Aproveitando a ciência comportamental para mitigar o risco de segurança cibernética

Shari Lawrence Pfleeger
Instituto de Proteção de Infraestrutura de Informação
Faculdade de Dartmouth
4519 Davenport St. NW
Washington, DC 20016
Telefone: +1 603 729-6023 E-mail:
shari.l.pfleeger@dartmouth.edu

Deanna D. Caputo (autora correspondente)
Corporação MITRE
7515 Colshire Drive
McLean, VA 22102-7539
Telefone: +1 703 983-3846 E-
mail: dcaputo@mitre.org

Aproveitando a ciência comportamental para mitigar o risco de segurança cibernética

Shari Lawrence Pfleeger (Dartmouth College)

Deanna D. Caputo (Corporação MITRE)

Resumo: A maioria dos esforços para melhorar a segurança cibernética concentra-se principalmente na incorporação de novas tecnologias abordagens em produtos e processos. No entanto, um elemento-chave da melhoria envolve o reconhecimento a importância do comportamento humano ao projetar, construir e usar tecnologias de segurança cibernética. Neste artigo de pesquisa, descrevemos por que incorporar uma compreensão do comportamento humano na segurança cibernética produtos e processos podem levar a uma tecnologia mais eficaz. Apresentamos dois exemplos: o primeiro demonstra como a alavancagem da ciência comportamental leva a melhorias claras, e o outro ilustra como a ciência comportamental oferece o potencial para aumentos significativos na eficácia da segurança cibernética. Com base no feedback coletado dos profissionais em entrevistas preliminares, restringimos nosso foco a dois aspectos comportamentais importantes: carga cognitiva e viés. Em seguida, identificamos comportamentos comprovados e potenciais descobertas científicas que têm relevância para a segurança cibernética, não apenas relacionadas à carga cognitiva e ao viés, mas também heurísticas e modelos de ciência comportamental. Concluímos sugerindo vários próximos passos para a incorporação descobertas da ciência comportamental em nosso design, desenvolvimento e uso tecnológico.

Palavras-chave: segurança cibernética, carga cognitiva, viés, heurística, comunicação de risco, modelos de saúde

2 Introdução

"**Apenas amadores atacam máquinas; profissionais atacam pessoas**"(Schneier, 2000).

Qual a melhor forma de lidar com ataques cibernéticos? A segurança cibernética promete proteção e prevenção, utilizando tanto a tecnologia inovadora quanto a compreensão do usuário humano. Quais aspectos do comportamento humano oferecem a maior promessa de tornar os processos e produtos de segurança cibernética mais eficazes? Qual o papel que educação e treinamento? Como podemos incentivar boas práticas de segurança sem prejudicar desnecessariamente

interromper ou irritar os usuários? Como podemos criar um ambiente cibernético que forneça aos usuários todas as funcionalidade de que precisam sem comprometer a segurança empresarial ou nacional? Investigamos as respostas a essas questões examinando a literatura da ciência comportamental para identificar teorias da ciência comportamental e resultados de pesquisas que têm o potencial de melhorar a segurança cibernética e reduzir riscos. Neste artigo, relatar nossas descobertas iniciais, descrever diversas áreas da ciência comportamental que oferecem informações particularmente úteis aplicações para segurança e descrever como usá-las em um processo geral de redução de risco.

O restante deste artigo está organizado em cinco seções. A Seção 2 descreve alguns dos problemas que um uma solução tecnológica por si só não consegue resolver. A Seção 3 explica como usamos um conjunto de cenários para obter sugestões sobre os comportamentos que mais preocupam os designers e usuários de tecnologia. Seções 4 e 5 destacar diversas áreas da ciência comportamental com relevância demonstrada e potencial para a segurança tecnologia. Finalmente, a Seção 6 sugere possíveis próximos passos para a inclusão da ciência comportamental em design, construção e uso de tecnologia de segurança.

3 Por que a tecnologia por si só não é suficiente

Os meios de comunicação social expressam frequentemente a preocupação do sector privado quanto à responsabilidade pelos ataques cibernéticos e à sua ânsia por minimizar os riscos. O setor público tem preocupações semelhantes, porque aspetos da vida quotidiana (como como operação e defesa de infraestrutura crítica, proteção de informações de segurança nacional e operação dos mercados financeiros) envolvem tanto a regulamentação governamental quanto a administração do setor privado.[1] A preocupação do governo é justificada: a União dos Consumidores concluiu que o governo era a fonte de um quinto das violações de dados relatadas publicamente entre 2005 e meados de 2008 (Consumer's Union, 2008). A natureza mutável da tecnologia e do ambiente de ameaças aumenta os riscos para a informação e infraestrutura difícil de antecipar e quantificar.

[1] Veja, por exemplo, o vídeo em http://www.cbsnews.com/video/watch/?id=5578986n&tag=related;photovideo

Os problemas de resposta adequada a incidentes cibernéticos são exacerbados quando a tecnologia de segurança é percebido como um obstáculo para o usuário. O usuário pode ser sobrecarregado por dificuldades de segurança implementação, ou podem desconfiar, interpretar mal ou ignorar a segurança. Um estudo recente com usuários da Virginia A tecnologia ilustra o problema (Virginia Tech, 2011). Bellanger et al. examinaram as atitudes dos usuários e a "comportamento de resistência" de indivíduos diante de uma mudança obrigatória de senha. Os pesquisadores descobriram que, mesmo quando as senhas foram alteradas conforme necessário, as alterações foram intencionalmente atrasadas e a solicitação percebido como uma interrupção desnecessária. "As pessoas estão cientes de que uma violação de senha pode ter consequências graves, mas não afeta sua atitude em relação à implementação da política de segurança". Além disso, "quanto mais competência técnica os entrevistados têm, menos eles são a favor do aprimoramento da política. ...Numa implementação voluntária, essa competência pode ser um vetor de orgulho e realização. contexto obrigatório, o indivíduo pode sentir sua competência desafiada, desencadeando uma atitude negativa em direção ao processo".

No passado, as soluções para estes problemas variaram desde o controlo rigoroso e baseado na tecnologia dos computadores comportamento humano baseado (frequentemente com aplicação inconsistente ou às vezes rígida) para abrangente educação e treinamento de desenvolvedores e usuários de sistemas. Nenhum dos extremos obteve sucesso significativo, mas estudos recentes sugerem que uma combinação dos dois pode levar a resultados eficazes. Por exemplo, o Reino Unido O Escritório de Padrões em Educação, Serviços e Habilidades para Crianças (Ofsted) avaliou a segurança do ensino online comportamento em 35 escolas representativas no Reino Unido "Onde a provisão para segurança eletrônica foi excelente, as escolas conseguiram administrar os sistemas em vez de bloqueá-los. Na melhor prática observada, os alunos foram ajudados, desde muito cedo, avaliar o risco de acesso aos sítios e, assim, adquirir gradualmente competências que os ajudaria a adotar práticas seguras mesmo quando não fossem supervisionados" (Ofsted, 2010). Em outras palavras, palavras, os comportamentos de segurança mais bem-sucedidos foram exibidos em escolas onde os alunos foram ensinados comportamentos apropriados e, em seguida, confiáveis para se comportarem de forma responsável. O relatório do Ofsted compara a abordagem a ensinar as crianças a atravessar a rua com segurança, em vez de depender dos adultos para acompanhá-las do outro lado da rua todas as vezes.

