptos, inclui os sistemas de autenticação baseados em biometria. Esses sistemas envolvem alguma forma de dados obtidos a partir da fisiologia do usuário, tais como *scan* de retina, impressões digitais, padrões de voz e assim por diante. Embora convenientes por estarem sempre com o usuário (ao contrário de objetos identificadores ou *tokens*) e por não estarem sujeitos às limitações de memória (como as senhas textuais), tais sistemas não podem ser alterados em caso de comprometimento (e.g., por ataque), o que limita o seu uso. Também não se conhecem as taxas de erro para as várias tecnologias biométricas, i.e., em que proporção esse tipo de autenticação é confiável, nem como estas se comparam às taxas de erro humano. Existe ainda uma preocupação quanto às questões de privacidade, ligadas ao uso de técnicas biométricas de identificação, que poderiam gerar muitos problemas, se a tecnologia fosse usada para vigiar e monitorar amplamente os cidadãos em seu dia-a-dia.

Cada uma das alternativas a senhas textuais mencionadas acima envolve decisões e renúncias, que vão desde o custo ao tempo de autenticação, passando por questões de privacidade e por outras considerações especiais, que incluem o local de uso, equipamento, a habilidade de alteração ou sua substituição em caso de comprometimento. Além disso, limitações físicas (e.g., não possuir o dedo polegar), considerações de saúde (e.g. perigo de contaminação por uso de equipamento em local público) e confiabilidade também devem ser levadas em conta, na implementação de procedimentos de segurança. Indiscutivelmente, tais alternativas parecem promissoras, principalmente se aplicadas aos sistemas adequados. No

entanto, pelas razões mencionadas acima e com o suporte da literatura (Ives et al., 2004; Kuo et al., 2006), acreditamos que as senhas textuais ainda continuarão sendo usadas por muitos anos. Em suma, muitas técnicas promissoras foram inventadas e testadas, sem que, no entanto, alguma delas tenha conseguido superar a conveniência e o baixo custo das senhas tradicionais.

A Memória e as Senhas

Sabe-se que as memórias não são armazenadas de forma integral e, mesmo quando estabelecidas e consolidadas, não são permanentes (Schacter, 2001). Por esta razão, somos melhores na generalização e na abstração de conhecimentos do que na retenção do registro literal de eventos. O esquecimento ocorre continuamente, enfraquecendo o traço de memória do que foi aprendido. De fato, esquecer é uma função essencial ao bom funcionamento da memória: seria impossível, e pouco prático, lembrar com riqueza de detalhes todas as informações que já vivenciamos. Quando se fala de senhas, o que se coloca como limitação é, na verdade uma característica valiosa do ser humano, que lhe permite pensar, conviver, sobreviver e se adaptar: o esquecimento.

Sabe-se que muitas das deficiências dos sistemas de autenticação por senhas se originam das condições de funcionamento da memória humana. Assim, os maus hábitos, como escrever ou reusar a senha, bem como as falhas de memória, acontecem simplesmente porque, na impossibilidade de memorizar suas senhas, as pessoas desenvolvem estratégias não seguras. Os estudos da Psicologia Cognitiva que têm pesquisado o funcionamento da memória, têm mostrado consistentemente que:

- guardar informações literais, ou detalhes superficiais como a exata ordem em que os caracteres aparecem em uma senha, é uma tarefa cognitivamente difícil e suscetível a falhas (Reyna & Brainerd, 1995);
- as pessoas tendem a ter facilidade de lembrar de informações em que o significado esteve envolvido na codificação, especialmente se combinado com a presença de pistas

compatíveis na hora do teste de recordação (S. C. Brown & Craik, 2000) - o que geralmente não é o caso das senhas aleatórias ou geradas pelo sistema;

- com a falta de uso e a passagem do tempo, traços literais, como a estrutura da senha ou a fonte, tendem a se perder;
- o fato de processar informações de natureza similar interfere em seu registro mnemônico (F. N. Dempster & Brainerd, 1995; Pergher & Stein, 2003), acarretando perda de parte ou de toda a informação.

Memória e Esquecimento no Uso de Senhas

Pela sua importância na autenticação, os aspectos literais de uma senha, como por exemplo, a fonte (i.e., o sistema ao qual a senha se refere) ou a estrutura da senha desempenham um papel fundamental. Parece-nos claro, então, que as senhas são um tipo especial de memória literal, pois precisamos lembrar os caracteres exatos que as compõem, na ordem exata, além do sistema em particular ao qual uma determinada senha se refere, para que o acesso seja liberado. O esquecimento de um único caractere, ou a atribuição de uma senha ao sistema errado, significa acesso negado.

No intuito de explicar, em termos cognitivos, como as senhas são criadas, memorizadas e recordadas, é interessante aludir aos estudos da memória humana e do esquecimento que vêm sendo contemplados na Psicologia há mais de um século. Ao longo desse tempo, várias teorias foram desenvolvidas, na tentativa de esclarecer a estrutura e o funcionamento da memória. Assim, existe certo consenso no que se refere aos três estágios que compõem o que costumamos chamar de memória, quais sejam: (a) codificação, isto é, o momento em que uma senha recém criada é registrada na memória; (b) armazenamento, o processo pelo qual uma senha é guardada na memória enquanto não é utilizada, e (c) recuperação, ou o momento em que a senha é resgatada da memória para ser utilizada, como, por exemplo, no caixa-eletrônico (Schwartz & Reisberg, 1991). Da mesma forma que em outras atividades que envolvem o uso da memória, não se pode afirmar o estágio no qual ocorreram os problemas de esquecimento ou confusão no uso de senhas.

O esquecimento, ou o fenômeno pelo qual as informações que já estiveram na memória deixam de estar disponíveis, tem sido explicado, historicamente, por várias teorias. Estas teorias podem ser classificadas, de forma ampla, em três grupos: (1) as teorias que postulam que as memórias, com a passagem do tempo, enfraquecem e desaparecem gradualmente, até serem completamente extintas (deterioração); (2) as teorias que pregam que, na verdade, a informação não se perde, estando em algum lugar na memória, mas não podendo ser acessada, embora esteja intacta (falha de recuperação) e (3) as teorias que consideram o esquecimento uma conseqüência do efeito de novas memórias sobre outras já existentes ou vice-versa (interferência). Atualmente as modernas teorias de interferência explicam o esquecimento em termos dos processos responsáveis, tanto pela deterioração quanto pela falha de recuperação.

Várias teorias de memória poderiam ser consideradas no estudo do uso de senhas, como, por exemplo, a teoria dos Níveis de Processamento (Craik & Lockhart, 1972), que propõe que as pessoas usam níveis diferentes de elaboração ao processar informações. Isso se dá num *continuum* que se inicia na percepção, passa pela atenção, rotulação e, finalmente, chega à significação. O ponto chave é que todos os estímulos percebidos são armazenados na memória, mas diferentes níveis de processamento (do mais superficial ao mais elaborado) contribuem na habilidade para acessar ou recuperar-se aquela informação. Assim sendo, informações processadas de forma mais elaborada são, em geral, melhor lembradas. Já a Teoria da Ativação e Monitoramento da Fonte (Johnson *et al.*, 1993), postula que a localização da fonte de um evento é essencial para uma memória precisa da experiência vivida. No caso das senhas, a fonte corresponde ao sistema ao qual a senha em questão se refere, o qual é essencial para o sucesso do processo de autenticação. Há ainda a Teoria do Traço Difuso (TTD) (C.J. Brainerd & Reyna, 1993), que classifica as memórias quanto ao tipo de representação: literal (detalhes superficiais e específicos) ou de essência (significado), as quais são armazenadas paralelamente e de forma independente para um mesmo evento. As

memórias literais e de essência estão sujeitas a processos de esquecimento (como sendo a desintegração gradual dos traços) de natureza diferenciada, onde as taxas de esquecimento são mais altas para representações literais do que para as de essência. Os caracteres exatos de uma senha, na ordem e formato exatos, são informações literais. Então, embora não se possa classificar as senhas como informações puramente literais (já que é possível e comum associar significado às senhas), parece adequado considera-las como um tipo especial de memória literal.

A codificação de memórias literais parece ser comparável à codificação de seqüências de palavras não-relacionadas (Anderson & Paulson, 1977). Isto significa que a representação literal parece ser de natureza perceptual, pois envolve a codificação da ordem em que os caracteres ou palavras aparecem, bem como a codificação de sua fonte. Essa representação é mais forte imediatamente após sua aquisição, mas vai enfraquecendo à medida que o tempo passa. Os resultados encontrados por Anderson e Paulson confirmam a decrescente disponibilidade de informações literais após um intervalo. Eles também observaram que os traços que sobrevivem à perda inicial ficam menos vulneráveis à ação do tempo. Sabe-se ainda que memórias literais são mais vulneráveis aos efeitos de interferência e que se tornam inacessíveis mais rapidamente do que as memórias de essência (Reyna & Brainerd, 1995).

Assim, uma vez que a codificação de informações literais é suscetível à interferência do ambiente, algumas informações acabam não sendo codificadas (Titcomb & Reyna, 1995). Por isso, é mais provável que as pessoas se baseiem em representações literais imediatamente após uma experiência, porém, com a passagem do tempo, as representações de essência se tornam mais disponíveis. Como exemplo, é possível que o usuário lembre que a senha se refere ao aniversário da irmã, mas não consiga lembrar se usou o ano de nascimento com quatro ou seis dígitos, o mês por extenso, ou pior, se era o aniversário da irmã mais velha ou da mais nova.

Memória e Repetição

Os pesquisadores de memória têm observado, desde os estudos pioneiros de Ebbinghaus há mais de um século, que repetir a informação melhora a memória para aquilo que foi repetido (Schacter, 2001). Embora não seja possível dissociar completamente os processos de aquisição, armazenamento e recuperação de memórias, sabe-se que a maneira pela qual um material é codificado na memória pode afetar sua recuperação (Schwartz & Reisberg, 1991). McDermott e Chan (2006) observaram que o uso de repetição na codificação (fase de aquisição de informações) leva a probabilidades mais altas de recordação precisa no futuro. No entanto, repetições mecânicas, nas quais a pessoa apenas repete o estímulo em silêncio para si mesma, não produzem efeitos duradouros na aprendizagem, ao passo que repetir enquanto se pensa sobre o significado do material a ser evocado, ou estabelecendo conexões com informações já aprendidas (repetição elaborativa), leva a uma melhor memória (Craik & Lockhart, 1972).

O caráter das senhas seguras, que deveriam ser conjuntos de caracteres aleatórios e sem sentido (contendo letras, números e símbolos), torna a repetição elaborativa difícil. Porém, se uma associação com informações significativas for criada, pode ser que a repetição elaborativa de uma senha, no momento de sua geração, tenha um impacto em sua codificação na memória. Afinal, segundo Ebbinghaus (1885), lembrar itens sem sentido é mais difícil que lembrar materiais significativos.

Recordação com Auxílio de Pista

A recordação livre permite recordar a informação em qualquer ordem e possibilita observar o tipo de organização que os indivíduos realizam para sua recuperação. No caso da recordação com pista, o processo de recuperação é ativado pelos traços do estímulo fornecidos ao indivíduo. Nesta situação, verificamos que as pistas assumem um papel facilitador do processo de recuperação da informação processada (Lockhart, 2000). Na verdade, vários estudos já mostraram que informações que pareciam perdidas podem ser

recuperadas com o auxílio de pistas que nos lembram de como codificamos a experiência inicialmente (Schacter, 2001).

A eficácia do auxílio de pista na recordação de senhas já foi observada em um estudo pioneiro de Lu & Twidale (2003), depois replicado com resultados similares por Hertzum (2006). O estudo partiu do pressuposto que revelar alguns caracteres da senha, cuidadosamente escolhidos, bastaria para ativar a memória dos usuários legítimos para esta senha, mas não seria suficiente para um invasor mal-intencionado. Para ambos os estudos, entretanto, o objetivo era melhorar a experiência do usuário ao autenticar-se, e a segurança ficou em segundo plano. Os participantes realmente lembraram-se mais das senhas acessadas com as pistas do que daquelas sem elas, porém muitas das senhas criadas eram fracas ou previsíveis.

APRESENTAÇÃO DA TESE

A presente tese de Doutorado partiu do interesse em promover uma aproximação entre os campos de estudo da memória e da segurança de sistemas, isto é, promover práticas seguras de uso de senhas sem gerar-se uma sobrecarga para a memória dos usuários. O interesse principal deste trabalho foi identificar os principais fatores que comprometem a recordação de senhas, e com isso derivar subsídios para motivar uma melhora na usabilidade de senhas como técnica de autenticação. Ao identificar tais fatores, buscamos promover uma aproximação viável entre a tecnologia de segurança da informação e as condições cognitivas de seus usuários humanos. Buscou-se também investigar os aspectos mnemônicos envolvidos na aquisição, armazenamento e recuperação de senhas, através da testagem dos efeitos de repetição elaborativa e de recordação com auxílio de pista na geração e memorabilidade de senhas.

A presente tese foi estruturada em forma de artigos científicos a serem submetidos para publicação. A primeira parte (Introdução) traz uma revisão da literatura multidisciplinar envolvendo o uso de senhas e seus problemas relacionados, na qual são discutidos tanto os

aspectos tecnológicos quanto fatores humanos. Após, é apresentada uma breve descrição de soluções alternativas às senhas, as quais foram propostas ao longo do tempo. Posteriormente, é feita uma revisão dos estudos de memória relevantes ao uso de senhas.

A segunda parte da presente tese de doutorado consiste na apresentação desta pesquisa. O primeiro estudo, "*The Revenge Effects of Passwords*" (Os Efeitos de Vingança das Senhas) descreve o levantamento realizado com o objetivo de identificar as características das populações que encontram dificuldades para lembrar-se de suas senhas. Neste estudo foram entrevistados 263 participantes de ambos os sexos, com idades entre 18 e 93 anos, e com escolaridade variada. Os dados indicaram que o número de usos de senhas é o fator que mais influencia o desempenho da memória para senhas. Assim, usuários com escolaridade mais alta, por possuírem várias senhas, mostraram uma maior tendência ao esquecimento e à confusão no uso de senhas. Ao contrário das expectativas, o declínio cognitivo devido ao envelhecimento não parece ter tido efeito na memória para senhas.

Com o segundo estudo, intitulado "*Exploring Cognitive Psychology to Help Password Security*" (Explorando a Psicologia Cognitiva para Auxiliar na Segurança de Senhas) buscou-se explorar idéias baseadas na Psicologia Cognitiva visando a melhorar o desempenho da memória no uso de senhas. O estudo investigou o efeito da repetição elaborativa e do uso de pista como auxílio à recordação de senhas, visando a promover a geração de senhas fortes e recordáveis através de dois experimentos. O Experimento 1 avaliou o efeito da repetição elaborativa e do auxílio de pista na composição, tamanho e potencial de segurança das senhas geradas, comparadas com um grupo-controle. A recordação das senhas foi testada em dois momentos: após 5 minutos e uma semana depois. Uma vez que a maioria dos usuários teve sucesso em lembrar-se das senhas, um segundo experimento foi realizado para que se pudesse examinar as causas do bom desempenho da memória no Experimento 1, tanto nos grupos experimentais quanto no grupo controle. O Experimento 2 então, buscou avaliar se o bom

desempenho permaneceria após um intervalo maior de tempo decorrido, ou seja, cinco semanas depois da geração da senha.

Por último, são apresentadas as considerações finais, com base na revisão da literatura e nos resultados dos estudos que sintetizam os principais achados desta pesquisa, bem como suas possíveis implicações. Além disso, são apontadas sugestões para futuras investigações e limitações dos trabalhos realizados.

OBJETIVOS

Como sistema de autenticação, a utilização de senhas tem se revelado um meio prático e de baixo custo para diversos serviços (e.g., identificação bancária ou acesso a contas de *e-mail*). Contudo, o seu uso depende de capacidades cognitivas por parte do usuário que vão de encontro às exigências de segurança, gerando conseqüências como o esquecimento da própria senha. Essa limitação, apesar de conhecida por administradores de rede e organizações, tem sido pouco estudada.

Tendo em vista que a abordagem da SI ao enfrentar os problemas relacionados ao uso de senhas parece se concentrar nos aspectos tecnológicos, verificou-se a necessidade de melhor compreender os fatores humanos e buscar alternativas de incluí-los em sistemas de autenticação por senhas. Dessa forma, o objetivo desse trabalho abrange dois campos de estudo. Em primeiro lugar, busca conhecer melhor a realidade dos usuários através da identificação dos principais fatores que comprometem a recordação de senhas, incluindo setores da população que não haviam sido estudados anteriormente, tais como pessoas de baixa escolaridade, que constituem a maioria da população mundial, e pessoas de idade mais avançada, já que o envelhecimento da população é uma realidade em muitos países, incluindo o Brasil. O segundo campo de estudo diz respeito aos aspectos mnemônicos, investigados através da testagem do impacto de duas variáveis mnemônicas - repetição elaborativa e recordação com auxílio de pista - na geração, codificação, armazenamento e recuperação de senhas.