Esta abordagem está no cerne da nossa pesquisa. A nossa hipótese principal é que, se os humanos usarem computadores os sistemas recebem as ferramentas e informações de que necessitam, são ensinados sobre o significado do uso responsável e, então, confiável para se comportar adequadamente em relação à segurança cibernética, os resultados desejados podem ser obtidos sem a segurança é percebida como onerosa ou penosa. Ao compreender o papel do comportamento humano e aproveitando as descobertas da ciência comportamental, os designers, desenvolvedores e mantenedores de informações a infraestrutura pode abordar obstáculos reais e percebidos à produtividade e proporcionar uma gestão mais eficaz segurança. Essas mudanças comportamentais levam tempo, portanto, os planos para iniciar a mudança devem incluir tempo suficiente propor a mudança, implementá-la e torná-la parte da cultura ou prática comum.

Estão a começar a surgir outras evidências (Predd et al., 2008; Pfleeger et al., 2010) que apontam para a importância de entender os comportamentos humanos ao desenvolver e fornecer segurança cibernética.[2]Há interesse particular em usar a confiança para mitigar riscos, especialmente online. Por exemplo, a União Europeia financiou um projeto multidisciplinar de vários anos sobre confiança online (iTrust),[3]documentando as muitas maneiras que a confiança pode ser criada e quebrada. Agora, estão sendo desenvolvidas estruturas para analisar o grau de qual confiança é construída e mantida em aplicações de computador (Riegelsberger, Sasse e McCarthy, 2005). De forma mais ampla, uma literatura rica e relevante em ciências comportamentais aborda problemas críticos de segurança, como como desvio dos funcionários, conformidade dos funcionários, tomada de decisão eficaz e o grau em que emoções (Lerner e Tiedens, 2006) ou condições estressantes (Klein e Salas, 2001) podem levar a situações mais arriscadas escolhas dos tomadores de decisão.[4]Ao mesmo tempo, há muitas evidências de que os avanços tecnológicos podem ter consequências não intencionais que reduzem a confiança ou aumentam o risco (Tenner, 1991). Por estas razões, concluem que é importante incluir o elemento humano ao projetar, construir e usar recursos críticos sistemas.

[2]Veja o Primeiro Workshop Interdisciplinar sobre Segurança e Comportamento Humano, descrito em http://www.schneier.com/blog/archives/2008/06/security_and_.html http://www.cl.cam.ac.uk/~rja14/shb08.html

[3]Veja os artigos do workshop emhttp://www.informatik.uni-trier.de/~ley/db/conf/itrust/itrust2006.html

[4]O programa da National Science Foundation está interessado nas conexões entre ciências sociais e segurança cibernética. Anunciou um novo programa que incentiva cientistas da computação e cientistas sociais a trabalharem juntos (Ciberespaço Seguro e Confiável, descrito em http://www.nsf.gov/pubs/2012/nsf12503/nsf12503.htm?WT.mc_id=USNSF_25&WT.mc_ev=click).

Para entender como projetar e construir sistemas que incentivem os usuários a agir de forma responsável ao utilizá-los, identificámos dois tipos de descobertas da ciência comportamental: aquelas que já demonstraram um efeito bem-vindo na implementação e utilização da segurança cibernética e aqueles com potencial para tal efeito. No primeiro caso, documentamos as descobertas relevantes, para que os profissionais e pesquisadores possam determinar quais abordagens são mais aplicáveis ao seu ambiente. No segundo caso, estamos projetar uma série de estudos para testar resultados promissores da ciência comportamental em um ambiente de segurança cibernética com o objetivo de determinar quais resultados (com estratégias associadas para reduzir ou mitigar os efeitos comportamentais problemas que refletem) são os mais eficazes.

No entanto, aplicar descobertas da ciência comportamental a problemas de segurança cibernética é uma tarefa enorme. Para maximizar a provável eficácia dos resultados, usamos um conjunto de entrevistas para obter informações dos profissionais opiniões sobre comportamentos preocupantes para que pudéssemos nos concentrar naqueles percebidos como mais significativos. descrevem as entrevistas e os resultados na Seção 3. Essas descobertas sugerem hipóteses sobre o papel de comportamento no enfrentamento de questões de segurança cibernética.

4 Identificando Aspectos Comportamentais da Segurança

Os designers e desenvolvedores de tecnologia de segurança podem aproveitar o que se sabe sobre as pessoas e seus percepções para fornecer uma segurança mais eficaz. Um ex-chefe de segurança do aeroporto israelense disse:

> "Eu digo que a tecnologia deve apoiar as pessoas. E as pessoas qualificadas devem estar no centro da nossa conceito de segurança e não o contrário" (Amos, 2010).

Para implementar este tipo de segurança centrada no ser humano, os tecnólogos devem compreender o comportamento ciências à medida que projetam, desenvolvem e usam a tecnologia. No entanto, traduzir os resultados comportamentais para uma ambiente tecnológico pode ser um processo difícil. Por exemplo, os projetistas de sistemas devem abordar o elementos humanos obscurecidos pela mediação do computador. Os consumidores que fazem uma compra online confiam que o

O comerciante representado pelo site não está apenas pegando seu dinheiro, mas também cumprindo sua obrigação para fornecer bens em troca. O consumidor infere o envolvimento humano do comerciante online por trás do cenas. Assim, em algum nível, o comprador e o vendedor são humanos realizando uma transação habilitada por um sistema projetado, desenvolvido e mantido por humanos. Pode não haver contato humano real nem contato direto conhecimento dos outros atores humanos envolvidos, mas o processo de transação reflete sua contraparte humana.

A prevenção ou mitigação de incidentes adversos de segurança cibernética requer ações em várias etapas: concepção da tecnologia sendo incorporada na infraestrutura; implementando, testando e mantendo a tecnologia; e usar a tecnologia para fornecer produtos e serviços essenciais. A ciência comportamental tem abordaram noções de segurança cibernética nessas atividades por muitos anos. De fato, Sasse e Flechais (2005) note que os sistemas seguros são sistemas sociotécnicos nos quais devemos usar uma compreensão de ciência comportamental para “evitar que os usuários sejam o 'elo mais fraco'”. Por exemplo, algumas cientistas investigaram como os mecanismos de confiança afetam a segurança cibernética. Outros relataram descobertas relacionados com a concepção e utilização de sistemas cibernéticos, mas a relevância e o grau de efeito ainda não foram testado.

Algumas das ligações entre a ciência comportamental e a segurança são específicas de certos tipos de sistemas. Por exemplo, Castelfranchi e Falcone (1998 e 2002) analisam a confiança em sistemas multiagentes a partir de uma perspectiva comportamental. Eles veem a confiança como tendo vários componentes, incluindo crenças que devem ser mantidas para desenvolver a confiança (o contexto social, conforme descrito por Riegelsberger, Sasse e McCarthy (2003)) e relações com interações anteriores (o contexto temporal do Riegelsberger-Sasse-McCarthy estrutura). Eles usam fatores psicológicos para modelar a confiança em sistemas multiagentes. Além dos fatores sociais e preocupações temporais, acrescentamos expectativas de realização, onde alguém confia em alguém ou outra pessoa espera algo em troca (Baier, 1986). Esta pesquisa comportamental lança luz sobre a natureza da expectativa do usuário e na confiabilidade percebida das interações mediadas pela tecnologia e tem implicações importantes relacionadas ao projeto de sistemas e processos de proteção.

Sasse e Flechais (2005) veem a segurança a partir de três perspectivas distintas: produto, processo e panorama.

- **Produto.**Esta perspectiva inclui o efeito dos controlos de segurança, como as políticas e mecanismos sobre as partes interessadas (por exemplo, designers, desenvolvedores, usuários). Os controles envolvem requisitos que afetam a carga de trabalho física e mental, comportamento e custo (humano e financeiro). Os usuários confiam que o produto manterá a segurança enquanto executa a tarefa principal.

- **Processo.**Este aspecto aborda como as decisões de segurança são tomadas, especialmente nos estágios iniciais de coleta e projeto de requisitos. O processo deve permitir que os mecanismos de segurança sejam “um parte integrante do design e desenvolvimento do sistema, em vez de ser 'adicionado'” (Sasse e Flechais, 2005). Porque “mecanismos que não são empregados na prática, ou que são usados incorretamente, fornecem pouca ou nenhuma proteção”, os projetistas devem considerar as implicações de cada mecanismo sobre carga de trabalho, comportamento e fluxo de trabalho (Sasse e Flechais, 2005). A partir disso perspectiva, as partes interessadas devem confiar no processo para que possam tomar decisões adequadas e decisões eficazes, especialmente sobre suas tarefas principais

- **Panorama.**Este aspecto descreve o contexto em que a segurança opera. Porque a segurança é geralmente não é a tarefa principal, os usuários provavelmente “procurarão atalhos e soluções alternativas, especialmente quando os usuários não entendem por que seu comportamento compromete a segurança...Uma segurança positiva cultura, baseada numa compreensão partilhada da importância da segurança...é a chave para alcançar comportamento desejado” (Sasse e Flechais, 2005). Nessa perspectiva, o usuário visualiza a segurança mecanismos como essenciais mesmo quando parecem intrusivos, limitantes ou contraproducentes.

4.1 Criação de Cenários

Como os tipos de infraestrutura e as ameaças são vastos, usamos os resultados das entrevistas para restringir nossa investigação para aquelas áreas da ciência comportamental com potencial demonstrado ou provável de aumentar a confiança de um ator em usando qualquer infraestrutura de informação. Para orientar nossas entrevistas, trabalhamos com duas dúzias de jornalistas americanos funcionários do governo e da indústria familiarizados com questões de proteção de infraestrutura de informação para definir três cenários de ameaças relevantes para a proteção da infraestrutura de informação. A metodologia e os resultados

as análises foram conduzidas pelo primeiro autor do artigo e envolveram cinco etapas:

- *Escolhendo tópicos.*Selecionamos três tópicos de segurança para discutir, com base em eventos recentes.
  a combinação dos três pretendia representar uma (reconhecidamente incompleta, mas) significativa
  uma série de preocupações típicas, cuja discussão revelaria áreas subjacentes propícias para
  melhoria.
- *Criar um cenário representativo e realista para cada tópico.*Usando nosso conhecimento de cibersegurança recente
  incidentes e ataques, criamos um cenário de ataque para cada tópico plausível, retratando um ciberataque
  problema de segurança cuja solução seria bem-vinda pela indústria e pelo governo.
- *Identificar pessoas com autoridade para tomar decisões sobre produtos de segurança cibernética e seu uso*
  *entrevista sobre os cenários.*Identificamos pessoas da indústria e do governo que estavam
  disposto a participar de entrevistas.
- *Realização de entrevistas.*As nossas discussões centraram-se em duas questões: Serão estes cenários realistas?
  e como a segurança cibernética poderia ser melhorada em cada situação?
- *Analisando os resultados e suas implicações.*Analisamos os resultados dessas entrevistas e suas
  implicações para nossa pesquisa.

Cenário 1: Melhorando a conscientização sobre segurança entre os construtores de infraestrutura de informação

A segurança raramente é a principal tarefa de quem utiliza a infraestrutura de informação. Normalmente, os usuários buscam informações, analisar relacionamentos, produzir documentos e executar tarefas que os ajudem a entender situações e agir. Da mesma forma, os desenvolvedores de sistemas geralmente se concentram nessas tarefas principais antes incorporando segurança em uma arquitetura ou design. Além disso, os desenvolvedores de sistemas frequentemente implementam requisitos de segurança escolhendo mecanismos de segurança que sejam fáceis de construir e testar ou que atendam a alguns outro objetivo técnico do sistema (por exemplo, confiabilidade). Os desenvolvedores raramente levam em consideração a usabilidade do mecanismo ou a carga cognitiva adicional que ele coloca no usuário. O cenário 1 descreve maneiras de melhorar

conscientização sobre segurança entre os desenvolvedores de sistemas para que a segurança tenha mais probabilidade de ser útil e eficaz.

Suponha que os engenheiros de software estejam projetando e construindo um sistema para dar suporte à criação e transmissão de documentos sensíveis entre os membros de uma organização. Muitos aspectos da criação e gerenciamento de documentos transmissão são bem conhecidos, mas mecanismos de segurança para avaliar a sensibilidade, rotular documentos a transmissão adequada e segura de documentos tem apresentado dificuldades há muitos anos. Em nossa cenário, os engenheiros de software são encarregados de projetar um sistema que solicita informações de documentos criadores, modificadores e leitores, para que uma designação de confiança possa ser atribuída a cada documento. Segurança As questões incluem a compreensão dos tipos de informações relacionadas à confiança necessárias, a determinação do papel de uma mudando o ambiente de ameaças e definindo a frequência com que as informações de confiança devem ser atualizados e reavaliados (particularmente à luz de incidentes de segurança cibernética que podem ocorrer durante a vida do documento). Além disso, os engenheiros de software devem implementar algum tipo de resumo de confiança designação que terá significado para criadores, modificadores e leitores de documentos.

Esta designação de confiança, diferente da classificação de sensibilidade do documento, representa o grau de em que tanto o conteúdo quanto o provedor (ou modificador) podem ser confiáveis e por quanto tempo. Por exemplo, um documento sobre a capacidade militar emergente de uma nação pode ser altamente confidencial (ou seja, altamente sensíveis), independentemente de o fornecedor de informações ser altamente confiável (porque, por exemplo, ele tem forneceu repetidamente informações altamente úteis no passado) ou não (porque, por exemplo, ele frequentemente fornece informações incorretas ou enganosas).

Existem dois aspectos importantes da conscientização de segurança dos engenheiros de software. Primeiro, eles devem ser capazes de selecionar mecanismos de segurança para implementar a designação de confiança que lhes permitam equilibrar a segurança com os requisitos de desempenho e usabilidade. Este equilíbrio implica apreciar e acomodar o papel da segurança no contexto mais amplo da finalidade pretendida do sistema e dos seus múltiplos usos. Em segundo lugar, os utilizadores deve ser capaz de confiar que o mecanismo de segurança apropriado foi escolhido. Confiança significa que o mecanismo em si deve ser apropriado à tarefa. Por exemplo, o Modelo de Integridade Biba (Biba, 1977), um sistema de As políticas de segurança informática, expressas como regras de controlo de acesso, visam garantir a integridade dos dados.