PROBLEMAS E HIPÓTESES

Muitas das deficiências dos sistemas de autenticação por senhas se originam das condições de funcionamento da memória humana. Se não fosse necessário lembrar-se de senhas, elas poderiam, com certeza, ser muito seguras (isto é, totalmente aleatórias e tão longas quanto as limitações do sistema permitissem), e conter todos os tipos de caracteres. Apesar disso, Brown et al. (2004) apontam que a literatura é bastante escassa ao fornecer procedimentos claros, passo a passo, que auxiliem na geração e recordação de senhas. A maioria dos poucos artigos existentes não leva em consideração as limitações cognitivas impostas pela natureza humana. Assim, as pessoas são obrigadas a conviver com um dilema entre a segurança e a conveniência.

Para que se possa abordar o dilema das senhas de forma eficiente, se faz necessário conhecer os principais fatores que comprometem a recordação de senhas. Assim, é importante investigar se variáveis como sexo, idade e escolaridade, que tradicionalmente impactam outras tarefas diárias, têm efeito na memória para senhas. Ainda, sabendo-se que a sobrecarga de informações acarreta um ônus para a memória, faz-se necessário incluir o número de senhas que o indivíduo possui entre as variáveis investigadas. Esperávamos encontrar mais problemas de memória entre participantes mais velhos e, talvez, entre os de escolaridade mais baixa. Examinamos também a possibilidade de encontrar diferentes níveis de esquecimento e confusão de senhas entre homens e mulheres. Além disso, com base em estudos anteriores (A. Adams & Sasse, 1999; S. Brostoff & Sasse, 2003; Carstens et al., 2004; Dhamija & Perrig, 2000), esperava-se que o número de usos de senhas influencie negativamente o desempenho da memória.

É possível que boa parte das senhas esquecidas ou confundidas não tenha sido bem codificada e, por isso, seja vulnerável ao tempo e à interferência de outras senhas. Além disso, muitas vezes o usuário lembra-se de uma senha, mas usa-a com o sistema errado. Isto é um claro exemplo em que se desfez a conexão entre a fonte da informação (o sistema) e o evento

original (nesse caso, a senha). Sendo a fonte uma informação literal, é esperado que seja esquecida mais rapidamente (Reyna & Brainerd, 1995). Titcomb e Reyna (1995) apontam que o teste imediato e o questionamento repetido após um intervalo fortalecem as memórias literais, o que leva a crer que seria interessante investigar se a repetição elaborativa de uma senha, no momento de sua geração, terá um impacto na qualidade da senha gerada, bem como em sua codificação na memória.

Uma outra possibilidade pouco explorada na literatura se refere a possibilidade de se reduzir os problemas de esquecimento e confusão de senhas com a apresentação sistemática de uma pista semântica. O uso de senhas para autenticação se constitui num claro exemplo de teste de memória, mais especificamente, um teste de recordação livre. Segundo a Teoria do Traço Difuso, em testes de recordação livre, em primeiro lugar, tenta-se acessar diretamente às informações literais, e, quando esse processo falha, passa-se a tentar reconstruir o evento a ser lembrado, a partir das informações de essência disponíveis. A pesquisa em Psicologia Cognitiva indica que o auxílio de pista aumenta as probabilidades de acesso direto aos traços de memória em comparação com a recordação livre (1995) (C.J. Brainerd *et al.*, 2002). Por essa razão, parece interessante estudar os efeitos do uso de pista como auxílio na recordação de senhas.

CAMPO DE PESQUISA

As entrevistas do levantamento foram realizadas individualmente nos locais de trabalho, estudo ou lazer dos participantes, em Porto Alegre, RS, Brasil.

Os estudos experimentais foram realizados nos laboratórios de Informática das Faculdades de Administração e Psicologia da PUCRS, respectivamente, em Porto Alegre, RS, Brasil.

METODOLOGIA DE PESQUISA

Estudo de Levantamento

Foram realizados dois estudos preliminares, tendo como base o estudo de Brown e colegas (2004) com o objetivo de gerar um instrumento que permitisse pesquisar o uso de senhas em nossa realidade, para que, a partir dos dados levantados, fosse possível identificar as implicações do uso de senhas no desempenho da memória humana. Um estudo de levantamento preliminar foi realizado com a finalidade de adaptar o instrumento de Brown e colegas (2004) para a realidade brasileira. Na seqüência, através de um levantamento piloto, buscou-se testar o instrumento traduzido e adaptado. O instrumento resultante (Anexo A), (aprovado pelo Comitê de Ética da PUCRS, ofício no. 446/06 – CEP - Anexo B) foi utilizado no Estudo de Levantamento.

No estudo de levantamento participaram 263 pessoas de ambos os sexos, com idades entre 18 e 93 anos e níveis de escolaridade de baixo a superior. As entrevistas foram realizadas individualmente, de forma dirigida. Com o pacote estatístico SPSS, foram realizadas análises descritivas e de freqüência, correlações e, por fim, uma análise de regressão logística, incluindo os fatores que mais pareciam influenciar os problemas de memória com senhas.

Estudos Experimentais

Os estudos experimentais realizados integram dois experimentos cujas amostras foram compostas por estudantes universitários, distribuídos aleatoriamente em três grupos: controle (GC), experimental-pista (GEP) e experimental-repetição (GER).

O material utilizado consistiu de páginas *web* para cadastro, onde eram solicitados os dados demográficos e a identificação do participante através de seu nome de usuário, bem como a digitação de uma nova senha, e sua confirmação, ou *login*, onde o indivíduo deveria digitar seu nome de usuário e senha experimental. Por ocasião do cadastro, ao GEP também era solicitada a criação de uma pista (palavra ou expressão curta), a qual lhes era apresentada

nos *logins* posteriores. A pista deveria ser suficiente para ajudá-los a lembrar a senha, mas não para revelá-la a intrusos. Ao GER, foram apresentadas quatro páginas subseqüentes para a confirmação adicional da senha.

Os dados foram organizados em um Banco de Dados, com o Programa Microsoft Excel e analisados através do pacote estatístico *SPSS*. A distribuição dessas variáveis foi descrita como média e desvio padrão ou freqüência e proporção, quando cabíveis. Foram computados, para cada sessão do experimento, os acertos por participante, os tipos de erros e o número de tentativas necessárias para acertar a senha. No primeiro encontro, foi também analisado o grau de segurança da senha gerada. Nas comparações entre encontros, foram realizadas análises de variância com medidas repetidas. Todos os tratamentos estatísticos utilizaram um α=0,05 para o teste das hipóteses.

SEÇÃO EMPÍRICA I

THE REVENGE EFFECTS OF PASSWORDS

ABSTRACT

Contrary to what most people would expect, natural cognitive decline caused by age doesn't seem to be of major impact on memory for passwords. A survey conducted to identify characteristics of the populations that face the most difficulties in handling passwords revealed that password forgetting and mix-ups happen more frequently as the number of password uses increases. Thus, better-educated users, by owning several passwords, were more prone to forget and/or mix them up. The survey data showed a positive correlation between the number of password uses and memory failures such as forgetting and confusion, suggesting that, as far as memory for passwords is concerned, interference effects are more significant than age. The sample consisted of 263 male and female volunteers, aged 18 to 93, from a variety of educational levels. Our data indicated that the number of password uses is the factor that impacts memory performance the most and the vast majority of respondents - across age groups and literacy levels - used passwords mostly to access their bank account(s). Most passwords were generated by the user, as opposed to the system. The average password length was 4.89 characters long, the majority being numeric only. Over half of all respondents admitted keeping physical records of at least one of their passwords. The strategies used by the subjects in order to memorize and/or remember their passwords, for the most part, go against the most basic security recommendations.

Keywords: passwords, human memory, information security

Introduction

Ironic unintended consequences of technology, that induces behavior which appears to cancel out the very reason for using it, are called *revenge effects* (Tenner, 1997). However, technology alone does not produce revenge effects. It is only when it is anchored in laws, regulations, customs and habits that an irony reaches its full potential. Password security is an example of such a technology, anchored in security guidelines that conflict with humans' customs, habits, and, worst of all, cognitive capabilities.

If you are like most regular people, you'll find it hard to enumerate, without thinking, all of your passwords. This is probably a frustrating exercise, since we often have so many of them that it becomes hard to remember them all without a conscious effort. When talking about passwords, almost everybody has a story to tell, or a complaint to make, people struggle to handle their passwords and often face problems due to the failure to remember the right password at the right time. It seems that when people try to keep track of their passwords, is when revenge effects arise, such as physical records or password reuse. In spite of this, little is known about who among us faces the most difficulties in handling passwords and what makes the revenge effects more evident.

How do we remember passwords? How many we forget? How many expired, are outdated and useless? And how can we get rid of them, since apparently we cannot forget on demand? What are the consequences of forgetting the ones that are still valid? Despite the many unanswered questions, the well-known flaws and the several alternatives created in the recent years, like for instance, graphical passwords or biometric authentication, textual passwords are still the nearly universal authentication mechanism, since they are very convenient. Both the processes of creation and authentication, take only a few seconds, system implementation is relatively simple and the cost is low.

The notion of a secure password has evolved over time, in response to the ever more sophisticated hacker attacks. In the beginning, a good password should only be remembered

and kept secret. Today, however, a good password has become something very difficult to remember, since it should be as safe against ill intentioned intruders as possible. Good passwords should be as long as possible, and contain upper and lower case letters, numbers, and other special characters (A. S. Brown et al., 2004). Besides, even good passwords should be replaced from time to time.

From a security standpoint, the rules are not essentially wrong. However, they do not take into account users' motivations, behavior, and cognitive capabilities. It is not surprising, then, that most users find it difficult to generate and remember secure passwords, especially when there are many to be remembered in their day-to-day activities.

The strength of a password is a function of the password space, which is calculated based on the character sets used in the password. But people tend to be biased by their preferences and knowledge when choosing passwords, what reduces the actual attack space, since not all possible combinations are equally likely. For instance, considering the four-digit ATM passwords, it is known that people find it relatively easy to remember dates and that dates are easily coded as four digit strings. Therefore, many people use dates for their four-digit passwords. This means the attack space is then reduced to the combinations that accommodate dates. If the potential attacker knows the potential victim or has access to his/her personal information, the attack space could be further reduced. In places with elevated theft and robbery rates, a person's stolen wallet or purse may prove itself a rich source for password guessing, for example, to access the person's bank account.

Although the literature is scarce, there are some aspects of password use that have been consistently found, such as the *bad habits* arisen as coping strategies, or the consensus that many of the deficiencies of password systems are related to human memory limitations.

The so-called bad habits in password generation and use include short passwords consisting of names, nicknames, addresses and birth dates, from the user himself or of relatives, or even simple words from the dictionary. The two most frequent and most

preoccupying bad habits, though, are the reuse of the same password for multiple accounts and the storage of the password in paper or electronic media (A. S. Brown et al., 2004; Morris & Thompson, 1979). In fact, cracking one password may reveal to the attacker more than the password itself. It may also reveal the person's strategy for creating passwords and thus lead to the cracking of different passwords from the same user, maybe even some more critical ones.

Most of the solutions proposed for password problems can be roughly classified in three categories: (1) pro-active measures, which identify and refuse weak passwords at generation time; (2) security technologies, such as cryptography, and (3) user training and education. None of them, however, addresses memory limitations (Dhamija & Perrig, 2000) and all have their own shortcomings. The pro-active measures of increasing the enforcement of password guidelines may drive users towards writing their "strong" passwords down, as explicitly stated by one of Adams and Sasse's (1999) participants as well as by one of ours. Security technologies, although good and necessary, are not immune to Social Engineering (i.e., the practice of obtaining confidential information by manipulating users) attacks, and user education, though fundamental, is not sufficient to make users fully compliant with security guidelines.

Cognitive Aspects of Password Authentication

One of the main components of password authentication systems based on knowledge, such as passwords, is human memory. Actually, one of such systems' main weaknesses stems from their need for the precise recall of the secret information (Dhamija & Perrig, 2000). Unfortunately, precise recall (without cues) is not the strength of human cognition. In password use, the tiniest mistake, for instance, forgetting one character or mixing-up the order of two characters, is enough for authentication failure.

In human memory research, there is a rich knowledge base, which may be used to explain the problems with password forgetting. The most important findings, as far as

passwords are concerned, have already been enumerated by Sasse, Brostoff and Weirich (2002). The human limitations at precise recall generate an irreconcilable conflict with the requirements for secure passwords, since users often prefer to decrease the memory burden at the expense of security (Wiedenbeck et al., 2004).

All the above mentioned bad habits, as well as the memory lapses in password use, happen simply because, unable to memorize all their passwords, people resort to developing non-secure strategies. Over time, Cognitive Psychology studies have consistently shown that: (a) the storage of verbatim information or superficial details, such as the exact order in which characters occur in a password, is a difficult task (Reyna & Brainerd, 1995); (b) people tend to be able to remember meaningful information easily (S. C. Brown & Craik, 2000) – what generally is not the case with system generated or random passwords; (c) with lack of use and the passage of time, verbatim traces, such as the password structure or its source, tend to decay more rapidly than its meaning (C. J. Brainerd *et al.*, 1995); (d) the processing of similar information interferes with the mnemonic record of these information, causing the information to be lost in part or in full (F. N. Dempster & Brainerd, 1995; Pergher & Stein, 2003).

The vast body of knowledge in the psychology of memory should be considered by the information security industry in their efforts to make password authentication more viable. Although most of the existing studies on password usage and practices have not considered cognitive theories, they do point out interesting facts about the use of passwords.

Related Work on Password Usage and Practices

The existing literature on password selection and memorization criteria is surprisingly scarce. Several studies have addressed different aspects of password usage and practices, but most of them focused on the passwords themselves instead of their users with their inherent limitations (A. Adams & Sasse, 1999; Morris & Thompson, 1979; Riley, 2006). Three

studies focusing on the user were found and they were used as a foundation to this work (A. S. Brown et al., 2004; Carstens et al., 2004; Zviran & Haga, 1999).

Back in (1979), Morris and Thompson examined a large password database (3289 passwords), when they already observed some behavior trends in password generation (e.g., choice of short passwords or selecting words from the dictionary) that continue to occur these days, such as the frequent use of short dictionary words as passwords. They found out that a simple dictionary attack could crack a third of the passwords.

Zviran and Haga (1999) surveyed 860 users at the DoD, the American Department of Defense. The study looked at the characteristics of user-generated passwords, such as password length, composition (numeric, alphabetic, etc.), lifetime (how often the password was changed), and selection method (based on meaningful details, random, other), as well as the relationship among these characteristics and data attributes, such as data importance and sensitivity. The frequency of changing a password was found to be related to the importance or sensitivity of the data being protected. They also found that a password's composition (i.e., types of characters used in a password), frequency of changing, and method of selection are related to how difficult it is to remember it, while the number of characters in a password is not, that is to say, a longer password is not necessarily harder to recall, but complex ones, composed of all kinds of characters are, in fact, difficult to remember. In addition, they found that length, change frequency, and selection method are not related to writing a password down, while its composition is. People were more likely to write down passwords if they contained different sorts of characters. In other words, writing a password down was related to the difficulty of recalling it. Usage frequency was related to recall difficulty and to writing down a password. The authors argued that ten years after the Morris and Thompson (1979) study, the users were still choosing passwords that were made up of personal details meaningful to the user, relatively short, comprised of alphanumeric characters, rarely changed and often written down. The study examined only one password per user, since their main

focus was password characteristics, not users. Also, as they said it themselves, this was a case study at one government organization at one time, although the authors did expect that their findings would not be atypical.

Other studies also found that users tend to choose short passwords that relate to information with personal significance. Adams and Sasse (1999) conducted an internet survey with 139 respondents from organizational employees along with 30 detailed interviews, in order to get data about user behavior and perceptions regarding password systems. Four factors were identified that influence the effective use of passwords: multiple passwords, password content, perceived compatibility with work practices, and user perception regarding corporate security and importance of information. Besides, the authors found evidences that many users simply do not follow security recommendations and end up choosing weak passwords. It was found that people tend to choose passwords with the least number of characters allowed. People also tend to base their passwords on information with personal significance, which can be easily found out, and many times, the same password is reused across several different systems. Similar to Zviran & Haga (1999), Adams and Sasse (1999) addressed only organization employees. No information was available regarding the ages or education levels of the sample in their study.

In another study, Riley (2006) surveyed 315 undergraduate students to assess their password practices for online accounts. She used a 101-item self-report questionnaire developed through pilot tests. Some were applied in paper form and some over the web. Her main findings were that users use simplistic practices to develop passwords, in other words, they tend to revert to the simplest possible strategies. Most students surveyed never changed their passwords unless it was required. Also, users did not vary password complexity depending on the nature of the site, whether it protects sensitive information, like a bank, or not, like for instance, instant messaging. She speculated that having multiple accounts makes it harder to recall many unique passwords. Above all, she found that users were more aware

of secure behaviors they should use than it was thought, but behave less securely in fact. Riley's sample was composed of only undergraduates, who had previous technical knowledge of password practices. There might be a bias since the average user might lack this knowledge and the effects of age, education, or sex could not be examined.