O modelo define uma hierarquia de níveis de integridade e, em seguida, impede que os participantes corrompam os dados de um nível de integridade superior ao do assunto, ou de ser corrompido por dados de um nível inferior ao assunto. O modelo Biba foi desenvolvido para estender o modelo Bell-La Padula (1973), que aborda apenas confidencialidade dos dados. Assim, a compreensão e a escolha de políticas e mecanismos são aspectos importantes na que confiamos que os engenheiros de software exerçam discrição. Além disso, os engenheiros de software devem ser capazes de confiar na procedência, correção e conformidade com as expectativas dos mecanismos de segurança. Aqui, "proveniência" significa não apenas a aplicabilidade dos mecanismos e algoritmos, mas também a fonte de módulos arquitetônicos ou de implementação. Com a disponibilidade de módulos de código aberto e linha de produtos arquiteturas (ver, por exemplo, Clements e Northrup, 2001), é provável que algumas partes de algumas mecanismos de segurança terão sido criados para uma finalidade diferente, geralmente por uma equipe diferente de engenheiros. Os construtores e modificadores do sistema atual devem saber até que ponto confiar nos módulos de outra pessoa.

Cenário 2: Aumentar a consciência situacional durante uma "Evento cibernético"

Consciência situacional é o grau em que uma pessoa ou sistema sabe sobre uma ameaça no ambiente. Quando uma emergência se desenrola, as pessoas e os sistemas envolvidos na observação do seu desenrolar devem determinar o que já aconteceu, o que está acontecendo atualmente e o que provavelmente acontecerá no futuro; então, eles fazem recomendações de reação com base em sua consciência situacional. As pessoas ou sistemas percebendo a situação têm diferentes graus de confiança nas informações que coletam e nos provedores dessas informações. Quando um evento cibernético está ocorrendo, as informações podem vir de fontes primárias (como como sensores em sistemas de controle de processos ou medições de atividade de rede) e fontes secundárias (como como intérpretes humanos ou automatizados de tendências).

Considere analistas usando um sistema de computador que monitora a rede de sistemas de energia nos Estados Unidos Estados. O próprio sistema interage com uma rede de sistemas, cada um dos quais coleta e analisa dados sobre estações de geração e distribuição de energia e seus pontos de acesso. Os analistas notam uma série de falhas de rede em todo o país: primeiro, uma central eléctrica na Califórnia falha, depois uma no Missouri, e assim por diante

durante as primeiras horas do evento.[5]Os analistas devem determinar não apenas o que realmente está acontecendo mas também como responder adequadamente. A segurança e o comportamento humano estão envolvidos de muitas maneiras. Primeiro, a O analista deve saber se deve confiar nas informações que estão sendo reportadas ao seu sistema de monitoramento. Por exemplo, o analista está visualizando uma falha no ponto de acesso ou no sistema de monitoramento? Em seguida, o analista deve ser capaz de saber quando e se ela tem informações suficientes para tomar uma decisão sobre qual as reações são apropriadas. Esta decisão deve ser tomada no contexto de uma situação em evolução, onde alguns evidências inicialmente consideradas confiáveis são eventualmente determinadas como não confiáveis (e vice-versa). Finalmente, a O analista deve analisar os dados relatados, formular hipóteses sobre as possíveis causas e, em seguida, determinar qual interpretação dos dados usar. Por exemplo, a sequência de falhas é resultado de dados incorretos transmissão, um ataque cibernético, falhas aleatórias do sistema ou simplesmente as várias empresas de energia tendo comprou parte do seu software do mesmo fornecedor (cujo sistema agora está falhando)? Escolhendo o uma interpretação errada pode ter consequências sérias.

Cenário 3: Apoiando decisões sobre confiabilidade de transações de rede

No dia de Natal de 2009, um estudante nigeriano que voava de Amsterdã para Detroit tentou detonar uma bomba bomba para destruir o avião. Felizmente, a bomba causou poucos danos e os passageiros impediram o estudante de completar a tarefa pretendida. No entanto, ao analisar por que o aluno não foi detectado por uma variedade das telas de segurança do aeroporto, foi determinado que informações importantes nunca foram apresentadas ao tomadores de decisão apropriados (Baker e Hulse, 2009). Esta situação constitui o cerne do Cenário 3, onde um sistema consulta um conjunto interconectado de bancos de dados para encontrar informações sobre uma pessoa ou situação.

Neste cenário, um analista usa uma interface para uma coleção de repositórios de dados, cada um dos quais contém informações sobre crime e terrorismo. Quando o analista recebe um aviso sobre uma determinada pessoa de interesse, ela deve consultar os repositórios para determinar o que se sabe sobre essa pessoa. Existem muitos questões de segurança relacionadas a este cenário. Primeiro, o analista deve determinar o grau em que pode confiar

[5]De fato, nesta fase, pode não estar claro se o evento é realmente um evento cibernético. Um evento semelhante, com características semelhantes, ocorreu em 14 de agosto de 2003, nos Estados Unidos. Vejahttp://www.cnn.com/2003/US/08/14/power.outage/index.html

que todas as informações relevantes residem em pelo menos um dos repositórios conectados. Após o Tentativa de atentado de Natal, foi revelado que o Reino Unido negou um pedido de visto ao estudante, mas informações sobre a negação não estavam disponíveis para a Administração de Segurança de Transporte quando foram tomadas decisões sobre se o aluno deveria ser submetido a uma triagem de segurança adicional. Spira (2010) aponta salienta que o problema não é o número de bases de dados; é a falta de capacidade de pesquisar em toda a base de dados. "federação" de bases de dados.

Em seguida, mesmo que os itens relevantes sejam encontrados, os mais importantes devem estar visíveis no local apropriado. tempo. Libicki e Pfleeger (2004) documentaram as dificuldades em "coletar os pontos" antes de uma O analista pode dar o próximo passo para conectá-los. Se um "ponto" não estiver tão visível quanto deveria, pode ser negligenciados ou que receberam atenção insuficiente durante a análise subsequente. Além disso, Spira (2010) destaca a necessidade de visualizar as informações em seu contexto apropriado.

Em terceiro lugar, o analista também deve determinar o grau em que cada informação relevante pode ser confiável. Ou seja, ela não só deve saber a precisão e a atualidade de cada item de dados, mas também deve determinar se a própria fonte de dados é confiável. Existem vários aspectos neste último grau de confiança, como saber com que frequência a fonte de dados fornece as informações (ou seja, se são antigas notícias), sabendo se a fonte de dados é suficientemente confiável e se as circunstâncias podem mudar a confiabilidade da fonte. Por exemplo, Predd et al. (2008) e Pfleeger et al. (2010) apontam a Vários tipos de pessoas com acesso legítimo aos sistemas que realizam ações indesejadas. Um insider confiável pode se tornar uma ameaça devido a uma demissão pendente ou problema pessoal, desatenção ou confusão, ou ela tentar superar uma fraqueza do sistema. Portanto, a confiabilidade das informações e fontes deve ser re-avaliados repetidamente e talvez até previstos com base em previsões sobre um ambiente em mudança.

Por fim, o analista também deve determinar o grau de correção da análise. Qualquer análise envolve suposições sobre variáveis e sua importância, bem como as relações entre variáveis dependentes e variáveis independentes. Muitas vezes, é uma suposição equivocada que leva ao fracasso, e não dados falhos.

4.2 Análise dos Resultados

Os três cenários foram intrigantes para os nossos entrevistados e todos concordaram que eram realistas e relevantes e importante. No entanto, a análise dos cenários pelos entrevistados revelou menos comportamentos insights do que esperávamos. Em cada caso, o entrevistado viu cada cenário de sua perspectiva particular perspectiva, destacando apenas uma pequena parte do cenário para confirmar uma opinião que ele ou ela tinha. Por exemplo, um dos entrevistados utilizou o Cenário 3 para enfatizar a necessidade de partilha de informação; outro O entrevistado disse que a privacidade é uma preocupação fundamental, especialmente em situações como o Cenário 2, onde há o monitoramento deve ser equilibrado com a proteção da privacidade.