Instead of examining behaviors and perceptions about password systems or password characteristics, such as composition, some studies actually tried to focus on the user, that is, try to identify why the users act the way they do and what could be done. In order to develop a model to evaluate the impact that the human factor in password authentication had on Information Security, Carstens et al. (2004) carried out a survey (N=250) and an experimental study (N=30) with university students and employees. Their results indicated that passwords containing information that was meaningful to the individual were easier to remember, even if they contained additional characters, like for instance, symbols and punctuation characters. Other common behaviors observed include password reuse in different accounts, as well as writing down one or more passwords. Users with eight or more passwords were found to be at a higher risk of forgetting a password at least once a month, and they also tended to write more passwords down. Carstens et al. asked for participants' gender, age, and educational level in their survey, but didn't make inferences about that. Also, they regretted not doing a pilot testing of the instrument because they thought that, had they tested the questions in a preliminary research effort, they might have gotten better ones.

Brown and colleagues (2004) performed a survey with 218 undergraduate students to evaluate password generation and use. A questionnaire containing 19 items or password use categories, such as bank account or email was used so that, for each item, the participants were asked to describe the type of information used to generate or remember the respective password. The authors also classified the strategies used by participants for generating or remembering their passwords according to three factors: *entity, information* and *format* upon which the passwords were based. For most of the observed passwords, the most frequent

entity was the participant him/herself and relatives. The information most frequently used as a basis for passwords were names and dates, and the most common formats were full (intact, e.g., a date with day, month and year) and partial (e. g. a date with day and month only). This study was an important step in understanding password users and their memory problems. Nevertheless, by surveying a homogeneous population in age and education, it was not possible to make inferences about the impact of these factors on password usage in other age groups and with different educational levels. And that is exactly what the present study intended to do.

Survey Study

This study was designed to extend the work started by Brown et al. (2004) by expanding the sample to include different ages, specially elder people, both male and female, since an increasing elder population is a reality all over the world (United Nations, 2003), and different educational levels, since the majority of the population in the developing world is still composed of people with little or no education.

Since the focus of most studies reviewed in the literature mentioned above was mainly on the passwords and their characteristics, many questions about what people do in handling passwords and why, or what their passwords refer to, remain unanswered. By looking at password usage within different age groups, distinct educational levels, and comparing males and females, we intended to find out what characterizes the people that face the most memory issues when using passwords. In sum, our main goal was to investigate how age, sex, and educational level, as well as number of passwords owned, could affect people's memory for passwords.

METHOD

Participants

A sample of 263 unrewarded individuals participated in the study (Table 1). Three age groups were considered, according to the developmental stages described by Papalia and Olds (2000): young adults (age 39 or younger), middle-aged adults (age 40-64), and elder adults (65 or older). For each age group, participants were distributed according to their educational level and sex, that is, males and females who had not finished high school, and those who had at least a high school degree.

Table 1 - Distribution of participants by age, educational level, and sex.

	Education								TOTAL	
	Less Than High School				High School or More					
Sex / Age	Male		Female		Male		Female			
	N	%	N	%	N	%	N	%	N	%
≤ 39 years-old	13	4.9	16	6.1	24	9.1	38	14.4	91	34.5
Aged 40 to 64	32	12.2	18	6.8	11	4.2	39	14.8	100	38.0
≥65 years-old	12	4.6	23	8.8	21	8.0	16	6.1	72	27.5
TOTAL	57	21.7	57	21.7	56	21.3	93	35.3	263	100

Instrument

The instrument was based on the questionnaire developed by Brown et al. (2004). It was adapted through a preliminary survey conducted to determine most frequently used password categories. The adapted questionnaire was adjusted after a pilot testing with a sample of 20 university students and personnel. It is available upon request to the first author.

The first section of the anonymous questionnaire included demographic information, such as sex, age, and educational level. Section 2 contained nine categories of password use (e.g., bank, email, etc.) plus the 'other' category. For each item or password use category on the questionnaire, the participants stated how many passwords they owned, whether the passwords were chosen by themselves or assigned to them by the system, password length,

and the type of characters used in the password (letters, numbers, etc.). The respondents also described - as generically as possible - the strategy used for generating or remembering each password. Section 3 included open ended questions about whether participants ever forgot or mixed-up a password, how frequently they were required to replace their passwords (if so), and whether they used to write any of their passwords down.

Procedure

With the intent of avoiding biases due to differences in reading and comprehension abilities stemming from different educational levels, the majority of interviews was conducted individually and took approximately 10 to 15 minutes per participant. The participants were surveyed at their work/study/leisure locations in Porto Alegre, RS, Brazil. After a general introduction, if the person agreed to participate, he/she signed the informed consent form and answered the questionnaire orally. For each item, the interviewer asked whether the participant used passwords for the item If so, for each password owned, the interviewer asked more information about the password, such as, length, composition, etc. and took note of the answers. The college students, after a brief initial orientation session, filled out their questionnaires in class under assistance and close supervision of three interviewers. We were especially careful regarding the memorization strategies, thus offering plenty of detailed explanation and examples to ensure that the participants' privacy was protected.

RESULTS

We performed descriptive analyzes of the data from Section 2 of the questionnaire regarding characteristics of the passwords, such as number of passwords per category and per participant, password length, and password composition. These data were compared within and between the demographic groups defined by sex, age, educational level. In addition, we conducted a logistic regression analysis to test the relationships between the relevant variables

(sex, age, educational level, and number of passwords) and memory problems. All statistic analyses used a level of significance ≤ .05.

Password Characteristics

The surveyed participants had a mean of 5.38 (*SD*=3.79) password uses (*Mdn*=4; *Mo*=2; range=1 to 29). Since the distribution was positively skewed, the median was used for subsequent analyses. The mean number of unique passwords was 3.98 (*SD*=2.37; *Mdn*=4; *Mo*=2; range=1 to 16), which means that for about every four passwords used, one was duplicated. Curiously, the percentage of unique passwords was negatively correlated (r_s=-0.625, p<.001) with the total number of password uses, that is, as the number of password uses increased, the number of unique passwords tended to decrease, since people tended to start repeating already used passwords.

Password composition was also examined. From a total of 1415 passwords reported by the sample, the majority, 62.6%, was numeric only, followed by alphabetic only (24.3%), alphanumeric (12.4%), and barely 0.7% contained numbers, letters and other characters. The sum of the sizes of all passwords owned by an individual was divided by the total number of passwords he or she had, thus yielding the average password length for that particular person. For the full sample, the mean password length was 4.89 (*SD*=1.06, *Mo*=5, *Mdn*=4.80). These findings make more sense when looking at the most represented categories of passwords, in Table 2, since some systems impose limitations on password composition and sizes, like for instance, banks often require passwords containing a fixed number of characters, usually 4 or 6 digits.

Password Uses

As shown in Table 2, the vast majority of participants used passwords for their bank accounts (including phone and online banking), followed by other cards, email, and the internet (including instant message, online news, etc.). Other uses (not shown in Table 2)

including computer, cell phone, insurance, school/library, alarm, and the 'others' category, accounted for less than 5% each, totaling 17.7% of all passwords.

Table 2. Most frequent password uses per group (percentage of all 1415 passwords)

	Total	Education		Age			Sex	
		< High School	≥ High School	≤39	40-64	≥ 65	M	F
Password uses								
Bank	53.4	39.8	60.2	31.8	40.0	28.2	44.3	55.7
Other cards	11.6	38.3	61.7	34.0	35.2	30.9	40.7	59.3
Email	9.9	15.1	84.9	59.0	27.3	13.7	38.8	61.2
Internet	7.7	15.7	84.3	50.0	18.5	31.7	53.7	46.3

Memory Problems

Whenever participants reported having already forgotten passwords, mixed passwords up, or both, we considered that the person had already experienced memory problems in password use. Nearly two thirds of our sample, 72.1% (189) did.

When we specifically looked at the group that reported memory problems, we observed that, over half of the participants owned more than four passwords, the majority, 62.4% held at least a high school degree, 55.5% were under 65 years of age, and 59.3% were women. Those who did have memory problems and were better educated, used in average 7.2 passwords. On the other hand, from the 27.9% of individuals that never forgot nor mixed-up a password, most (59%) had not completed high school and owned in average 2.88 passwords. Confirming this tendency of better education corresponding to more passwords and consequently, more memory problems, a positive correlation (r_s=0,438, p<.001) was observed between education and number of password uses for users who reported memory problems. Over 90% of the participants with 10 or more passwords had at least a high school degree, and among those, roughly 95% had already forgotten or mixed-up their passwords.

In addition, over half of respondents who had memory problems (59.2%), admitted keeping a physical record of at least one of their passwords, 52.8% needed to reset their password at least once, and many passwords were reused for more than one account.

Memorization Strategies

The memorization strategies of all passwords were analyzed according to the classification scheme devised by Brown et al. (2004). Every password was analyzed in terms of the *entity* it refers to, the *information* or fact it was based upon, and the *format* in which it was coded. For example, if the strategy was "my girlfriend's birthday" and the password was composed of four digits, it was considered 'lover' for entity, 'date' for information, and 'partial' for format (the partial format was inferred given that dates are composed of day, month, and year, which require at least 6 digits).

Entity. Across the 1015 user-generated passwords, the most used entity was ***relative***, (41.8%), followed by *self* (37.3%), and *mixed* (6.3%). All the remaining entities (friend, acquaintance, lover, celebrity, fictional character, animal, organization, activity, and product) were grouped in the *others* category, since they accounted for less than 3% each.

Information. Regarding the information or fact used as a basis for generating the examined passwords, 39.5% of the passwords were based on ***dates***, followed by *numbers* (17.1%), *names* (12.5%), and *mixed* (10.9%). The *other* category includes *place*, *ID*, *aggressive words*, and *license plates*.

Format. The formats in which the information were coded to compose the passwords showed that a similar number of passwords were coded in ***partial*** (e.g., part of a date or a name) (29.2%) and ***combination*** (e.g., two dates combined) (29.8%) formats, followed by *full* (intact), (19.7%), *transformed* (e.g., phone number multiplied by 2), (9.1%) and *sequence* (e.g., first 5 odd numbers) (6.3%). The *other* category included *abbreviated*, *initials*, *backwards*, and *drawings* in the keyboard.

From all passwords examined for memorization strategies, 28.3% that were generated by the system and assigned to the users. A total of 71.7% (1015) were generated by the users and only those were considered for the analyses described below. These findings are supported in the literature (A. S. Brown et al., 2004). The mean number of passwords generated by the user is 3.76 (3.23) and by the system is 1.61 (1.63).

Considering that our sample used passwords mostly for bank-related uses, and that the majority of the passwords were numeric only, it was not surprising to find that a large number of passwords was said to be 'random' or not specified. Hence, the strategies that were not specified (23.0%) or random (10.9 %) were excluded for this analysis.

The Impact of Password Use on Memory

A Logistic Regression analysis tested the impact of sex, age, education, and number of passwords owned by an individual on reported memory problems, such as forgetting or confusion (mix-ups). Considering that the group aged between 40 and 64 stated having more memory problems than the other age groups, we decided to use this group as the reference for the age factor. An odds ratio equal to 1.0 represents full statistical independence, thus the odds that memory problems occur is larger as the odds ratios increase. Sex did not seem to affect memory problems (*OR*=.71 (.41-1.22), *p*=.264), and therefore the variable was removed from the model. The odds ratios were calculated for each factor with respect to memory problems and were adjusted to account for the interaction of all remaining factors (age, education, and number of passwords).

The results of this analysis are summarized in Table 3. The percentages were calculated in relation to the studied factor, either age, education, or number of passwords. For instance, in row 5 (Education < high school), the 71 individuals shown represent 62.3% of all respondents who did not finish high school, the remaining 37.7% did not report memory problems). When the factors were analyzed independently, age, educational level, and number of passwords were significant, see columns 4 and 5. However, when the contributions of age,

education, and number of passwords were taken into account, only the number of passwords owned by a participant remained significant for memory problems, that is to say, when age and education are equated, individuals who owned more than four passwords were 2.94 times more likely to forget or mix-up their passwords than their counterparts who owned up to four passwords.

Table 3. Odds ratios for memory problems, such as forgetting and confusion, calculated using logistic regression and adjusted for age, education and number of passwords.

		Memory Problems n (%)	***p***	**Odds Ratio (95% CI)**	***p***	**Adjusted odds ratio (95% CI)**
Age	39 or younger	71 (78.9)	.526	1.25 (.63-2.46)	.940	.97 (.48-1.99)
	Aged 40 to 64	75 (75)		1		1
	65 or older	43 (59.7)	.034	.49 (.26-.95)	.092	.55 (.28-1.1)
Education	< high school	71 (62.3)	.003	1	.108	1
	≥ high school	118 (79.7)		2.38 (1.37-4.14)		1.64 (.90-3.0)
# of passwords	≤ 4	80 (59.7)	<0.001	1	.001	
	> 4	109 (85.2)		3.87 (2.13-7.04)		2.94 (1.54-5.61)

Similar analyses were run for password forgetting and confusion, separately, and the results are comparable. Number of passwords was still the major factor, and for confusion problems, specifically, educational level remained significant even when the interaction of all factors was considered, which means that better-educated people with several passwords were more prone to mix them up, regardless of their age (Figure 1).

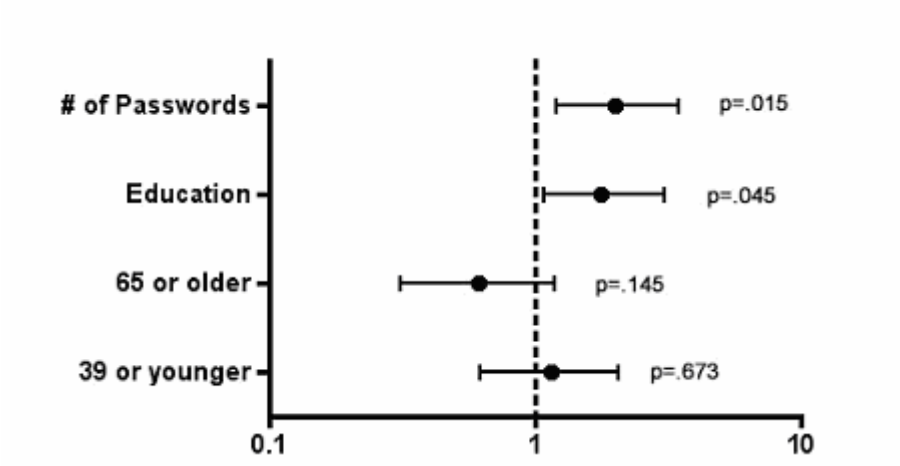

Figure 1. Odds Ratios of Confounding Passwords

Contrary to what might be expected, based on the literature, our results have indicated that age and educational level do not affect memory for passwords as much as the number of passwords owned by a person. Having several passwords, as a revenge effect, led people to experience memory problems, forgetting or mixing them up.

Discussion

Endorsing the literature, the present study indicated that security guidelines are routinely violated. That may be due to the fact that users have difficulties remembering their passwords. The vast majority of this survey's participants (72.1%) reported having already experienced memory problems while using passwords.

We found that, challenging the natural expectations of a worse memory performance for the elders, in the specific case of passwords, age did not play a significant role. We also found memory problems to be statistically independent from sex. Moreover, when the interaction of age, education and number of passwords owned was considered, the odds of a person having memory problems with passwords increased as the number of passwords increased. Interestingly, better educated users were at a higher risk of mixing passwords up. The odds of having problems of confusion, like for example, using the right password with the wrong account, or mixing-up the character order in a password, were also significant for education. We hypothesize that, as education improves, the number of password uses tend to

increase, so, people start repeating passwords and/or using variations of a "master" password. These slight variations would interfere with each other causing failed login attempts due to confusion.

These memory-related issues appear to be the origin of the revenge effects like reuse, password reset requests, or written copies of passwords. Unable to memorize many different passwords, users resorted to duplicate their existing passwords with new systems, practice that was already observed in other studies (A. Adams & Sasse, 1999; A. S. Brown et al., 2004; Carstens et al., 2004; Dhamija & Perrig, 2000). By reusing some passwords, they may also be at a higher risk of having the security of their accounts jeopardized. In line with previous findings (A. Adams & Sasse, 1999; A. S. Brown et al., 2004; Carstens et al., 2004; Dhamija & Perrig, 2000), over half of the participants, admitted keeping a physical record of at least one of their passwords. Carstens et al. (2004) found that users with eight or more passwords tended to write more passwords down. In our sample, we observed that, even among the people who reported having memory problems, more than half admitted writing passwords down.

This investigation is an expansion of previous studies with password users. Information about real password use was collected from male and female participants of different ages and educational levels. In addition to examining password usage, we were able to inspect a possible contribution of these factors to memory problems in password use.

Our data indicated that the factor that impacted password forgetting the most was the number of passwords owned by an individual. Adams and colleagues (1999, 1997) had already pointed out that having to remember multiple passwords decreased memorability and increased the cognitive overhead associated with the password mechanism. They recommended having no more than five different passwords. Besides, by owning more passwords, better-educated people appeared to be at a higher risk of experiencing memory problems when using passwords, such as forgetting and confusion. Users with eight or more

passwords were previously found to be at a higher risk of forgetting a password at least once a month (Carstens et al., 2004).

Most of the time the respondents in our study based their passwords on information of personal significance, especially dates related to family. Unlike Brown's study (A. S. Brown et al., 2004), where the *self* was the basis for over 90% of the passwords, our data showed a split between the *relative* and the *self* entities, with a slight advantage for *relative*, which might be due to cultural issues not considered in this study. Perhaps due to the nature of the systems using passwords that were mostly banks and often required numeric passwords, the majority of the passwords were based on dates, followed by numbers, names, and *mixed*. The format in which the passwords were coded was also often determined by the limitations of the system, like for instance, banks that limit their ATM passwords to 4 or 6 digits, and which was the most frequent category in this study, accounting for over half of the passwords (53.4%).

Even though security guidelines recommend using a mix of different characters in passwords, our findings indicated that only 0.7% (10) of the reported passwords presented a mix of numbers, letters and other characters. Although this study's participants were bound by the limitations of the systems for which they used passwords, this result suggests that, as found in the literature (Carstens et al., 2004; Riley, 2006; Zviran & Haga, 1999), the recommendations to use all kinds of characters in a password were not being followed. Interestingly, all 10 passwords (out of 1415) that did contain special characters were owned by better-educated respondents, with six of them, owned by people who were 65 years-old or older. Unlike other studies (Carstens et al., 2004; Riley, 2006; Zviran & Haga, 1999), the vast majority of passwords was numeric only, followed by alphabetic only, and alphanumeric - one more time, this may be due to the nature of the systems using passwords that were mostly banks.

The password dilemma of choosing secure and memorable passwords is a real problem that has been aggravated with the proliferation of e-commerce, online banking, and other online applications. The increasing concern with information security has lead some systems to strengthen password policies and many times block the creation of weak passwords. However, when forced to create stronger passwords that can't be forgotten, users are left with no alternative but write it down or reuse some older password they already know well.

The practices of choosing simple passwords, reusing existing ones, and writing passwords down are the revenge effects of password systems, since they put at risk the very purpose of using passwords, which is to protect information. By better understanding what factors contribute to memory problems with passwords, password system designers could try to account for these cognitive limitations and researchers could look for ways to minimize these effects by taking advantage of mechanisms already known to help memory for verbatim information such as the one required for passwords.

Human factor guidelines should be available to assist people in the development of strong passwords which are acceptable from an Information Security standpoint. This might be achieved by training users, raising their awareness about the threats associated with the use of weak passwords, as well as regarding good practices, such as mnemonics. Yan, Blackwell, Anderson and Grant (2004) demonstrated that passwords generated by means of a mnemonic technique, like for instance, the first letters of a song verse (e.g., for the verse "Welcome to the Hotel California!", create password "W2tHCa!"), can be as memorable as simple words and as secure as random strings.

However, Nielsen (2004) argued that user training is not enough to solve the password dilemma, because even if users are aware of the dangers and know what they should do, they will still choose passwords they can actually remember. One could expect that users create stronger passwords to protect more important or more sensitive information, but Zviran and

Haga (1999) found that how a password is chosen, the number of characters in a password, and password composition (type of characters used) were not affected by the level of data importance or sensitivity.

In sum, if the number of passwords owned is what makes it difficult to deal with passwords, it seems that the only reasonable solution is to have an acceptable number of them, a maximum of four to five, as recommended by Adams and Sasse (1999) and confirmed in this investigation. Since one cannot avoid having more passwords, Brown's suggestion of categorizing information sounds like a good way out. One could devise four or five categories of information, according to their importance (the inherent value of the data to the user) and sensitivity (the degree to which problems would arise if the information protected was known to others), and create four or five passwords that he/she can remember, with a strength level appropriate for each category, then, yes, reuse them.

Future research could test the efficacy of this approach, both for security and for memory performance. Moreover, two issues that were not examined in this study, the frequency of use of a password or how long a password has been used, are likely to have an impact on memory and their role in memory for passwords might be investigated as well. In addition, the vast body of knowledge from Cognitive Science, in general, and Memory research, specifically, could be used to help password learning and retrieval, and thus, avoid revenge effects.

References

Adams, A., & Sasse, M. A. (1999). Users are not the enemy. *Communications of the ACM, 42* (12), 41-46.

Adams, A., Sasse, M. A., & Lunt, P. (1997). Making passwords secure and usable. In H. Thimbleby, B. O'Conaill & P. Thomas (Eds.), *People & computers xii (proceedings of hci'97)* (pp. 1-19). Springer.

Brainerd, C. J., Reyna, V. F., & Brandse, E. (1995). Are childrens false memories more persistent then their true memories? *Psychological Science, 6*, 359-364.

Brown, A. S., Bracken, E., Zoccoli, S., & Douglas, K. (2004). Generating and remembering passwords. *Applied Cognitive Psychology, 18*(6), 641-651.

Brown, S. C., & Craik, F. (2000). Encoding and retrieval of information. In E. Tulving & F. Craik (Eds.), *The oxford handbook of memory.* (pp. 93-107). New York.: Oxford University Press US.

Carstens, D., McCauley-Bell, P., Malone, L., & DeMara, R. (2004). Evaluation of the human impact of password authentication practices on information security. *Informing Science Journal, 7*(1), 67-85.

Dempster, F. N., & Brainerd, C. J. (1995). *Interference and inhibition in cognition*. San Diego, CA: Academmic Press.

Dhamija, R., & Perrig, A. (2000). *Déjà vu: A user study using images for authentication.* Paper presented at the 9th USENIX Security Symposium, Denver, Colorado.

Morris, R., & Thompson, K. (1979). Password security: A case history. *Communications of the ACM, 22*, 594-597.

Nielsen, J. (2004). User education is not the answer to security problems. Retrieved January, 2005, from http://www.useit.com/alertbox/20041025.html

Papalia, D. E., & Olds, S. W. (2000). *Desenvolvimento humano*. Porto Alegre: Artes Médicas Sul.

Pergher, G. K., & Stein, L. M. (2003). Compreendendo o esquecimento: Teorias clássicas e seus fundamentos experimentais. *Revista Estudos de Psicologia, 14*, 129-155.

Reyna, V. F., & Brainerd, C. J. (1995). Fuzzy-trace theory - an interim synthesis. *Learning And Individual Differences*, *7*(1), 1-75.

Riley, S. (2006). Password security: What users know and what they actually do. Retrieved November, 2006, from http://psychology.wichita.edu/surl/usabilitynews/81/Passwords.htm

Sasse, M. A., Brostoff, S., & Weirich, D. (2002). Transforming the weakest link: A human-computer interaction approach to usable and effective security. In T. J. Regnault (Ed.), *Internet and wireless security* (pp. 243-258): London: IEE.

Tenner, E. (1997). *Why things byte back: Technology and the revenge of unintended consequences*. New York: Knopf.

United Nations, P. D. (2003). The ageing of the world's population. Retrieved January 15, 2003, from http://www.un.org/esa/socdev/ageing/popageing.html

Wiedenbeck, S., Waters, J., Birget, J.-C., Broditskiy, A., & Memon, N. (2004, July). *Passpoints: Design and evaluation of a graphical password system.* Paper presented at the Workshop on Usable Security Software, Rutgers.

Yan, J., Blackwell, A., Anderson, R., & Grant, A. (2004). Password memorability and security: Empirical results. *Security & Privacy*, 25-31.

Zviran, M., & Haga, W. J. (1999). Password security: An empirical study. *Journal of Management Information Systems, 15*(4), 161-185.

SEÇÃO EMPÍRICA II

EXPLORING COGNITIVE PSYCHOLOGY TO HELP PASSWORD SECURITY

ABSTRACT

Passwords have been used to secure the access to almost everything, from bank accounts to email. However, the requirements for secure passwords conflict directly with the cognitive abilities of the users, what generates countless problems. It is a fact that developing methods to generate secure and memorable passwords is difficult, yet the potential benefits certainly justify their pursuit, especially because it is likely that passwords remain widely used in the foreseeable future. As a step in this pursuit, two established concepts of Cognitive Psychology were explored with the intent to promote the generation of stronger and more memorable passwords: elaborative rehearsal and cue support to recall. Therefore, two experiments were conducted to investigate the effect of these two concepts on password generation and recall. Experiment 1 examined password composition, length, and security potential in experimental conditions (*repetition* and *cue* groups), as compared to a control group. Password recall was tested in two occasions, 5 minutes after password generation and after a week interval. The results from Experiment 1 indicate that people tend to observe only the requirements that are enforced, in spite of the instructions. Experiment 2 evaluated password recall after a longer delay: five weeks after generation. Group effects (experimental vs. control) on password recall were not observed. However, a possible confound was identified, the spacing effect, caused by the 5 minute login which, according to memory studies, facilitates the encoding process. In both experiments, recall levels were higher than expected. In addition, the errors from the unsuccessful login attempts were carefully analyzed and categorized. The observed error types suggest that, oftentimes, the users do remember the gist of the password, but forget the details of the format in which the password was coded.

Keywords: Passwords, human memory, elaborative rehearsal, cued recall, information security

1. Introduction

Although designing methods that will generate secure and memorable passwords is difficult, the potential benefits from developing such methods justify their pursuit, especially because the username-password combination is likely to remain widely used for years to come (Vu et al., 2007). Information security involves making information accessible to people who need it and are authorized to access it, while its integrity and confidentiality are maintained. Security, in general, and information security, in particular, is more of a social or human problem than a technological one. When there is intent to steal or access protected information, ill-intentioned individuals eventually find a way to take advantage of human nature and bypass security. Oftentimes, excessive technology gets in the way of security, by overcomplicating everyday tasks. When the security codes or procedures become too complex, people cannot remember them, so users write them down and post them on their computers, under keyboards or phones, or in their desk drawer. If the desired information is not in an obvious place, all one needs to do is ask, and someone will probably be willing to "help". In the end, information security is actually a systems' problem, but a system where the human is the most important component (D.A. Norman, 2004). Hence, it is necessary to design security systems and methods that take into account this human component, with all its capabilities and limitations.

User authentication is one of the key areas and one of fundamental challenges in information security. It is the process of determining whether one has access to a determined system or resource. Despite the well-known flaws and the several alternatives for authentication of legitimate users created in the recent years, such as biometric techniques, passwords are still the nearly universal authentication mechanism. Password authentication is popular because it is widely accepted by users, creation and authentication take only a few seconds, system implementation is simple and the cost is low. On the other hand, as long as users keep forgetting or mixing-up passwords, the very purpose of using passwords may be

put at risk. Nevertheless, the role that people play in information security is an important one that the literature has only begun to address.

In our technology-centered world, passwords are used to secure everything, from bank accounts to email. Developing strong passwords is necessary, but how to do this in a way that they can be remembered is still uncertain. It is clear that many of the deficiencies of password systems are related to human memory limitations (Sasse et al., 2002; Yan et al., 2004). In fact, Sasse et al. already compiled and enumerated the Cognitive Psychology findings that are most relevant for password use. They also concluded that existing human-computer interaction techniques could be used to prevent or address undesirable user behavior in password use. In an attempt to overcome password flaws, security researchers have tried to address the issue by devising alternative authentication methods.

Historically, several studies tried to address password-related cognitive limitations in different ways. Zviran and Haga (1990) devised a system named Cognitive Passwords, in which, the user and the system were involved in a dialog where the user answered a rotating set of five questions about highly personal facts and opinions, chosen from the predefined user's personal list of 20 question-answer pairs. A set of brief responses replaced a single password. After a three-month interval, cognitive passwords were much better remembered than traditional user-generated passwords. In addition, fact-based questions produced a higher rate of correct recall, as compared with opinion-based questions. The authors anticipated that issues of cost and time would raise concerns in the implementation of cognitive passwords, and that further research was needed. Unfortunately, those issues were perhaps considered too important, since traditional passwords are still the norm. Another recent study (Just, 2004) explored the design and evaluation of challenge-question systems as well, although its focus was on recovering lost credentials rather than primary authentication.

Cognitive passwords are a variation of associative passwords (S. L. Smith, 1987), a technique in which the user stores word pairs, so that, upon login, for every word presented to

him/her, she/he has to enter the associated pair word. Several approaches were tested by Pond, Podd, Bunnel and Henderson (2000): (a) when the users receive a list of cues and generate responses, (b) when the user generates both cue and responses, or (c) when the users choose a theme and then generate cues and responses within that theme. The average recall and guessing rates were comparable in the three groups.

Another alternative to traditional passwords, passphrases are sequences of words or other text used to control the access to a computer system, program or data (Porter, 1982). A passphrase is similar to a password in usage, but is generally longer for added security. Passphrases are often used to control both access to, and operation of, cryptographic programs and systems. Passphrases are particularly applicable to systems that use the passphrase as an encryption key. However, passphrases were also susceptible to memorability problems, since people often remembered the gist of the phrase but not its exact wording. Moreover, with longer strings, the likelihood of mistyping was also larger.

Instead of using long passphrases, one could take advantage of the mnemonic techniques and create phrase-based passwords containing the first letter of every word in a sentence, or replacing a word by a number or special character. Mnemonic passwords were found to be as strong as random passwords and just as memorable as naively selected passwords (Yan et al., 2004). Yet, Kuo et al. (2006) suggested that, since most users choose phrases from existing sources (e.g., a popular movie quote) rather than creating their own phrases, with increasing computer power, efficient mnemonic dictionaries can be created in the near future and then mnemonic passwords can be easily cracked.

In order to exploit the natural characteristics of human visual memory, and knowing that memory performance is better at recognition than at recall (Schwartz & Reisberg, 1991), the graphical password concept was created. The idea behind graphical authentication relies on the knowledge that visual memory is extremely powerful (Madigan, 1983). Two main system types are used for graphical authentication, recognition-based and location-based

(Renaud & De Angeli, 2004). Recognition-based systems require the user to select target images embedded amongst a set of distractors. Location-based systems, on the other hand, require the user to touch predetermined areas of an image in a fixed sequence for authentication. Furthermore, these systems rely on the ability to click on a specific position on the screen with some level of precision.

Examples of recognition-based systems are Passfaces and Déjà Vu. Passfaces (S. Brostoff & Sasse, 2000) is a system in which, at enrollment, the user selects four faces (photos), stored as his/her password. At login, they must correctly select their faces from a grid of nine faces displayed on the screen, one in each of four grids of nine faces. The grids are presented one at a time on the screen, and the order of presentation remains constant, as do the faces contained in each grid. However, no grid contains faces found in the other grids, and the order of faces within each grid is randomized. Before the enrollment is completed, users practice login a few times. Brostoff and Sasse (2000) tested Passfaces with college students for access control to a university resource. It proved to be very memorable over long intervals and less login failed attempts were observed, as compared with traditional passwords. Yet, with the computer power at the time of testing, the authentication process was slow and not heavily used. Passfaces is now a commercial application.

Besides aiming at improving memorability, Déjà Vu (Dhamija & Perrig, 2000), another system that uses images in user authentication, also intended to prevent users from choosing weak password, writing them down or sharing them. In Déjà Vu, the users first create a portfolio, by selecting a specific number of abstract images from a larger set presented by a server, then go through a training phase, to improve memorability of the portfolio images. At login, for each authentication challenge, the server creates a challenge set, which consists of portfolio and decoy images. If the user correctly identifies all portfolio images, she/he is authenticated. Déjà Vu performance for infrequently used passwords was good, but the authors recommended further investigations to evaluate whether this would

change with frequency of use, large or multiple portfolios. In addition, it was detected to be vulnerable to observer attacks.

Examples of location-based authentication systems are Jiminy and Passpoints. In Jiminy (Renaud & De Angeli, 2004), a user chooses a specific location in several images and has to identify it precisely at future logins. Passpoints (Wiedenbeck et al., 2004) allows the users to use images of their choice and choose any arbitrary sequence of points in the picture, for example, with 5 or 6 click points one can make more passwords than with 8-character passwords over the alphabet. In order to log in, the user has to click close to the chosen points (the tolerance can be adjusted). In the creation phase, the user chooses a number of points on the picture, not too close to one another; they have to remember the points and their click order. When the password is successfully created, the user begins the learning phase, where he or she practices her/his password and gets feedback, until the correct password is entered ten times. Then there is the retention phase, when the user actually logs in, in different intervals. The learning phase takes significantly longer for Passpoints users, than for alphanumeric password users. Besides there are several sources of error: number of points, click order, etc., and inputting the password also takes longer. Interestingly, the differences in input times did not seem to be due primarily to the mechanics of movement and selection, but rather to time needed to think about the chosen location and determine precisely where to click.

In terms of the essential criteria, graphical password systems are memorable, unpredictable and relatively convenient, but password generation and authentication take longer than traditional passwords and they are also less accessible, since visually impaired users or people with motor disabilities will have difficulty using them. Another issue arises when the user forgets the password, since it is not possible to send a “reminder” by email, and it is necessary to register a new password (Wiedenbeck et al., 2004). Location-based systems also require extra decision-making time. Finally, it is necessary to consider possible

interference effects, in case the user owns several graphical passwords, where new passwords may be confounded with old ones and vice-versa. Radvansky and Copeland (2006) observed consistent evidence of interference among similar images.

It turns out that many promising techniques were invented and tested; however, so far, none of them has beaten the convenience and cost of traditional passwords.

1.1 Password strength

It could be said that strong security depends on strong passwords. Strong passwords are usually defined by a set of policies slightly different across systems, nonetheless, most agree that a strong password should be resistant to password guessing attempts, be it human or automated guessing (Beverstock, 2003). A good password is one that is easily remembered by its owner, but not easily guessed by everybody else. Strong passwords are at least 8-character long, and contain different types of characters: numbers, symbols, and letters in both, upper and lower case. These passwords should not contain words from a dictionary in any language, nor dates, addresses, or ID numbers that could be easily connected to the user. Such passwords are probably not easily remembered by a person, and thus, are also probably not easily guessed.