No entanto, muitos dos entrevistados apresentaram boas sugestões para moldar o caminho a seguir. Por exemplo, um deles disse que há muito a aprender com algoritmos de comando e controle, onde os atores militares aprenderam a lidar com a percepção de risco, a incerteza, a informação incompleta e a necessidade de tomar uma decisão decisão importante sob pressões extremas. Existe uma rica literatura abordando a tomada de decisão sob pressão, de Ellsberg (1964) a Klein (Klein, 1998; Klein, 2009). Em particular, os modelos de Klein de a tomada de decisão adaptativa pode ser aplicável (Klein e Calderwood, 1991; Klein e Salas, 2001). Embora a metodologia do cenário não fosse uma abordagem estruturada de geração de ideias, na medida do possível, nos esforçamos para ser imparciais em nossa interpretação das respostas dos entrevistados. Não estávamos tentando reunir apoio a ideias preconcebidas e estavam genuinamente tentando explorar novas ideias onde a ciência comportamental poderia ser aproveitado para resolver problemas de segurança.

Vários temas emergiram das entrevistas:

- **A segurança está interligada com a maneira como os humanos se comportam ao tentar atingir uma meta ou executar uma tarefa.**A separação da tarefa primária da secundária, bem como seu impacto no comportamento do usuário, foi expressa pela primeira vez de forma clara em Smith et al. (1997) e elaborada no âmbito da segurança por Sasse et al. (2002). Nossas entrevistas reconfirmaram que, na maioria dos casos, a segurança é secundária à segurança do usuário. tarefa principal (por exemplo, encontrar uma informação, processar uma transação, tomar uma decisão).

Quando a segurança interfere, a pessoa pode ignorar ou até mesmo subverter a segurança, já que a pessoa está recompensado pela tarefa principal. Em certo sentido, a pessoa confia no sistema para cuidar da segurança preocupações. Essa perspectiva pode levar a pelo menos dois eventos indesejáveis. Primeiro, quando confrontado com incerteza sobre a segurança de um curso de ação, a pessoa*trusts*que o sistema tem garantiu a segurança da ação (por exemplo, quando um usuário abre um anexo assumindo que o o sistema verificou se há vírus ou, como no Cenário 3, os usuários presumiram que o bombardeiro não era um risco de segurança porque seu nome não foi revelado pelo sistema de segurança). Em segundo lugar, quando, no passado, os recursos de segurança impediram ou retardaram a conclusão da tarefa, um usuário subverte a segurança porque ele ou ela não pode mais*confiar*o sistema para permitir a conclusão eficaz das tarefas no futuro. Assim, compreender a ciência comportamental (em vez da segurança em si) pode oferecer novas maneiras de projetar, construir e usar sistemas cuja segurança seja compreendida e respeitada pelo usuário.

- Entrevistados observados em todos os cenários**como limitações na memória ou na capacidade de análise interferiu com um analista'capacidade de executar**. Um entrevistado observou a abundância de informações geradas por sistemas automatizados e a crescente probabilidade de que dados importantes eventos passariam despercebidos (Burke 2010). Nas ciências comportamentais, o termo*carga cognitiva* refere-se à quantidade de estresse colocado na memória de trabalho. Abordado pela primeira vez por Miller (1956), que afirmou que a "memória de trabalho" de uma pessoa poderia lidar com no máximo cinco a nove pedaços de informações de uma só vez, a noção foi estendida por Chase e Simon (1973) para abordar a memória sobrecarga durante a resolução de problemas. Vários resultados empíricos (ver, por exemplo, Scandura, 1971) sugerem que os indivíduos variam em sua capacidade de processar uma determinada quantidade de informação.

- ***Cegueira desatenta*é um aspecto particular da carga cognitiva que desempenhou um papel em cada cenário.**Reconhecido pela primeira vez por Mack e Rock (1998) e extensivamente estudado por Simons e seus colegas (ver, por exemplo, Simons e Chabris, 1999 e Simons e Jensen, 2009), cegueira desatenta refere-se à incapacidade de uma pessoa de perceber eventos inesperados quando concentrando-se em uma tarefa principal. Por exemplo, a cegueira desatenta pode fazer com que um analista

Cenário 2 para não ver um padrão na falha de usinas de energia (por exemplo, que todas as usinas de energia falham estavam em áreas que enfrentavam seca severa), ou para levar um analista no Cenário 3 a ignorar uma aviso do pai do terrorista porque a atenção estava restrita ao próprio terrorista.

- **Há um viés significativo na maneira como cada entrevistado pensa sobre segurança.**Este preconceito reflecte a experiência, os objetivos e a especialização do entrevistado, evidenciando-se na forma como duas pessoas veem a mesma situação de maneiras muito diferentes. Por exemplo, entrevistados com empregos que se concentram principalmente na privacidade, pensei nos cenários como proteção de dados de pessoas de fora, mas não considerar a corrupção inadvertida. Ao compreender os vieses, os designers e desenvolvedores de segurança podem antecipar percepções prováveis e levá-las em conta ao projetar abordagens para encorajar o bem comportamento de segurança.

- **Existe um elemento significativo de risco em cada cenário e os tomadores de decisão têm dificuldade em tempo, tanto na compreensão da natureza do risco (expresso como uma combinação de probabilidade e impacto) e equilibrar múltiplas percepções do risco para tomar a melhor decisão no tempo disponível.**Existe uma literatura considerável sobre percepção e comunicação de risco, com artigos importantes incluídos nas compilações de Mayo e Hollander (1991) e Slovic (2000). Ao aplicar as descobertas da ciência comportamental ao design, desenvolvimento e uso do sistema, os usuários podem ser mais conscientes do provável impacto de suas decisões relacionadas à segurança.

As entrevistas revelaram como os profissionais (ou seja, usuários e desenvolvedores) envolvem e não envolvem a segurança. preocupações relacionadas em seu processo de tomada de decisão. Vários pontos ficaram claros para nós como resultado dessas discussões:

- Os profissionais não têm um entendimento comum sobre segurança.

- Os profissionais não têm uma consciência maior de como a segurança pode afetar todo o seu trabalho funções e papéis. Por exemplo, as pessoas se sentem confortáveis revelando pequenas quantidades de informação em cada situação, mas não percebem a facilidade com que as informações podem se agregar em um quadro completo

isso se torna uma preocupação de segurança.

- Os profissionais têm experiência limitada na análise de uma situação para identificar as medidas de segurança necessárias relacionamentos.
- A combinação de foco estreito com uma grande (e muitas vezes crescente) quantidade de informação continua a causar falha em “conectar os pontos”. Encontrar um padrão ou conexão entre apenas um poucos pontos dentro de um grande conjunto de dados é semelhante ao problema de identificar uma constelação em uma estrela-céu noturno cheio. Algumas pessoas conseguem encontrar a Ursa Maior facilmente, enquanto outras veem apenas muitas estrelas. Nossas entrevistas deixaram claro que os profissionais precisam de treinamento e assistência para identificar aspectos importantes de uma situação e em saber como e quando focar.