The strength of a password can be assessed by means of cracking programs, which provide measures of the proportions of passwords cracked during a certain period of time using various methods (Vu et al., 2007). However, these programs usually have a cost and demand time. On the other hand, numerous applications provide password testers, in which passwords earn higher scores as they get longer and include more character classes (Kuo et al., 2006). Most of these testers provide feedback using qualitative progress bars, often using colors.

The human inclination towards choosing passwords composed of information that is meaningful to them, and that can thus be remembered, significantly reduces the amount of unique passwords that are actually used. Password composition and size are not the only

problem with strong passwords, though. An aggravating circumstance, as far as password memory is concerned, is that most password users these days have multiple passwords to remember, that is, they use passwords for different applications (e.g., bank, email, computers, etc.). Having multiple passwords reduces memorability and increases insecure practices, such as writing passwords down (A. Adams & Sasse, 1999). Moreover, the users not only have to remember passwords, but also the system and username with which they are associated. Unfortunately, that is not current practice for most users, since most of them still tend to decrease the memory burden at the expense of security (Pilar da Silva *et al.*, 2006). It would be wise, then, to have the strength of a password always be proportional to the value or importance of the data being protected, so that a user could focus on memorizing just a few strong passwords, for those most sensitive accounts.

The study of human memory and forgetting in psychological research has been going on for more than a century. Human memory research offers substantive knowledge that might be used to elucidate password forgetting problems. The Information Security industry could make use of this solid knowledge to find ways of better using human cognitive abilities in benefit of information security.

1.2 Memory and Repetition

Although it is not possible to completely dissociate the processes of acquisition, storage, and retrieval of memories, it is known that the way a material is encoded in memory may affect its retrieval (Schwartz & Reisberg, 1991). McDermott and Chan (2006) noticed that the use of rehearsal at encoding (study phase) leads to higher probabilities of accurate recall later on. Nonetheless, rote repetitions per se do not produce lasting effects in learning. Information that is analyzed deeply, is better recalled than information that is analyzed superficially, that is, thinking hard about the meaning of some information improves the likelihood that it will be recalled at a later time (elaborative rehearsal). In other words, rehearsing while thinking about the meaning of the material to be evoked, or establishing

connections with already known information, however, leads to better memory (Craik & Lockhart, 1972). Secure passwords should be random, meaningless strings, containing letters, numbers, and symbols, all of which makes elaborative processing hard. Thus, the elaborative rehearsal of a password, at its creation, may have an impact on its encoding in memory. After all, recalling meaningless items is harder than remembering meaningful ones (Ebbinghaus, 1885).

1.3 Cued Recall

Another poorly explored possibility in the literature is whether password forgetting and mix-up problems can be reduced with the systematic presentation of a *semantic* cue. Most existing password systems use primarily free recall. However, cognitive research indicates that cue support increases the probabilities of direct access to memory traces as compared to free recall (C.J. Brainerd *et al.*, 2002).

Free recall permits to recollect information in any order and enables the observation of the kind of organization used by individuals in retrieving this information. With cued recall, the retrieval process is activated by the traces of the stimulus that are provided to the individual. In this case, the cues act as a facilitator for the retrieval of the processed information (Schacter, 2001).

The efficacy of cue support in password recall has been observed in a pioneer study, funded by the American National Science Foundation (Lu & Twidale, 2003), and later replicated by Hertzum (2006), with similar results. For both studies, the goal was to develop a tool to improve the user experience of logging in, and security was secondary. The study assumed that a few carefully revealed hints would refresh a legitimate user's memory, but would not be sufficient to help an ill-intentioned potential intruder. At the time of registration, a user selected a number of characters within the password to be displayed at future logins. Before accepting this choice, the system performs a back-end dictionary check to verify that the password is not easily guessed. The system displays the password with the chosen

characters revealed, and the remaining characters replaced with stars. It then takes a snapshot of the masked password and distorts the image, to provide additional security against password cracking software. The transformed image is stored both on the local machine and on the host server. Participants did remember significantly more passwords with hints than without hints. They were also more confident in the correctness of their memory. However, despite the aid of the cue, many passwords were weak and/or fit into predictable patterns, both in password generation and in the choice of the to-be-revealed characters.

The exact match of the characters in a password, in its exact order and format is verbatim information. Human memory studies have consistently shown that people have difficulty remembering verbatim information after a delay, as well as that verbatim memory traces fade faster than semantic ones (Reyna & Brainerd, 1995).

1.4 Present Study

With the goal of exploring ideas based on Cognitive Psychology for improving memory performance in password use, this study investigated the effect of cue support and elaborative rehearsal on password generation and recall. Cue support to recall and elaborative rehearsal are two well-established cognitive artifacts, known to help memory. In Experiment 1, we conducted an initial evaluation of the effect of elaborative rehearsal (experimental group *repetition*) and of cue support (experimental group *cue*) on password composition, size, and security potential, as compared to a control group. Password recall was tested after a 5-min delay and a week later. Experiment 2 evaluated memory performance after a longer delay: 5 weeks from password generation. In addition, the errors committed in unsuccessful login attempts were examined and categorized.

2. Experiment 1

The purpose of this experiment was to examine what kind of passwords would be generated by the participants when basic security restrictions were imposed, as summarized

by Smith (2002) The participants were distributed into three groups: experimental group *repetition*, that rehearsed the password several times following creation, experimental group *cue*, that created a cue at password generation, and *control* group, that just created a password, with no further instructions. We also wanted to verify if their recall ability would be affected by the different generation condition and by the test time. Users were randomly assigned to either the *control* group or one of the experimental groups (*cue* or *repetition*).

Regarding the *repetition* condition, we expected password recall to be facilitated if, at generation time, the user rehearsed it several times while thinking about its meaning and/or establishing connections with information he/she already knew, as well as higher precise recall rates for the experimental groups at the second meeting, a week later, as compared with the *control* group. We also hypothesized that the creation of a semantic cue at password generation time would support the creation of stronger passwords, since the individuals could rely on the cue for help at future login occasions. In addition, we expected the systematic presentation of the cue at login time to facilitate password recall.

Given that a previous study results suggested that better-educated people tend to have more passwords and, consequently, be more prone to password forgetting and mix-ups (Pilar da Silva et al., 2006), the participants for all experiments were college students.

2.1 METHOD

The experimental design was a 2 x 3 mixed-model factorial; with *Test time* (immediate vs. one week after) manipulated within subject, and *groups* (control vs. repetition vs. cue) manipulated between subjects. The dependent variables were the characteristics of the generated password, password strength score, memory performance in terms of precise recall, type of error (if any), and number of attempts needed to remember the correct password and successfully log in.

Participants. Participated in Experiment 1, 128 first- or second-year volunteer students from PUCRS University majoring in Psychology, Economics and Business

Administration, the mean age was 21.2 (*SD*=5.08). They were randomly assigned to one of the three groups: 40 in the control group, 45 in the cue group, and 43 in the repetition group.

Materials. The materials consisted of web pages, similar to regular website user registration and login pages. The web pages were developed in PHP programming language, using a MySQL database. The interface was simple and uncluttered containing only the required input fields (Figure 1). The registration page also listed the password requirements. The demographic data was recorded in the database and the login trials were logged in flat text files.

Figure 1 - Login screen

Procedure. The participants were tested in groups at the computer lab, where each student took one computer and accessed the experiment registration page over a web connection. Participants were instructed to create a new fictitious password, not to be used in real settings and different from any real passwords they already had, for their own security. All participants were told the password should not be written down and be kept secret. In order to avoid an extra memory burden, the username to be used by participants followed a simple pattern, which consisted of their first name concatenated to the initials of their other names and their year of birth with four digits. The password was to follow basic security guidelines, which were presented with task instructions and remained visible on the screen

through the registration process. These guidelines were the following: “contain at least 6 characters”; “contain at least two letters (of either case)”; “contain at least two digits”; “special characters are allowed”; and “do not use the username as a password”.

Both experimental groups were also told to try and associate their password with information that was meaningful to them. In order to help them to better encode their password, the *repetition* group was instructed that, as they reconfirmed their password, they were to think about its meaning or the associations established with previously known information.

In the **study phase**, or generation phase, the user entered demographic data and username on the registration web page. The restrictions mentioned above were enforced via software, both in the client (at the user machine) and server (after submitting the registration data) sides. Feedback about the strength of the generated passwords was not provided to the participants. The user was then asked to enter a new password and re-enter it, to confirm. The *repetition* group repeated the confirmation step four times, each in a subsequent page. The *cue* group was instructed that, they would also create a cue, which was supposed to be a word or a phrase, close enough to the password as to help its recall, but vague enough, so it did not make sense to other people. They were also reminded that the cue would be visible on the screen and it might be seen by other people.

Following study, there was a **buffer activity**, which consisted of a psychological attention test (Cambraia, 2003), in order to purge the experimental password from participants' working memory. The test was chosen because it contains only images, and thus, the information manipulated would not directly interfere, or compete with the password in memory. In the **test phase**, or login occasions, participants were reminded of what their username consisted of, then proceeded to enter it followed by the password. They had a maximum of ten attempts to get the correct password. If the password was incorrect, they saw a sad face image and the message "The password was incorrect. Please try again". They were

not given feedback about the error. After the tenth strike, if they still could not remember the right password, they would see a happy face image with the message "Thanks for participating."

The individuals returned to the lab the following week to log in again. The procedure was identical to the immediate login. At the end of the experiment, the log files were compiled and analyzed.

2.2 RESULTS

In order to evaluate the impact of elaborative rehearsal and cue support on password generation and recall, we analyzed password characteristics, forgetting rates, and types of errors at unsuccessful login attempts. The significance levels were set at $\alpha<.05$.

2.2.1 Password Characteristics

The number of characters per password ranged from 6 to 15, with an overall mean length of 7.2 (*SD*=1.79) characters. Group mean lengths were 7.0 (*SD*=1.5) characters for the *repetition* group, 7.8 (*SD*=2.3) characters for the *cue* group, and 6.85 (*SD*=1.21) characters for the *control* group. A one-way ANOVA, with *group* (control, cue or repetition) as a between-subjects factor, was conducted on mean password length as the dependent measure. There was a main effect for group ($F(2,125)=3.604$, $p<.04$). Pairwise comparisons indicated that participants from the experimental condition *cue* created longer passwords than the *control* group ($p=.04$), but no difference was observed between the two experimental groups or between the *control* group and the *repetition* group ($p>.05$).

In order to quantitatively compare the strength of the generated passwords, we developed a scoring system to classify passwords based on their length and composition (Table 1). The system was based on most common security recommendations, widely popular in the literature (R. E. Smith, 2002), and some online password testers. We also developed a program to test the passwords and rank them based on this scoring system. In addition, the passwords that contained a dictionary word had their score decreased by 5 points.

Table 1. Password Strength Scores

Score	Rating
10 or less	very weak
10 to 19	Weak
20 to 29	Acceptable
30 to 39	Strong
40 or more	very strong

The strength of the passwords generated was evaluated according to the scale shown in Table 1. Most participants, 78.9% (n=101), used the weakest passwords allowed, including only letters and numbers. Slightly stronger passwords were used by 18.8% (n=24) of participants, containing a combination of upper and lower case letters and digits in their passwords. Strong passwords including special characters, letters, and numbers, were used by only 2.4% (n=3) of the participants. A one-way ANOVA, with group (control vs. cue vs. repetition) as a between-subjects factor, showed no significant effect of group on mean password strength scores.

2.2.2 Number of Login Attempts

Participants were allowed to try and remember their passwords up to ten times. A 3 (Group: control vs. cue vs. repetition) × 2 (Test time: immediate vs. 1 week) repeated measures ANOVAs with test time as a within-subject factor and group as a between-subjects factor indicated a significant main effect only for test time on overall number of attempts needed to login successfully, $F(1, 115)=4.87$, p <.03. Overall, the mean number of login attempts for the individuals who remembered their passwords (n=118) was 1.08 attempts (SD=.42, range=1 to 4) in the immediate login, and 1.34 attempts (SD=1.19, range=1 to 10) in the one week interval.

There were no significant effects of group or interactions (p>.05). Yet, it is important to note that in average the number of attempts needed to log in was low across groups, indicating that remembering the password was relatively easy for all participants.

2.2.3 Forgetting

We examined the password forgetting rates for the participants who were not able to login successfully either at the immediate login or at the one week interval. There was almost no password forgetting, since ceiling effects were detected in password recall, both in the control and experimental groups. All participants were able to recall the password in the immediate login. Most participants were able to log in with only one attempt in the immediate test (5 minutes after password creation): 97.7% of the *repetition* group, 86.7% of the *cue* group, and 100% of the control group. After one week, we still observed very low forgetting rates: only 1 (2.5%) participant in the *control* group, 4 (8.9%) participants in the *cue* group, and 5 (11.6%) participants in the *repetition* group were unable to recall their passwords at the second meeting.

2.2.4 Error types

In the past, the information technology community has concentrated efforts on reducing or eliminating the risk of malicious intruders. However, research has indicated that human error makes up as much as 65% of incidents causing economic loss for a company (Carstens et al., 2004). Thus, in order to enable the investigation of possible techniques to minimize errors in password recall, it is important to determine what types of errors users are prone to make when trying to remember a password to log in. The several attempts made by our users before they could successfully log in were examined by two independent judges and a categorization scheme was devised, as follows:

(a) error of amplitude, when characters were added or removed from the password;

(b) error of structure, when the participants recalled all components, but the order of the characters was mixed-up, or letters were typed in the wrong case;

(c) error of both structure and amplitude;

(d) semantic error, when a somehow related word was used instead of the original one;

(e) unrelated error or non-apparent relation to the experimental password.

The errors from the unsuccessful login attempts were classified according to the categories above, with an agreement level of 83% between the judges. It is important to note that each participant may have committed more than one error, since they had up to ten chances to try and remember the correct password at each experimental session.

Analysis of the error types among those who needed more than one attempt to log in, showed that the most common errors were of type a (amplitude). Table 2 lists the percentages of errors of each type, committed at given test times. For example, at the 5-minute test, 72.7% of all errors committed by participants from the *cue* group were of amplitude, while 18.2% were errors of structure, and 9.1% were errors of both amplitude and structure; no semantic or unrelated errors were committed by this group at this test time. Failure to remember the password promptly, especially for the *cue* group, suggested that the participants remembered the gist of the password, but not its verbatim details. This type of error is consistent with the memory literature in that verbatim traces fade faster (Reyna & Brainerd, 1995).

Table 2. Percentages of Error Types by group in two test times

	Amplitude		Structure		Amplitude & Structure		Semantic		Unrelated	
Time Group	5 min	1 week	5 min	1 week	5 min	1 week	5 min	1 week	5 min	1 week
Cue	72.7	50.0	18.2	15.4	9.1	15.4	0.0	19.2	0.0	0.0
Repetition	0.0	39.7	100.0	10.3	0.0	29.3	0.0	0.0	0.0	20.7
Control	0.0	23.5	0.0	5.9	0.0	0.0	0.0	11.8	0.0	58.8

2.3 Discussion

Experiment 1 results indicated that people do not follow security recommendations unless these are enforced somehow. The experimental passwords were required to satisfy basic security guidelines, yet most of them scored only as acceptable, suggesting that the ability to choose all kinds of characters or instructions to do so, do not have a strong effect on people's password generation practices. However, it is important to keep in mind that too much enforcement, may yield revenge effects such as the bad habit of writing passwords

down, since when users usually prefer to decrease the memory burden at the expense of security (Pilar da Silva et al., 2006). In fact, almost twenty years ago, De Alvare and Schultz (1988) already observed that security does not improve when password complexity increases, because users simply create a physical record of difficult passwords. Moreover, many of the passwords started with a capital letter and had the digits appended at the beginning or the end (perhaps to satisfy the restrictions), which is a predictable pattern and is likely to decrease password strength.

The recall rates were higher than we had expected, especially in the *control* group. The vast majority of users in all groups were able to login successfully with up to three attempts, which is the standard in most real applications, including banks. Many of the participants were, in fact, able to remember the gist of their password, but had a hard time remembering its exact composition and order, which is supported by memory theories, such as Fuzzy-Trace Theory (FTT) (Reyna & Brainerd, 1995). Fuzzy Trace Theory postulates that forgetting is the gradual disintegration of the features of a memory trace, when pieces of it (e.g., source information) become dissociated from one another. After a delay, since traces may become disintegrated, it is possible to forget some aspects of an experience while remembering others. In addition, FTT studies show that verbatim memories, such as passwords, are more vulnerable to interference effects and become inaccessible faster than gist memories, and that forgetting rates are higher for verbatim than for gist representations.

Interestingly, the errors observed in the immediate login were all from the experimental groups, with the *repetition* group having all observed errors from the same type (structure), that is, a change in case or character order, indicating that the password was remembered but some of its details were perhaps not well encoded or remembered. The errors from the *cue* group were mostly amplitude errors, that is to say, the users remembered the gist of the password (e.g., “grandmas’ initial + birthday”) but did not remember the exact format in which the password was coded (for instance, whether the birthday was coded as “21APR”

or “214”). After one week, most of the errors from the *cue* group were still of type amplitude, or amplitude and structure combined, which means that they remembered the information used in the password, but were unsure about its format.

3. Experiment 2

With Experiment 2 we intended to verify whether a longer interval would potentialize the effects of the experimental conditions on password recall, as well as on the type of error committed before successful login attempts. Since people are more likely to rely on verbatim representations immediately after experiencing an event, but shift to gist after a delay, we hypothesized that users would need more attempts to remember the experimental password after a longer delay, 5 weeks from password generation (or 4 weeks after their last login).

3.1 METHOD

The experimental design was also a 3 x 2 mixed-model factorial, with *groups* (control vs. repetition vs. cue) manipulated between subjects and *test time* (one week vs. five weeks) manipulated within subject. The dependent variables were memory performance in terms of precise recall, type of error (if any), and number of attempts needed to remember the correct password.

Participants. The participants in Experiment 2 were 87 volunteer Psychology and Business Administration students from PUCRS University: 29 in the control group, 28 in the cue group, and 30 in the repetition group. Most students rated themselves as having a good or advanced computer skill level and over half had been using the internet for seven years or more.

Materials. The materials were the same used in Experiment 1, followed by an online debriefing questionnaire.

Procedure. The study phase and buffer activity were identical to Experiment 1. In the test phase, all participants were reminded of their username format, then proceeded to enter it

followed by the password. The *cue* group was also reminded that, upon entering their user name, their cue would be displayed on the screen. As in Experiment 1, participants had a maximum of ten attempts to remember the correct password. After a successful login or the expiration of the allowed number of attempts to log in, the participants answered a debriefing questionnaire involving their former experience with technology, as well as about their choices in the creation of the experimental password.

3.2 RESULTS

3.2.1 Number of Login Attempts

As in Experiment 1, participants had up to ten chances to try and get the correct password. Overall, the individuals who remembered their passwords (n=79) needed in average 1.30 (SD=1.20) attempts to successfully log in, in the one week meeting, and an average of 1.53 (SD=1.43) in the 5-week interval. In the third meeting, individuals in the *repetition* group needed slightly more attempts to log in (M=1.85, SD=1.99), as compared to the second meeting (M=1.11, SD=.32). No differences were observed in the other groups.

The mean number of login attempts for the subjects who remembered their passwords, were submitted to 3 (Group: control vs. cue vs. repetition) × 2 (Test time: 1 week vs. 5 weeks) repeated measures ANOVAs with test time as a within-subject factor and group as a between-subjects factor. There was no main effect of either group or test time. However, a significant interaction between group and test time was observed, $F(1, 76)=5.301$, p <.03, that might explain why the *repetition* group needed more login attempts at the third meeting. No other interactions were observed.

3.2.2 Forgetting

The password forgetting rates for the 87 participants were examined for the logins after one week and after five weeks from password generation, respectively. In the week following password creation, 2 (7.1%) in the *cue* group forgot their passwords, as did 2

(6.7%) in the *repetition* group. No participants in the *control* group forgot their passwords after one week. After five weeks, only 2 (6.9%) participants in the *control* group, 3 (10.7%) participants in the *cue* group, and 3 (10%) participants in the *repetition* group failed to recall their passwords.

3.2.3 Error types

The errors made from the several login attempts made by users who, in the end, managed to remember their password, were examined and classified according to the categories described in Section 2.2.4., as compared to their stored password.

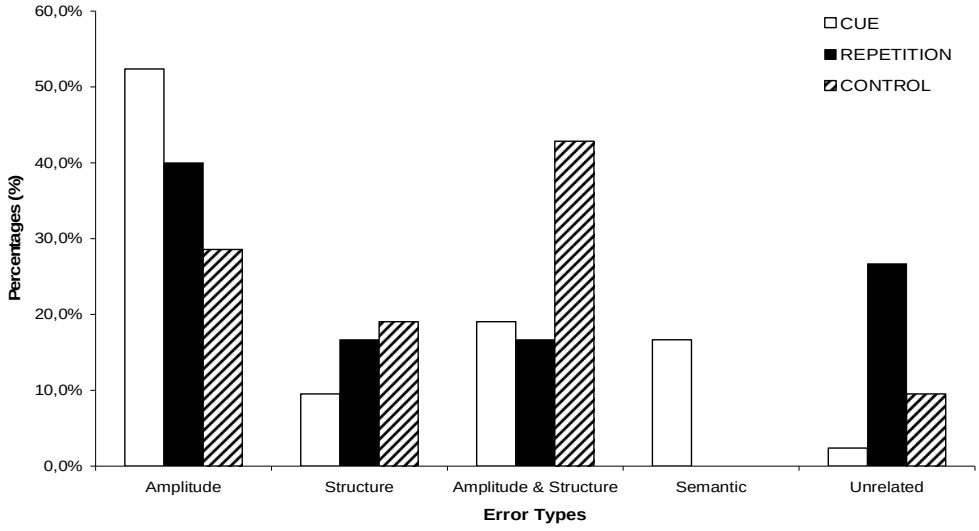

Figure 2. Error types per group five weeks after password generation.

Five weeks after the creation of the experimental password, the most common errors were still of type a (amplitude). Figure 2 depicts the percentages of error types per group in the third meeting. The comparison of the percentages for errors of types a, b, and c, between groups was not statistically significant (p>.05), however, significant differences were found between the *cue* group and each of the other two groups (*repetition* and *control*) for semantic errors (χ^2=9.19 p=.01), as well as between the two experimental groups for unrelated errors (χ^2=10.03 p<.01).

3.3 Discussion

Experiment 2 results suggest that the experimental conditions were not of major impact in password recall. However, researching the literature, we identified a possible confound, that may explain our excellent levels of recall across all groups, including the *control* group. Vu et al. (2007) observed that users who created a password and logged in after a 5-minute interval, presented better memory performance in password recall than those who just created their password and returned after one week, even if they re-entered their password once at creation, as is common practice in most applications. Now, all of our groups did log in after a 5-minute interval, in which a non verbal attention task was administered as a buffer activity. This finding is also supported in traditional memory literature, and it is called "*Spacing Effect*" (F. N. Dempster, 1988). The spacing effect refers to the finding that, for a given amount of study time, spaced presentations yield better learning than do massed presentations. The improvement in recall was found to be more observable with longer intervals between repetitions (lags). Thus, we speculate that the immediate login actually worked as a repetition after a 5-minute lag and that it benefited all participants alike, more than the experimental manipulations of elaborative rehearsal and cued recall.

Although only a small portion of participants committed errors when trying to log in, the analysis of the errors in the unsuccessful login attempts suggest that the effects of the experimental manipulations should be further investigated.

General Discussion and Implications of the Results

This study sought to explore Cognitive Psychology in order to find ways to help facilitate password generation and recall. The results indicate that, consistently with our former survey study (Pilar da Silva et al., 2006), users avoid mixing letters, digits, and symbols within a password, which is probably because that makes it harder to recall. Moreover, from the users in all three groups who volunteered to answer the debriefing questionnaire, most (close to 70%) said that reusing passwords is a current practice for them,

while an equivalent number said they do not write passwords down, giving the idea that these two bad habits are somehow mutually exclusive, that is, a user tends to resort to either reusing passwords or writing them down as a coping strategy. In addition, close to 80% of all experimental passwords followed only the minimal requirements (enforced via software) in composing their passwords, that is, included only numbers and letters and contained no more than six characters. Most of the remaining participants created slightly stronger passwords, including upper and lower case letters and digits, while only 3 participants (2.4%) used special characters as well. This result confirms previous findings in that people tend to choose the weakest passwords allowed by the system (Zviran & Haga, 1999).

Finding methods to help password users to generate passwords they can actually remember is a necessity. The quest for such methods is hard and requires the exploration of different paths. By learning from well-established science disciplines, we can eventually find some answers. Memory research began more than a century ago when Ebbinghaus (1885) published the results of his seminal work on memory with himself as the subject. Since then, Experimental Psychology has learned a lot about how we learn and remember or forget.

In Experiment 1, the assumption that users, if instructed or trained, may create stronger passwords was disconfirmed. In both experimental conditions (*cue* and *repetition*) we tried to engage participants to deeply process their password, while *control* group participants simply created a password according to the restrictions, without instructions on how to process it. Except for the participants from the *cue* group, who created longer passwords, which suggests that perhaps the students trusted the cue as an efficient aid, results showed that both experimental conditions had little effect on password generation, including their security potential. Most passwords contained one-case only letters and digits, as required, and only three participants used symbols as well. According to Nielsen (2004), user education should not be the main approach to countering security problems. It is indeed

necessary, but simply insufficient because security should be approached systemically, and the user is only one of its components.

As far as recall is concerned, we observed a much better performance than we had anticipated, in both Experiments 1 and 2. Memory performance for the *control* group was comparable to the experimental groups, if not better, but we considered our first test, logging in 5 minutes after password generation, might have played a reinforcing role, acting as a learning repetition after a lag. All three groups logged in after five minutes, and this may have acted as a confound.

It is also possible that the instructions to attribute meaning to the password or to associate it with personally meaningful information might have influenced the encoding of the experimental password by focusing attentional resources on its gist at the expense of superficial details. Given that gist memory is more persistent than verbatim memory, in the end, this may have helped password recall, despite the need for two or three login attempts in order to correctly remember the detailed format.

In line with Vu et al. (2007)'s findings, it seems that the number of login attempts allowed is not as significant as it was thought. Brostoff and Sasse (2003) suggested that reconsidering the "*3-strikes*" policy commonly applied to password login systems could be an immediate way of reducing the demand on system administrators and help desks. They predicted that requests for password reminders could be reduced by up to 44% by increasing the number of strikes from three to ten. Contrary to Brostoff and Sasse (2003)'s findings, most users in both experiments were able to log in, in up to three attempts, at all three intervals (5-minute, 1-week, and 5-weeks), and most individuals who tried more than three times, failed in the end.

Further research is needed to investigate the impact of password creation with and without the spacing effect, which, if confirmed, may be promptly used in most sensitive systems, such as banks, by having the user re-confirm their password after a 5-minute delay.

The benefits could be measured starting from a drop in password reset requests. The aid of the cue should also be further exploited since an effective cue may be able to, at least, help users to remember which password goes with which account. Although we did not find a significant effect of the cue on password recall, our users seemed to trust the help of the cue to create longer passwords. In addition, the failed login attempts by *cue*-group participants were often highly close to the actual password. If we consider password recall as a function of error type, we can observe indications that inducing deeper processing seems to help preserve the gist of the passwords, see the low percentages of unrelated errors amongst experimental groups, especially within the *cue* group. On the other hand, since the *control* group had a better performance, we might speculate that this deeper processing may have shifted encoding resources away from the verbatim details. It may be possible to work on the generation of better cues, though, that can effectively bridge that gap and help to rescue verbatim details for successful exact password matches.

Overall, most errors committed in unsuccessful login attempts were of amplitude, especially among the experimental groups. However, the sample of errors was small, since the forgetting rates were low. Nevertheless, research has already indicated that overload human information processing capabilities is the actual cause of human error, which can be referred to as system induced errors (Wickens, 1992). Hence, the errors should also be investigated further, with a larger sample, since knowing the types of errors users tend to make might enable researchers to work on possible ways to minimize them.

References

Adams, A., & Sasse, M. A. (1999). Users are not the enemy. *Communications of the ACM, 42* (12), 41-46.

Alvare, A. M. D., & Jr., E. E. S. (1988, August 29, 1988). *A framework for password selection.* Paper presented at the USENIX UNIX Security Workshop, Portland, OR, USA.

Beverstock. (2003). Passwords are dead! (long live passwords?). Retrieved June, 2003, from http://www.giac.org/certified_professionals/practicals/gsec/3017.php

Brainerd, C. J., Wright, R., Reyna, V. F., & Payne, D. G. (2002). Dual-retrieval processes in free and associative recall. *Journal of Memory and Language, 46*, 120-152.

Brostoff, S., & Sasse, M. A. (2000). *Are passfaces more usable than passwords: A field trial investigation.* Paper presented at the HCI 2000.

Brostoff, S., & Sasse, M. A. (2003). *Ten strikes and you're out: Increasing the number of attempts can improve password usability.* Paper presented at the CHI 2003 Workshop on Human-Computer Interaction and Security Systems. Ft. Lauderdale. Florida.

Cambraia, S. V. (2003). *Teste ac: Atenção concentrada* (3 ed.). São Paulo: Vetor Editora Psico-Pedagógica Ltda.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11*, 671-684.

Dempster, F. N. (1988). The spacing effect - a case-study in the failure to apply the results of psychological-research. *American Psychologist, 43*(8), 627-634.

Dhamija, R., & Perrig, A. (2000). *Déjà vu: A user study using images for authentication.* Paper presented at the 9th USENIX Security Symposium, Denver, Colorado.

Ebbinghaus, H. (1885). *Memory: A contribution to experimental psychology*.New York: New York: Teachers College, Columbia University.

Hertzum, M. (2006). Minimal feedback hints for remembering passwords. *Interactions, 13*(3), 38-40.

Just, M. (2004). Designing and evaluating challenge-question systems. *IEEE Security & Privacy*, 32-39.

Kuo, C., Romanosky, S., & Cranor, L. (2006). *Human selection of mnemonic phrase-based passwords.* Paper presented at the Symposium on Usable Privacy and Security, Pittsburgh, PA, USA.

Lu, B., & Twidale, M. B. (2003). Managing multiple passwords and multiple logins: Mifa minimal-feedback hints for remote authentication. *INTERACT'03*, 821-824.

Madigan, S. (1983). *Image memory*.Hillsdale, N.J.: Lawrence Erlbaum Associates.

McDermott, K. B., & Chan, J. C. K. (2006). Effects of repetition on memory for pragmatic inferences. *Memory & Cognition*, *34*(6), 1273-1284.

Nielsen, J. (2004). User education is not the answer to security problems. Retrieved January, 2005, from http://www.useit.com/alertbox/20041025.html

Norman, D. A. (2004). *Emotional design: Why we love (or hate) everyday things*.New York: Basic Books.

Pilar da Silva, D. R., Gomes, C. F. A., & Stein, L. M. (2006). The revenge effects of passwords. Unpublished.

Pond, R., Podd, J., Bunnel, J., & Henderson, R. (2000). Word association computer passwords: The effect of formulation techniques on recall and guessing rates. *Computers & Security, 19*(7), 645-656.

Porter, S. N. (1982). A password extension for improved human factors. *Computers and Security, 1*(1), 54-56.

Radvansky, G. A., & Copeland, D. E. (2006). Situation models and retrieval interference: Pictures and words. *Memory, 14*(5), 614-623.

Renaud, K., & De Angeli, A. (2004). My password is here! An investigation into visuo-spatial authentication mechanisms. *Interacting with Computers, 16*, 1017-1041.

Reyna, V. F., & Brainerd, C. J. (1995). Fuzzy-trace theory - an interim synthesis. *Learning And Individual Differences*, *7*(1), 1-75.

Sasse, M. A., Brostoff, S., & Weirich, D. (2002). Transforming the weakest link: A human-computer interaction approach to usable and effective security. In T. J. Regnault (Ed.), *Internet and wireless security* (pp. 243-258): London: IEE.

Schwartz, B., & Reisberg, D. (1991). *Learning and memory*:New York: W. W. Norton.

Smith, R. E. (2002). *The strong password dilemma authentication: From passwords to public keys*:Addison-Wesley.

Smith, S. L. (1987). Authenticating users by word association. *IEEE Computer Security, 6*(6), 464-470.

Vu, K.-P. L., Proctor, R. W., Bhargav-Spantzel, A., Tai, B.-L. B., Cook, J., & Schultz, E. E. (2007). Improving password security and memorability to protect personal and organizational information. *International Journal of Human-Computer Studies, 65*, 744-757.

Wickens, C. D. (1992): Human Factors Issues in Information Access. In: Proceedings of the Human Factors Society 36th Annual Meeting .

Wiedenbeck, S., Waters, J., Birget, J.-C., Broditskiy, A., & Memon, N. (2004, July). *Passpoints: Design and evaluation of a graphical password system.* Paper presented at the Workshop on Usable Security Software, Rutgers.

Yan, J., Blackwell, A., Anderson, R., & Grant, A. (2004). Password memorability and security: Empirical results. *Security & Privacy*, 25-31.

Zviran, M., & Haga, W. J. (1990). Cognitive passwords: The key to easy access control. *Computers & Security, 9*, 723-736.

Zviran, M., & Haga, W. J. (1999). Password security: An empirical study. *Journal of Management Information Systems, 15*(4), 161-185.

CONSIDERAÇÕES FINAIS

A literatura tem mostrado consistentemente que grande parte das deficiências dos sistemas de autenticação baseados em senhas se deve às condições de funcionamento da memória dos usuários (Ives et al., 2004; Renaud & De Angeli, 2004; Sasse et al., 2002; Yan et al., 2004) que são incompatíveis com as demandas da segurança da informação. A revisão da literatura também permitiu verificar que existem muitas diferenças entre as abordagens ao problema da segurança e memorabilidade das senhas. A maioria dos estudos existentes descreve sistemas de autenticação alternativos, muitas vezes apresentando um protótipo e sua avaliação. Várias alternativas promissoras foram criadas, mas todas envolvem custos e benefícios que restringem seu uso a sistemas mais específicos, como por exemplo, técnicas biométricas de reconhecimento de faces podem ser muito apropriadas para a segurança de aeroportos, mas o custo não se justificaria num sistema de matrículas universitário. As senhas, ao contrário, por seu baixo custo e alta conveniência, podem ser usadas quase que indiscriminadamente. Por essa razão, é bastante provável que as senhas ainda continuem sendo usadas por muito tempo.