Com base nos resultados das nossas discussões de cenários, reduzimos o nosso foco à carga cognitiva e ao viés como princípios organizadores para uma investigação de teorias relevantes da ciência comportamental e resultados de pesquisas que prometem sistemas mais seguros. Também buscamos informações sobre heurísticas e modelos que podem ser úteis para nos ajudar a transmitir informações de segurança cibernética e implementar resultados. Nas próximas duas seções, examinamos as descobertas da ciência comportamental que já foram foi demonstrado que tem impacto na segurança cibernética e naqueles com potencial para tal.

5 Áreas da Ciência Comportamental com Relevância Demonstrada

Começamos esta seção examinando várias descobertas importantes da ciência comportamental que foram demonstradas relevantes para a segurança cibernética em geral e para a proteção da infraestrutura de informação em particular. Então, no Na próxima seção, examinaremos pesquisas em ciências comportamentais que têm potencial para melhorar a segurança cibernética. Além disso, incluímos descrições de heurísticas e modelos relacionados à saúde que podem auxiliar os designers em Construindo uma boa segurança em produtos e processos. Em cada caso, documentamos as possíveis implicações de cada um.

5.1 Resultados com relevância demonstrável para a segurança cibernética

As descobertas da ciência comportamental melhoram o produto, o processo e o panorama nesses exemplos.

O reconhecimento é mais fácil do que a recordação

A literatura da ciência comportamental demonstra que o reconhecimento é significativamente mais fácil do que a recordação. Rock e Engelstein (1959) mostraram às pessoas uma única forma sem sentido, a capacidade dos participantes de se lembrarem diminuiu rapidamente, mas eles puderam reconhecê-lo quase perfeitamente um mês depois. Em outras palavras, perguntando participantes para recordar uma forma sem que lhes fossem mostrados exemplos teve muito menos sucesso do que exibir uma coleção de formas e pedindo-lhes que identificassem qual delas lhes havia sido mostrada inicialmente. Nas duas décadas seguintes, muitos estudos empíricos em larga escala reforçaram essa descoberta. Por exemplo, Standing (1973) mostrou aos participantes um conjunto de imagens complexas; o número de imagens em cada conjunto variou de 10 para 10.000. Os participantes conseguiram reconhecer subconjuntos deles com 95% de precisão.

Dhamija e Perrig (2000) estudaram o quão bem as pessoas se lembram de imagens em comparação com senhas, e descobriram que as pessoas conseguem reconhecer a imagem escolhida com mais segurança do que lembrar de uma senha selecionada. Este resultado está sendo aplicado à autenticação do usuário para o computador; o usuário seleciona uma imagem como imagem de autenticação ou seleciona uma senha de uso único com base em um formato ou configuração. Da mesma forma, Zviran e Haga (1990) mostraram que mesmo mecanismos de desafio-resposta baseados em texto e mecanismos associativos As senhas são uma melhoria em relação à recuperação de senhas sem ajuda.

Produtos comerciais estão usando esses resultados. Lamandé (2010) relata que a autenticação GrIDSure sistema (http://www.gridsure.com) foi integrado ao Unified Access Gateway (UAG) da Microsoft plataforma. Este sistema permite que um usuário se autentique com uma senha única baseada em um padrão de quadrados escolhidos de uma grade. Quando o usuário deseja acessar, é apresentada uma grade contendo números atribuídos aleatoriamente; ela então insere como sua senha os números que correspondem aos números escolhidos padrão. Como os números da grade exibidos mudam cada vez que a grade é apresentada, o padrão permite

a senha inserida seja um código de uso único. Muitos pesquisadores (ver, por exemplo, Sasse, 2007; Bond, 2008; Biddle, Chiasson e van Oorschot, 2009) examinaram aspectos da segurança e usabilidade.

Outros produtos comerciais usam imagens chamadas Passfaces. Introduzido há mais de dez anos (Brostoff e Sasse, 2000) e avaliados repetidamente (Everitt et al., 2009), os Passfaces oferecem uma opção que aborda o desvantagens de produtos como o GrIDSure. No entanto, o estudo da Consumer's Union (2008) e outros documentar o grau em que o usuário médio gerencia múltiplas senhas — às vezes dezenas! a segurança em geral leva a problemas que também são compartilhados com o reconhecimento de imagem: interferência.

Interferência

Mudanças frequentes em um item memorizado interferem na memorização da nova versão do item. Ou seja, a versão mais recente do item compete com as anteriores. A frequência de mudança é importante; por exemplo, Underwood (1957) descobriu que, em estudos em que os participantes eram solicitados a memorizar apenas algumas listas anteriores, seu nível de esquecimento foi muito menor do que em estudos onde o os participantes foram solicitados a memorizar muitas listas anteriores. Wixted (2004) aponta que mesmo diferentes as coisas podem interferir em algo que um sujeito está tentando memorizar: "...memórias recentemente formadas que ainda não tiveram a oportunidade de se consolidar são vulneráveis à força interferente da atividade mental e formação de memória (mesmo que a atividade interferente não seja semelhante ao material aprendido anteriormente)."

Em estudos empíricos que aplicam essas descobertas à memorização de senhas, Sasse, Brostoff e Weirich (2002) mostraram que as falhas de login aumentaram drasticamente à medida que as mudanças de senha necessárias se tornaram mais frequentes. Além disso, Brostoff e Sasse (2003) mostraram que permitir mais tentativas de login levou a um maior sucesso sessões de login; eles sugerem que sistemas tolerantes resultam em melhor conformidade do que aqueles muito restritivos.

Everitt et al. (2009) e Chiasson et al. (2009) examinaram o uso de múltiplas senhas gráficas. Eles descobriram que os usuários com múltiplas senhas gráficas cometiam menos erros ao lembrá-las, não crie senhas que estejam diretamente relacionadas aos nomes das contas e não use senhas semelhantes em todas as contas

várias contas. Além disso, mesmo depois de duas semanas, as taxas de sucesso de recall permaneceram boas com gráficos senhas e eram melhores do que aqueles com senhas de texto. Assim, parecia haver menos interferência com objetos gráficos do que com objetos textuais.

Estudos recentes abordaram preocupações adicionais sobre recall e interferência. Por exemplo, Jhawar et al. (2011) sugerem que um bom design pode superar esses problemas e que a recordação gráfica pode formar a base para práticas de segurança eficazes.

Outros Estudos na Intersecção

Além das descobertas citadas acima, a maioria das quais são extraídas da psicologia cognitiva básica literatura, há muitos exemplos de estudos aplicados de outras disciplinas onde cientistas comportamentais estudaram diretamente problemas relacionados à cibernética. Por exemplo,

- **Sociologia**. Cheshire e Cook (2004) aplicaram resultados de pesquisas sociológicas experimentais a quatro diferentes categorias de interação mediada por computador. Eles oferecem orientação aos cientistas da computação sobre como construir confiança em redes online. Por exemplo, eles sugerem tratar a informática interação mediada como um problema arquitetônico, usando a natureza da mediação para moldar comportamento desejado. Eles distinguem entre parceiros aleatórios e fixos em uma transação, e sugerir mecanismos apropriados para interação com base nessa caracterização (ver Figura 1).