Apesar disso, existe uma carência de estudos a respeito de qual ou quais os aspectos do uso de senhas acarretam uma sobrecarga da memória, ou ainda, que características dos usuários contribuem para que este lembre ou esqueça suas senhas, especialmente nos países em desenvolvimento. No Brasil, embora freqüentemente sejam noticiados incidentes envolvendo quebra de segurança por mau uso de senhas em instituições bancárias ou governamentais, ainda não existem estudos examinando o uso de senhas por usuários de diversos grupos etários e de escolaridade variada. Dessa forma, o mapeamento do uso de senhas numa realidade brasileira, descrevendo características das senhas e práticas comuns entre diversos grupos populacionais, assim como o achado de que o número de senhas que uma pessoa possui se constitui no fator que mais contribui para o esquecimento ou confusão de senhas, tornam-se contribuições importantes do presente trabalho de tese.

A diversidade da amostra do Estudo de Levantamento, descrito na Seção Empírica I, permitiu observar-se os efeitos de variáveis como número total de senhas de um indivíduo, sexo, idade e escolaridade sobre os problemas de memória reportados, tais como esquecimento e confusão de senhas, assim como possibilitou a análise das diversas interações dessas variáveis. Detectou-se ainda que, ao contrário das expectativas, o declínio cognitivo devido ao envelhecimento não parece acrescentar prejuízo à memória para senhas. No entanto, seria interessante complementar este estudo considerando-se a importância da informação a ser protegida, a freqüência de uso e o tempo de vida da senha, isto é, há quanto tempo a pessoa possui determinada senha, bem como sua familiaridade com a tecnologia, pois é possível que haja uma interação entre o número total de senhas e esses fatores.

Para que se possa melhorar a usabilidade de sistemas de autenticação por meio de senhas, é necessário que tanto projetistas quanto usuários estejam cientes da importância da segurança, das possíveis conseqüências do uso de senhas fracas e maus hábitos, e principalmente das capacidades e limitações dos usuários de senhas. Com esse objetivo, o papel do treinamento de usuários é muito importante, para que se aumente o nível de consciência e também para difundir boas práticas, como, por exemplo, o uso de técnicas mnemônicas, que permitem a criação de senhas tanto fortes quanto recordáveis, ou técnicas de *chunking* (Miller, 1956), onde o usuário agrupa a informação contida na senha em "pedaços" menores. Além disso, seria interessante considerar que é difícil para um ser humano lembrar-se de mais que quatro ou cinco senhas diferentes. Assim, uma vez que não se pode evitar ter várias senhas, talvez seja uma boa opção a idéia de utilizar "categorias de senhas" de acordo com a importância e criticalidade da informação a ser protegida.

Cabe destacar que consideramos de fundamental importância o estudo de métodos que possibilitem a utilização de técnicas cognitivas como auxílio à memorabilidade de senhas. Embora a concepção de métodos que permitam gerar senhas seguras e memoráveis seja difícil, os potenciais benefícios no desenvolvimento de tais métodos justificam sua busca.

Acreditamos que, se a recordação da senha for facilitada, ou se os usuários puderem contar com um auxílio eficaz para recordá-la, a criação de senhas mais seguras também se tornará mais viável. O Estudo de Levantamento nos permitiu conhecer os hábitos de uso de senhas em uma realidade brasileira e, a partir daí, buscar por métodos que facilitem a memorabilidade de senhas seguras. Infelizmente, embora o número de senhas que um usuário possui seja o fator que mais impacta sua memória, por questões logísticas, não nos foi possível testar o uso de diversas senhas por participante. Entretanto, sabendo-se que os usuários que possuem maior número de senhas são os que possuem maior escolaridade, na escolha da amostra para os estudos experimentais, buscamos participantes desse grupo. A senha experimental seria então adicionada às várias senhas que cada participante já possuía, acarretando um aumento na carga da memória. Assim, como um primeiro passo nessa busca, examinamos os efeitos de repetição e do auxílio de pista sobre a geração e recordação de uma senha experimental, cujos resultados estão descritos na Seção Empírica II.

Se observarmos os resultados obtidos na Seção Empírica II, podemos perceber que os dados não indicaram efeito das variáveis pista ou repetição sobre a geração e a recordação das senhas. Pelo contrário, os índices de recordação ligeiramente superiores do grupo controle sugerem que a indução de um processamento mais elaborado da senha pelas condições experimentais, no momento de sua geração, teria sido na verdade prejudicial à posterior recordação da mesma. Todavia, uma vez que os índices de esquecimento foram baixos em todos os grupos, incluindo o grupo controle, torna-se necessário um novo estudo que examine possíveis variáveis confundidoras, como por exemplo, a decorrência de intervalos diferentes entre *logins*.

Afinal, sabe-se que distribuir as repetições ao longo do tempo freqüentemente resulta em melhor memória que repetições sucessivas (Schacter, 2001). Dessa forma, uma possível explicação para os baixos níveis de esquecimento se refere ao teste imediato, depois de cinco minutos, ao qual foram submetidos, tanto os grupos experimentais quanto o grupo controle.

Esse intervalo pode ter funcionado como um efeito de espaçamento (F. N. Dempster, 1988) tendo beneficiado todos os participantes mais do que as manipulações experimentais de repetição elaborativa e auxílio de pista. Entretanto, refinamentos metodológicos podem ser integrados aos procedimentos. Um destes refinamentos refere-se à variável temporal na testagem da memória. A testagem envolvendo um grupo que não seja submetido ao teste imediato permitirá detectar se houve realmente o efeito de espaçamento, ou o reforço da aprendizagem da senha, após intervalo de cinco minutos. Outra questão metodológica importante diz respeito à variável instrução, que poderia ser manipulada de modo a reforçar a codificação dos aspectos literais das senhas.

Sabe-se que erros induzidos pelo sistema ou pelo *design* são aqueles em que as características do sistema excedem ou produzem inconsistências quanto às capacidades dos usuários (Wickens, 1992). Por essa razão, faz-se necessário explorar a conexão entre as capacidades humanas e as necessidades da segurança da informação ao projetar-se sistemas de autenticação. Pesquisas indicaram que o erro humano é responsável por cerca de 65% dos incidentes que causam prejuízo econômico às organizações (Carstens *et al.*, 2006). Contudo, para que se possa investigar possíveis maneiras de minimizar os erros na recordação de senhas, é importante determinar que tipos de erros os usuários tendem a cometer.

Assim, uma contribuição adicional deste trabalho se refere à análise criteriosa e à categorização dos tipos de erros cometidos pelos usuários em tentativas de *login* mal-sucedidas, possibilitando que eles sejam abordados de maneira diferenciada em trabalhos futuros, de acordo com a natureza de cada tipo. Foi detectado que, de um modo geral, a maior parte dos erros observados foi de amplitude (i.e. em que caracteres foram removidos ou adicionados à senha), especialmente entre os grupos experimentais. Além disso, não foram observados erros semânticos no grupo controle. No entanto, o estudo apresenta uma limitação importante, pois o tamanho da amostra de erros era pequeno, já que os índices de esquecimento foram baixos. Estudos futuros, envolvendo amostras maiores, e talvez criando

situações que favoreçam o erro, seriam necessários para melhor investigar estes fatores, uma vez que conhecendo-se os erros que as pessoas tendem a cometer, é possível buscar maneiras eficazes de minimizá-los.

Ao introduzir a presente tese, buscamos refletir sobre o dilema dos sistemas de autenticação por meio de senhas quanto à segurança e usabilidade. Apesar dos avanços tecnológicos, verificamos que o uso de senhas como método primário de autenticação tende a persistir por muitos anos. Por meio do estudo de levantamento, conseguimos observar o uso de senhas na nossa realidade junto a grupos populacionais bastante diversos. Nossos resultados mostraram que, embora a escolaridade não tenha efeito direto no esquecimento e confusão de senhas, pessoas com mais acesso à educação tendem a ter também maior acesso à tecnologia e, por isso, estão mais expostas a um maior número de senhas. Tal resultado tem uma implicação importante em tempos de inclusão digital, pois ao disponibilizar a tecnologia à população, se faz necessário considerar os perigos digitais aos quais ela estará exposta. Os resultados sugerem igualmente que, apesar de poder contar com um auxílio na codificação (repetição elaborativa) ou na recuperação (pista), a maioria das pessoas ainda tende a reduzir a carga da memória à custa da segurança. Para isso, utilizam-se de estratégias como escrever a senha em papel ou reusar a mesma senha para vários sistemas, práticas essas que tornam mais fácil a descoberta da senha por um intruso mal-intencionado. Isso acontece mesmo quando os usuários são instruídos adequadamente a respeito do que se constitui numa senha segura.

Os resultados dos estudos aqui descritos identificaram importantes vulnerabilidades produzidas através das ações de usuários de senhas, bem como algumas de suas causas. O uso desses resultados por pesquisadores e profissionais de tecnologia da informação poderá permitir a implementação de algumas das sugestões aqui apresentadas para que se possam minimizar essas vulnerabilidades. Através do aumento de funções automatizadas, treinamento e conscientização de usuários, e levando-se em conta suas capacidades e limitações será possível produzir sistemas mais seguros e usáveis. Sabendo-se que o número de senhas que

um usuário possui é o fator que mais compromete sua memória para senhas, uma opção seria a utilização de senha única, ao menos, num mesmo sistema, como, por exemplo, uma aplicação de *Internet Banking*.

Em suma, a maior preocupação da indústria de SI tem sido buscar maneiras de garantir a segurança através de novas tecnologias que minimizem o papel do usuário e o impacto de seus erros. Por outro lado, a maioria dos usuários não têm interesse nos mecanismos que serão usados para resolver os problemas da autenticação, desde que estes sejam resolvidos, o que infelizmente até agora não aconteceu. Cria-se então um paradoxo e, no caso das senhas, surgem os efeitos de vingança, tais como escrevê-las ou reusá-las. Acreditamos que não se trata de falta de informação ou de motivação. O comportamento dos usuários apenas reflete a tentativa de sobrevivência num mundo que, de uma hora para outra, se descobriu dependente de tecnologias que supostamente vieram para melhorar suas vidas. Verdade ou não, não é possível voltar no tempo e simplesmente descartar essas tecnologias. Assim, enquanto não compreendermos o que os usuários de senhas sabem, como pensam e agem, ou quais são seus objetivos, não seremos capazes de resolver tais problemas.

Neste momento, as pesquisas sobre usabilidade da segurança de informação ainda estão na fase de compreender os problemas. As soluções talvez ainda estejam distantes, mas é desta maneira que a ciência em geral avança, lentamente, persistentemente, até encontrar respostas viáveis.

REFERÊNCIAS BIBLIOGRÁFICAS

Adams, A., & Sasse, M. A. (1999). Users are not the enemy. *Communications of the ACM, 42* (12), 41-46.

Adams, A., Sasse, M. A., & Lunt, P. (1997). Making passwords secure and usable. In H. Thimbleby, B. O'Conaill & P. Thomas (Eds.), *People & computers xii (proceedings of hci'97)* (pp. 1-19). Springer.

Alvare, A. M. D., & Jr., E. E. S. (1988, August 29, 1988). *A framework for password selection.* Paper presented at the USENIX UNIX Security Workshop, Portland, OR, USA.

Anderson, J., & Paulson, R. (1977). Representation and retention of verbatim information. *Journal of Verbal Learning and Verbal Behavior, 16*, 439-451.

Aranha, A. C. n. d. (2005). A sociedade e a segurança da informação. Retrieved October 11, 2005, from http://www.microsoft.com/brasil/technet/Colunas/AnnaCarolinaAranha/Seguranca.mspx

Beverstock. (2003). Passwords are dead! (long live passwords?). Retrieved June, 2003, from http://www.giac.org/certified_professionals/practicals/gsec/3017.php

Brainerd, C. J., & Reyna, V. F. (1993). Memory independence and memory interference in cognitive development. *Psychological Review, 100* (1), 42-67.

Brainerd, C. J., Reyna, V. F., & Brandse, E. (1995). Are childrens false memories more persistent then their true memories? *Psychological Science, 6*, 359-364.

Brainerd, C. J., Wright, R., Reyna, V. F., & Payne, D. G. (2002). Dual-retrieval processes in free and associative recall. *Journal of Memory and Language, 46*, 120-152.

Brostoff, S. (2004). *Improving password system effectiveness.* University College London, London.

Brostoff, S., & Sasse, M. A. (2000). *Are passfaces more usable than passwords: A field trial investigation.* Paper presented at the HCI 2000.

Brostoff, S., & Sasse, M. A. (2003). *Ten strikes and you're out: Increasing the number of attempts can improve password usability.* Paper presented at the CHI 2003 Workshop on Human-Computer Interaction and Security Systems. Ft. Lauderdale. Florida.

Brown, A. S., Bracken, E., Zoccoli, S., & Douglas, K. (2004). Generating and remembering passwords. *Applied Cognitive Psychology, 18*(6), 641-651.

Brown, S. C., & Craik, F. (2000). Encoding and retrieval of information. In E. Tulving & F. Craik (Eds.), *The oxford handbook of memory*. (pp. 93-107). New York.: Oxford University Press US.

Cambraia, S. V. (2003). *Teste ac: Atenção concentrada* (3 ed.). São Paulo: Vetor Editora Psico-Pedagógica Ltda.

Carstens, D., Malone, L., & McCauley-Bell, P. (2006). Applying chunking theory in organizational password guidelines. *Journal of Information, Information Technology, and Organizations, 1*, 97-113.

Carstens, D., McCauley-Bell, P., Malone, L., & DeMara, R. (2004). Evaluation of the human impact of password authentication practices on information security. *Informing Science Journal, 7*(1), 67-85.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11*, 671-684.

Dempster, F. N. (1988). The spacing effect - a case-study in the failure to apply the results of psychological-research. *American Psychologist, 43*(8), 627-634.

Dempster, F. N., & Brainerd, C. J. (1995). *Interference and inhibition in cognition*. San Diego, CA: Academmic Press.

Dhamija, R., & Perrig, A. (2000). *Déjà vu: A user study using images for authentication.* Paper presented at the 9th USENIX Security Symposium, Denver, Colorado.

Ebbinghaus, H. (1885). *Memory: A contribution to experimental psychology*. New York: New York: Teachers College, Columbia University.

FolhaOnline. (2004). Brasileiro é condenado à prisão por invasão de sites de bancos. Retrieved January 5, 2005, from http://www1.folha.uol.com.br/folha/informatica/ult124u14866.shtml

Garancis, P. (2004). My gate is locked, is yours? A look at implementing a strong password policy. Retrieved February, 2004, from http://www.giac.org/certified_professionals/practicals/GSEC/3861.php

GloboOnline. (2005). Estagiário desvia r$ 3 milhões do inss. Retrieved November 29, 2005, from http://oglobo.globo.com/online/sp/189451733.asp

Hertzum, M. (2006). Minimal feedback hints for remembering passwords. *Interactions, 13*(3), 38-40.

Ives, B., Walsh, K. R., & Schneider, H. (2004). The domino effect of password reuse. *Communications of the ACM, 47*(4), 76-78.

Jeyaraman, S., & Topkara, U. (2005). *Have the cake and eat it too - infusing usability into text-password based authentication systems*: Purdue University.

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin, 114*, 3-28.

Just, M. (2004). Designing and evaluating challenge-question systems. *IEEE Security & Privacy*, 32-39.

Kuo, C., Romanosky, S., & Cranor, L. (2006). *Human selection of mnemonic phrase-based passwords.* Paper presented at the Symposium on Usable Privacy and Security, Pittsburgh, PA, USA.

Lockhart, R. S. (2000). Methods of memory research. In E. Tulving & F. Craik (Eds.), *The oxford handbook of memory* (pp. 45-57). New York: Oxford University Press US.

Lu, B., & Twidale, M. B. (2003). Managing multiple passwords and multiple logins: Mifa minimal-feedback hints for remote authentication. *INTERACT'03*, 821-824.

Madigan, S. (1983). *Image memory*. Hillsdale, N.J.: Lawrence Erlbaum Associates.

McDermott, K. B., & Chan, J. C. K. (2006). Effects of repetition on memory for pragmatic inferences. *Memory & Cognition, 34*(6), 1273-1284.

Miller, G. A. (1956). The magical number seven, plus or minus two. Some limits on our capacity for processing information. The *Psychological Review*, 63, 81-97.

Morris, R., & Thompson, K. (1979). Password security: A case history. *Communications of the ACM, 22*, 594-597.

Nielsen, J. (2004). User education is not the answer to security problems. Retrieved January, 2005, from http://www.useit.com/alertbox/20041025.html

Norman, D. A. (1990). *The design of everyday things*. New York: Doubeday.

Norman, D. A. (2004). *Emotional design: Why we love (or hate) everyday things*. New York: Basic Books.

Papalia, D. E., & Olds, S. W. (2000). *Desenvolvimento humano*. Porto Alegre: Artes Médicas Sul.

Pergher, G. K., & Stein, L. M. (2003). Compreendendo o esquecimento: Teorias clássicas e seus fundamentos experimentais. *Revista Estudos de Psicologia, 14*, 129-155.

Pilar da Silva, D. R., Gomes, C. F. A., & Stein, L. M. (2006). The revenge effects of passwords. Unpublished.

Pond, R., Podd, J., Bunnel, J., & Henderson, R. (2000). Word association computer passwords: The effect of formulation techniques on recall and guessing rates. *Computers & Security, 19*(7), 645-656.

Porter, S. N. (1982). A password extension for improved human factors. *Computers and Security, 1*(1), 54-56.

Radvansky, G. A., & Copeland, D. E. (2006). Situation models and retrieval interference: Pictures and words. *Memory, 14*(5), 614-623.

Renaud, K., & De Angeli, A. (2004). My password is here! An investigation into visuo-spatial authentication mechanisms. *Interacting with Computers, 16*, 1017-1041.

Reyna, V. F., & Brainerd, C. J. (1995). Fuzzy-trace theory - an interim synthesis. *Learning And Individual Differences, 7*(1), 1-75.

Riley, S. (2006). Password security: What users know and what they actually do. Retrieved November, 2006, from http://psychology.wichita.edu/surl/usabilitynews/81/Passwords.htm

Sasse, M. A., Brostoff, S., & Weirich, D. (2002). Transforming the weakest link: A human-computer interaction approach to usable and effective security. In T. J. Regnault (Ed.), *Internet and wireless security* (pp. 243-258): London: IEE.

Schacter, D. (2001). The seven sins of memory. *Psychology Today, 34*(3), 62-66,87.

Schwartz, B., & Reisberg, D. (1991). *Learning and memory*: New York: W. W. Norton.

Sieberg, D. (2005). Hackers shift focus to financial gain. Retrieved 26 de setembro, 2005, from http://www.cnn.com/2005/TECH/internet/09/26/identity.hacker/index.html

Smith, R. E. (2002). *The strong password dilemma authentication: From passwords to public keys*: Addison-Wesley.

Smith, S. L. (1987). Authenticating users by word association. *IEEE Computer Security, 6*(6), 464-470.

Spector, Y., & Ginzberg, J. (1994). Pass-sentence: A new approach to computer code. *Computers & Security, 13*, 144-160.

Tenner, E. (1997). *Why things byte back: Technology and the revenge of unintended consequences*. New York: Knopf.

Titcomb, A. L., & Reyna, V. F. (1995). Memory interference and misinformation effects. In F. N. Dempster & C. J. Brainerd (Eds.), *Interference and inhibition in cognition* (pp. 263-294). New York: Academic Press.

United Nations, P. D. (2003). The ageing of the world's population. Retrieved January 15, 2003, from http://www.un.org/esa/socdev/ageing/popageing.html

Van Vleck, T. (1997). Java password generator. Retrieved July, 2006, from http://www.multicians.org/thvv/qpw.html

Vu, K.-P. L., Proctor, R. W., Bhargav-Spantzel, A., Tai, B.-L. B., Cook, J., & Schultz, E. E. (2007). Improving password security and memorability to protect personal and organizational information. *International Journal of Human-Computer Studies, 65*, 744-757.

Wickens, C. D. (1992): Human Factors Issues in Information Access. In: Proceedings of the Human Factors Society 36th Annual Meeting

Wiedenbeck, S., Waters, J., Birget, J.-C., Broditskiy, A., & Memon, N. (2004, July). *Passpoints: Design and evaluation of a graphical password system.* Paper presented at the Workshop on Usable Security Software, Rutgers.

Wikipedia. (2005). Wikipédia, a enciclopédia livre. Retrieved October 10, 2005, from http://pt.wikipedia.org/wiki/Segurança_da_informação

Winograd, E., & Soloway, R. M. (1986). On forgetting the locations of things stored on special places. *Journal of Experimental Psychology, 115*, 366-372.

Yan, J., Blackwell, A., Anderson, R., & Grant, A. (2004). Password memorability and security: Empirical results. *Security & Privacy*, 25-31.

Zviran, M., & Haga, W. J. (1990). Cognitive passwords: The key to easy access control. *Computers & Security, 9*, 723-736.

Zviran, M., & Haga, W. J. (1999). Password security: An empirical study. *Journal of Management Information Systems, 15*(4), 161-185.

ANEXOS

ANEXO A

Este questionário se refere ao uso de códigos de identificação privados, mais conhecidos como *senhas*. Para proteger sua privacidade, não há identificação pessoal nesta folha. A informação que você fornecer será usada unicamente para fins de pesquisa.

SEXO: **M** / **F** ESCOLARIDADE: _____ Escolaridade da MÃE: _____
IDADE: _____ OCUPAÇÃO: _____ CURSO/SEMESTRE: _____

Você usa senhas? Marque com "x" as apro-priadas:	Quantas senhas para esse item?	Quem escolheu? (usuário/ sistema)	Tamanho (nº caracteres)	Que caracteres contém? (nºs,letras,outros)	TIPO DE INFORMAÇÃO (Exemplos: ano de nascimento de familiar; nome da namorada + placa do carro dela; telefone antigo próprio)
1. ( ) Banco Cartão Bankfone Internet Cheques					
2.( ) Outros Cartões					
3. ( ) Email					
4. ( ) Computa-dor Trabalho Casa					
5. ( ) Telefone/ Celular					
6. ( ) Internet (exs.: ZH, Orkut, Skype, lojas virtuais)					

Você usa senhas? Marque com "x" as apro-priadas:	Quantas senhas para esse item?	Quem escolheu? (usuário/ sistema)	Tamanho (nº caracteres)	Que caracteres contém? (nºs,letras,outros)	TIPO DE INFORMAÇÃO (Exemplos: ano de nascimento de familiar; nome da namorada + placa do carro dela; telefone antigo)
7. ( ) Escola/ Curso/ Faculdade					
8. ( ) Seguro/ Plano de Saúde					
Outros:					
( )					
( )					
( )					
( )					
( )					

Agora, por favor responda as questões a seguir:

a) Alguma vez você esqueceu alguma das senhas acima? ( ) SIM ( ) NÃO
Se "SIM", por favor liste as senhas (pelo número no quadro acima) e o que você fez/como agiu quando isso aconteceu?

b) Alguma vez você se confundiu ou misturou alguma(s) das senhas acima?
( ) SIM ( ) NÃO
Se "SIM", liste quais (pelo número no quadro acima) e o que aconteceu?

c) Você é forçado a mudar alguma dessas senhas regularmente (por exemplo, a cada 6 meses)?
( ) SIM ( ) NÃO
Se "SIM", liste quais (pelo n° no quadro acima) e com que freqüência você precisa mudá-las.

d) Você costuma guardar uma cópia escrita (em papel ou meio eletrônico) de uma ou mais senhas? ( ) SIM ( ) NÃO

e) Quantas vezes você precisou solicitar uma nova senha?

f) Já teve uma de suas contas "invadida"?
( ) SIM ( ) NÃO

ANEXO B

PONTIFÍCIA UNIVERSIDADE CATÓLICA DO RIO GRANDE DO SUL
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
COMITÊ DE ÉTICA EM PESQUISA - CEP - PUCRS

Ofício 446/06-CEP

Porto Alegre, 28 de abril de 2006.

Senhor(a) Pesquisador(a):

O Comitê de Ética em Pesquisa da PUCRS apreciou e aprovou seu protocolo de pesquisa registro 06/02937, intitulado: "O dilema das senhas".

Sua investigação está autorizada a partir da presente data.

Relatório parcial e final da pesquisa deve ser encaminhado a este CEP.

Atenciosamente,

Prof. Dr. José Roberto Goldim
COORDENADOR DO CEP-PUCRS

Ilmo(a) Sr(a)
Dout Denise Ranghetti Pilar da Silva
N/Universidade

ANEXO C

Pontifícia Universidade Católica do Rio Grande do Sul
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
COMITÊ DE ÉTICA EM PESQUISA

Ofício 0209/07-CEP

Porto Alegre, 07 de março de 2007.

Senhor(a) Pesquisador(a):

O Comitê de Ética em Pesquisa da PUCRS apreciou e aprovou seu protocolo de pesquisa registro CEP 07/03587, intitulado: "**Identidade secreta: o jogo da memória"**

Sua investigação está autorizada a partir da presente data.

Relatórios parciais e final da pesquisa devem ser entregues a este CEP.

Atenciosamente,

Prof. Dr. José Roberto Goldim
COORDENADOR DO CEP-PUCRS

Ilmo(a) Sr(a)
Profa Lilian Milnitsky Stein
N/Universidade

PUCRS
Campus Central
Av. Ipiranga, 6690 – 3ºandar – CEP: 90610-000
Fone/Fax: (51) 3320-3345
E-mail: cep@pucrs.br
www.pucrs.br/prppg/cep

ANEXO D

TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

Prezado(a) participante:

Sou estudante de doutorado na Faculdade de Psicologia da Pontifícia Universidade Católica do Rio Grande do Sul. Estou realizando uma pesquisa sob supervisão da professora Lilian Milnitsky Stein, cujo objetivo é identificar os principais usos de senhas.

Sua participação envolve a resposta a um questionário anônimo, envolvendo o usos de senhas, bem como características das mesmas, estratégias usadas para lembrar delas e possíveis problemas devido ao esquecimento de senhas.

A participação nesse estudo é voluntária e se você decidir não participar ou quiser desistir de continuar em qualquer momento, tem absoluta liberdade de fazê-lo.

Na publicação dos resultados desta pesquisa, sua identidade será mantida no mais rigoroso sigilo. Serão omitidas todas as informações que permitam identificá-lo(a).

Mesmo não tendo benefícios diretos em participar, indiretamente você estará contribuindo para a compreensão do fenômeno estudado e para a produção de conhecimento científico.

Quaisquer dúvidas relativas à pesquisa poderão ser esclarecidas pela pesquisadora, fone (51)9988-1073, no Grupo de Pesquisas, 3320-3633, ramal 225, ou pela entidade responsável – Comitê de Ética em Pesquisa da PUCRS, fone 3320 3345.

Atenciosamente

_____

Denise R. P. Silva
Matrícula: 05290236-8

Porto Alegre, 31 de outubro de 2005

_____

Profa. Dra. Lilian M. Stein
Matrícula: 032022

Consinto em participar deste estudo e declaro ter recebido uma cópia deste termo de consentimento.

_____

Nome e assinatura do participante

Porto Alegre, ____ de _____ de 200_

ANEXO E

TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

Prezado(a) participante:

Sou estudante de doutorado na Faculdade de Psicologia da PUCRS. Estou realizando uma pesquisa sob supervisão da professora Lilian Milnitsky Stein, cujo objetivo é estudar maneiras de melhorar a memorabilidade de senhas.

Sua participação envolve a participação em um experimento no qual você criará uma senha fictícia, a qual não deve ser idêntica à nenhuma senha real que você já possua, para sua própria segurança. Ao assinar este termo, você se compromete a não utilizar esta senha em nenhum contexto real, uma vez que ela será visível aos experimentadores. A senha criada deverá seguir as recomendações de segurança, as quais lhe serão apresentadas previamente à realização do experimento.

O experimento será realizado em duas ou três etapas e você precisará da senha para efetuar o *login* nos momentos posteriores. Para o bom andamento do experimento, a senha não poderá ser anotada em papel e deverá ser mantida em segredo.

A participação nesse estudo é voluntária e se você decidir não participar ou quiser desistir de continuar em qualquer momento, tem absoluta liberdade de fazê-lo.

Na publicação dos resultados deste estudo, sua identidade será mantida no mais rigoroso sigilo. Serão omitidas todas as informações que permitam identificá-lo(a).

Após a conclusão do experimento, você terá a oportunidade de assistir a uma aula-treinamento sobre segurança da informação, na qual você aprenderá maneiras de melhor administrar suas senhas, como criá-las para que sejam seguras e de forma que você não as esqueça. Além disso, você estará contribuindo para a compreensão do fenômeno estudado e para a produção de conhecimento científico.

Quaisquer dúvidas relativas à pesquisa poderão ser esclarecidas pela pesquisadora, fone (51)9988-1073, no Grupo de Pesquisas, 3320-3500, ramal 7741, ou pela entidade responsável – Comitê de Ética em Pesquisa da PUCRS, fone 3320-3345.

Atenciosamente

_____ Porto Alegre, __ de _____ de 200_

Denise R. P. Silva

Matrícula: 05290236-8

_____

Profa. Dra. Lilian M. Stein (Matrícula: 032022)

Consinto em participar deste estudo e declaro ter recebido uma cópia deste termo de consentimento.

_____ Porto Alegre, ____ de _____ de 200__

Nome e assinatura do participante

ANEXO F

RAPPORT / INSTRUÇÕES

Apresentação dos experimentadores e Grupo de Pesquisas em Processos Cognitivos.

Agradecimento.

Esse estudo visa investigar os aspectos do funcionamento da memória relacionados ao uso de senhas. Para que o experimento seja concluído com sucesso, nós contamos com a sua participação e colaboração.

Após a conclusão do experimento, vocês terão a oportunidade de assistir a uma aula-treinamento sobre segurança da informação, na qual serão discutidas maneiras de melhor administrar suas senhas, como criá-las para que sejam seguras e de uma forma que você não as esqueça.

1. O experimento será realizado em três momentos. As tarefas a serem realizadas são muito simples, mas as instruções devem ser observadas à risca para o bom andamento do experimento. Hoje vocês criarão um nome de usuário e uma senha, a qual será necessária para efetuar o *login* nos momentos posteriores. Isto é, vocês retornarão ao laboratório na próxima semana e daqui a 4 semanas.
2. LOGÍSTICA:
    a. Desligar celulares
    b. Parar e aguardar instruções sempre que acabar uma tarefa. Não clicar OK!
    c. Usar sempre o mouse e não as teclas TAB e ENTER
    d. Se receber mensagens de erro do Windows ou do Internet Explorer, ignorar, cancelar ou clicar "não".
    e. Fazer silêncio
3. DURAÇÃO: a etapa de hoje deve durar em torno de meia hora, a próxima etapa, uns 10 minutos, e a 3ª. Etapa, cerca de 20 minutos.
4. Após a criação da senha, quando todos terminarem, será realizada uma outra tarefa sem o computador e depois disso, vocês utilizarão o computador novamente para efetuar o login. IMPORTANTE: Esse nome de usuário e senha não dão acesso a nada, servem apenas para fins experimentais, isto é, para observar seus efeitos na memória.
5. Por favor, não se preocupem em memorizar o nome de usuário, ele será criado seguindo uma regra muito simples e eu vou lembrá-lo sempre que vocês forem usá-lo. Isto é, o nome de usuário será o seu 1º. nome, seguido das iniciais de seus outros nomes e sobrenomes, seguido do seu ano de nascimento (4 dígitos), por exemplo, no meu caso, deniserps1966 ☺.
6. Para a sua própria segurança, a senha que será criada para este estudo deve ser fictícia, pois os experimentadores terão acesso a elas. Pedimos que não utilize essa senha em nenhum outro local real.

7. É importante que essa senha não seja anotada em papel ou meio eletrônico. Como o objetivo desse estudo é analisar o impacto do uso de senhas na memória, escrever a senha prejudicaria o resultado obtido.
8. Por favor, não revele a sua senha a ninguém.
9. É importante que o experimento seja realizado em silêncio, por isso eu gostaria que todas as dúvidas fossem esclarecidas agora, de modo que não seja necessário interromper durante as tarefas. As instruções para a criação da senha estarão presentes na tela de cadastro.
10. SOBRE A SENHA A SER CRIADA:
    - Essa senha deve ser composta de, no mínimo, 6 caracteres;
    - Deve conter pelo menos duas letras e pelo menos dois números (a utilização de outros caracteres também é permitida);
    - [Atribua algum significado à senha, ou associe com alguma informação significativa para você.] (instrução apenas aos grupos experimentais)
    - Não utilize o nome de usuário como senha.
11. As instruções para a criação da senha estarão presentes na tela de cadastro, como lembrete.

Instrução do grupo experimental 1 (repetição elaborativa):

Para lhe ajudar a gravar a sua senha, você irá confirmá-la várias vezes. Ao re-confirmar a sua senha, pense no significado que você atribuiu a essa senha ou na associação que você estabeleceu com algo que lhe é significativo.

Por exemplo, se a senha: "HSL-0101" estiver associada ao fato de que a sua irmã nasceu no Hospital São Lucas da PUC no dia 1º. de janeiro, ao repetí-la você pensará "hospital do nascimento da minha irmã".

Instrução do grupo experimental 2 [pista]:

Para lhe ajudar a lembrar de sua senha, você irá criar uma pista, que será uma palavra ou uma expressão curta. Essa pista deve ser suficiente para lhe ajudar a lembrar da senha, contudo não deve fazer sentido para outras pessoas. Isto é, a pista deve ser vaga o suficiente para dar margem a diversas possibilidades. Lembre-se, a pista que você criará para lhe ajudar a lembrar de sua senha será visível em sua tela.

Por exemplo, se a senha: "HSL-0101" estiver associada ao fato de que a sua irmã nasceu no hospital da PUC no dia 1º. de janeiro. Suponha que sempre que se fala em "PUC" você lembra do nascimento da sua irmã. Sua pista então poderá ser "PUC Maria".