	Continuidade	
	Parceiro aleatório	*Parceiro Fixo*
Freqüência *Iterado Interação*	• Solicitação por e-mail • Anexos de e-mail do desconhecido indivíduos	(nenhum)
Um tiro Interação	• Troca de bens digitais ponto a ponto • Jogos online “pickup”	• Comunidades online • Leilões online • Grupos de bate-papo • Jogos online multijogador massivos

Figura 1: Exemplo de recomendações arquitetônicas (Cheshire e Cook, 2004)

- **Economia**. Os economistas estudam o papel da reputação no estabelecimento da confiança, e esta literatura é frequentemente referenciado em trabalhos na intersecção entre economia e segurança cibernética. Por exemplo, muitos dos artigos nos Workshops sobre Economia da Segurança da Informação alavancaram resultados econômicos da pesquisa de reputação. Yamagishi e Matsuda (2003) propõem o uso de informações baseadas na experiência sobre reputação para abordar o problema dos limões: decepção na expectativa. Eles mostram que a decepção é substancialmente reduzida quando os comerciantes online podem mudar livremente suas identidades e cancelar suas reputações.
- **Psicologia e economia**. Existe uma interação entre os custos reais e os custos percebidos quando as pessoas interagem, principalmente online. A pesquisa nesta área abrange tanto a psicologia (a percepção) e economia (os custos reais). Datta e Chatterjee (2008) aplicaram algumas dessas pesquisas sobre a transferência de confiança nos mercados eletrônicos. Eles mostram que a transferência é completo somente se os custos de agência da intermediação estiverem dentro dos limites do consumidor.

Estes exemplos convencem-nos de que explorar mais exaustivamente a literatura das ciências comportamentais conduzirá a uma Base empírica para melhorias na qualidade e eficácia da defesa da segurança cibernética. Esta seção forneceu exemplos da aplicação direta da pesquisa em ciências comportamentais a problemas de cibersegurança segurança. Na próxima seção, consideramos outras áreas onde o aproveitamento da ciência comportamental pode gerar retorno benefícios significativos na proteção da infraestrutura de informações.

6 Áreas da Ciência Comportamental com Potencial Relevância

Há uma quantidade significativa de pesquisas em ciências comportamentais sobre métodos ou conceitos que influenciam um percepções, atitudes e comportamentos de uma pessoa ou grupo. Muitas descobertas podem ter influência no design, construção e utilização de proteção de infraestrutura de informação, mas a relevância e o grau de efeito têm

ainda não foi testado empiricamente.

Nesta seção, identificamos uma variedade de descobertas bem estudadas da ciência comportamental da psicologia, medicina comportamental e outras disciplinas onde as técnicas demonstraram afetar o comportamento relacionados à cognição e ao viés. Também descrevemos diversas heurísticas e modelos relacionados à saúde que potencial para melhorar a segurança cibernética. No entanto, ao contrário das conclusões da Seção 4, estas conclusões não foram avaliados especificamente em termos de mudança de comportamento relacionado à segurança cibernética. Nesta seção, apresentar cada descoberta da ciência comportamental, discutir uma amostra de resultados de pesquisa e descrever os possíveis implicações para a segurança cibernética.

6.1 Cognição

Cognição refere-se à maneira como as pessoas processam e aprendem informações. Existem várias descobertas de pesquisas sobre cognição humana que podem ser relevantes para a segurança cibernética.

Efeito de Vítima Identificável

O efeito da vítima identificável refere-se à tendência dos indivíduos de oferecerem maior ajuda quando uma situação específica, pessoa identificável (a vítima) é observada em dificuldades, quando comparada a uma grande e vagamente definida grupo com a mesma necessidade. Por exemplo, muitas pessoas estão mais dispostas a ajudar uma pessoa sem-teto que vive perto do escritório do que as centenas de moradores de rua que vivem na cidade.**(**Exemplo: K. Jenni e G. Loewenstein, “Explicando o 'Efeito da Vítima Identificável'”,*Revista de Risco e Incerteza,*14, 1997, págs. 235-257.)**Implicações:**Os usuários podem escolher uma segurança mais forte quando possíveis resultados negativos forem tangível e pessoal, em vez de abstrato.

Modelo de Probabilidade de Elaboração

O Modelo de Probabilidade de Elaboração descreve como as atitudes são formadas e persistem. Baseia-se na noção de que existem duas vias principais para a mudança de atitude: a via central e a via periférica. Central os processos são lógicos, conscientes e exigem muita reflexão. Portanto, os processos de rota central para

a tomada de decisões só é utilizada quando as pessoas estão motivadas e aptas a prestar atenção. O resultado da centralização o processamento de rotas é frequentemente uma mudança permanente de atitude, à medida que as pessoas adotam e elaboram os argumentos sendo feito por outros. Em contraste, quando as pessoas seguem o caminho periférico, elas não prestam atenção a argumentos persuasivos; em vez disso, são influenciados por características superficiais, como a popularidade do falante. Neste caso, a mudança de atitude tende a ser apenas temporária. A pesquisa se concentrou em como fazer com que as pessoas usem a rota central em vez da rota periférica. (Exemplo: RE Petty e JT Cacioppo,*Atitudes e Persuasão: Abordagens Clássicas e Contemporâneas*. Dubuque, IA: WC Brown, 1981. RE Petty e JT Cacioppo,*Comunicação e Persuasão: Central e Periférica Rotas para a mudança de atitude,*Nova York: Springer-Verlag, 1986.)**Implicações**:Uma das melhores maneiras de motivar os usuários a seguir a rota central ao receber uma mensagem de segurança cibernética é tornar a mensagem pessoalmente relevante. O medo também pode ser eficaz para fazer com que os usuários prestem atenção, mas apenas se os níveis de medo forem moderado e uma solução para a situação que induz o medo também é oferecida; o medo forte leva à luta ou fuga (reações físicas). A rota central leva à consideração de argumentos a favor e contra, e a final a escolha é cuidadosamente considerada. Essa distinção pode ser particularmente importante na conscientização sobre segurança treinamento.

Dissonância Cognitiva

A dissonância cognitiva é a sensação de desconforto que surge ao manter dois pensamentos conflitantes ao mesmo tempo. mente ao mesmo tempo. Uma pessoa muitas vezes sente uma forte dissonância quando acredita em algo sobre si mesma (por exemplo, “Eu sou uma boa pessoa”) e então faz algo contrário a isso (por exemplo, “Eu fiz algo ruim”). O O desconforto muitas vezes parece uma tensão entre dois pensamentos opostos. A dissonância cognitiva é uma motivador poderoso que pode levar as pessoas a mudar de uma das três maneiras: mudar o comportamento, justificar o comportamento mudando a atitude conflitante ou justificando o comportamento adicionando novas atitudes. A dissonância é mais poderoso quando se trata de autoimagem (por exemplo, sentimentos de tolice, imoralidade, etc.). (Exemplos: L. Festinger,*Uma Teoria da Dissonância Cognitiva*, Stanford, CA: Stanford University Press, 1957; L. Festinger e JM Carlsmith, “Consequências cognitivas da conformidade forçada”,*Revista de Anormalidade e Social*

Psicologia, 58, 1959, págs. 203-211.)**Implicações:**A dissonância cognitiva é central para muitas formas de persuasão para mudar crenças, valores, atitudes e comportamentos. Para fazer com que os usuários mudem seu comportamento cibernético, podemos primeiro mudar suas atitudes em relação à segurança cibernética. Por exemplo, um sistema pode enfatizar a importância de um usuário sensação de insensatez em relação aos riscos cibernéticos que está assumindo, permitindo que uma tensão dissonante seja injetada repentinamente ou que se acumule ao longo do tempo. Então, o sistema pode oferecer ao usuário maneiras de aliviar a tensão mudando seu comportamento.

Teoria Social Cognitiva

A Teoria Cognitiva Social é uma teoria sobre a aprendizagem baseada em duas noções-chave: (1) as pessoas aprendem por observar o que os outros fazem e (2) os processos de pensamento humano são essenciais para a compreensão da personalidade. a teoria afirma que parte da aquisição de conhecimento de um indivíduo pode estar diretamente relacionada à observação outros dentro do contexto de interações sociais, experiências e influências da mídia externa. (Exemplos: A. Bandura, "Aplicação Organizacional da Teoria Social Cognitiva",*Revista Australiana de Gestão*, 13(2), 1988, pp. 275-302; A. Bandura, "Agência Humana na Teoria Cognitiva Social",*americano Psicólogo*, 44, 1989, págs. 1175-1184.)**Implicações:**Levando em consideração gênero, idade e etnia, uma campanha de conscientização cibernética poderia reduzir o risco cibernético usando a teoria cognitiva social para permitir que os usuários identificar-se com um colega reconhecível e ter um maior senso de autoeficácia. Os usuários provavelmente imitar as ações dos colegas para aprender um comportamento apropriado e seguro.

Efeito espectador

O efeito espectador é um fenômeno psicológico em que é menos provável que alguém intervenha em uma situação situação de emergência quando outras pessoas estão presentes e podem ajudar do que quando ele ou ela está sozinho. (Exemplo**:**JM Darley e B. Latané, "Intervenção de espectadores em emergências: difusão de Responsabilidade,"*Revista de Personalidade e Psicologia Social,*8, 1968, págs. 377-383.)**Implicações:** Durante um evento cibernético, os usuários podem não se sentir obrigados a aumentar a consciência situacional ou tomar as medidas necessárias medidas de segurança porque esperam que outros ao seu redor o façam. Assim, os sistemas podem ser projetados

com mecanismos para combater esse efeito, incentivando os usuários a agir quando necessário.

6.2 Viés

O preconceito descreve a tendência de uma pessoa de ver algo de uma perspectiva específica. Essa perspectiva impede a pessoa de ser objetiva e imparcial. As seguintes descobertas sobre preconceito podem ser úteis na concepção, construção e utilização de infraestrutura de informação.

Viés do status quo

O viés do status quo descreve a tendência das pessoas de não mudar um comportamento estabelecido sem uma incentivo convincente para fazê-lo. (Exemplo: W. Samuelson e R. Zeckhauser, “Status Quo Bias in Decision Fazendo,"*Revista de Risco e Incerteza,*1,1988, págs. 7-59.)**Implicações:**Os usuários precisarão de recursos atraentes incentivos para mudar seu comportamento de segurança cibernética estabelecido. Por exemplo, infraestrutura de informação pode ser projetado para fornecer incentivos para que as pessoas suspeitem de documentos enviados de fontes desconhecidas. Da mesma forma, a infraestrutura pode fornecer aos designers, desenvolvedores e usuários feedback sobre seus reputações (por exemplo, “Sessenta e três por cento dos seus anexos nunca são abertos pelo destinatário.”) ou repercussões de suas ações (por exemplo, “Foi o defeito do seu projeto que permitiu essa violação”) para reduzir o status preconceito quo.

Efeitos de enquadramento

Os cientistas geralmente esperam que as pessoas façam escolhas racionais com base nas informações disponíveis para elas. A teoria da utilidade esperada baseia-se na noção de que as pessoas escolhem opções que proporcionam o maior benefício (ou seja, a mais útil para eles) com base nas informações disponíveis para eles. No entanto, há uma crescente literatura que fornece evidências de que quando as pessoas devem escolher entre alternativas que envolvem risco, onde o as probabilidades dos resultados são conhecidas, eles se comportam de forma contrária às previsões da teoria da utilidade esperada. Esta área de estudo, chamada teoria da perspectiva, é descritiva e não preditiva; os teóricos da perspectiva relatam sobre como as pessoas realmente fazem escolhas quando confrontadas com informações sobre cada alternativa.

Uma das primeiras descobertas na teoria da perspectiva (Tversky e Kahneman, 1981) demonstrou que a

O enquadramento de uma mensagem pode afetar a tomada de decisões. O enquadramento refere-se ao contexto em que alguém

interpreta informações, reage a eventos e toma decisões. Por exemplo, a eficácia de um medicamento pode ser

enquadrado em termos de número de vidas salvas ou número de vidas perdidas; estudos demonstraram que dados equivalentes

enquadrados de maneiras opostas (ganho vs. perda) levam a decisões dramaticamente diferentes sobre se e como

usam a mesma droga. O contexto ou enquadramento de um problema pode ser alcançado pela manipulação do

opções de decisão ou referindo-se às qualidades dos tomadores de decisão, como suas normas, hábitos e

temperamento. (Exemplos: D. Kahneman e A. Tversky, "Teoria da Prospecção: Uma Análise de Decisões

"Em risco",*Econométrica*, 47, 1979, pp. 313-327; A. Tversky e D. Kahneman, "A estruturação de

Decisões e a Psicologia da Escolha",*Ciência,*211,1981, págs. 453-458.)**Implicações:**Escolhas do usuário

sobre segurança cibernética podem ser influenciados por enquadrá-los como ganhos em vez de perdas, ou por apelar a

características específicas do usuário. Possíveis aplicações incluem a classificação de dados anômalos de uma intrusão

log do sistema de detecção, apresentando a interface para um firewall como admitindo tráfego (bom) em vez de bloqueando

tráfego (ruim), ou descrever uma atividade de mineração de dados como expondo comportamento malicioso.

Viés de Otimismo

Dadas as chances minúsculas de ganhar na loteria, é incrível que as pessoas comprem bilhetes de loteria. Muitas

as pessoas acreditam que se sairão melhor do que a maioria dos outros envolvidos na mesma atividade, então compram ingressos

apesar das evidências em contrário. Este viés de otimismo se manifesta de muitas maneiras, como a superestimação

a probabilidade de eventos positivos e subestimar a probabilidade de eventos negativos. (Exemplos: ND

Weinstein, "Otimismo irrealista sobre eventos futuros da vida",*Revista de Personalidade e Social*

*Psicologia*39(5), novembro de 1980, pp. 806–820; D. Dunning, C. Heath e JM Suls, "Autodeterminação falhada

Avaliação: Implicações para a Saúde, Educação e Local de Trabalho",*Ciência Psicológica no Público*

*Interesse*5(3), 2004, págs. 69–106.)**Implicações:**Por subestimarem o risco, os usuários podem pensar que

são imunes a ataques cibernéticos, mesmo quando outros se mostraram suscetíveis. Por exemplo,

o viés do otimismo pode permitir o spear phishing (mensagens que parecem vir de uma fonte confiável, tentando obter

acesso não autorizado a dados de uma organização específica). O viés do otimismo também pode induzir as pessoas a ignorar medidas preventivas de cuidado, como curativos, porque acreditam que é improvável que sejam afetados. contrariando o viés do otimismo, os sistemas podem ser projetados para transmitir o impacto do risco e a probabilidade de maneiras que se relacionem às experiências reais das pessoas.

Viés de controle

O viés de controle refere-se à tendência das pessoas de acreditarem que podem controlar ou influenciar os resultados que desejam. claramente não pode; este fenômeno é às vezes chamado de ilusão de controle. (Exemplo: EJ Langer, "The Ilusão de Controle",*Revista de Personalidade e Psicologia Social*32(2), 1975, págs. 311-328.)

Implicações: Os usuários podem estar menos propensos a usar medidas de proteção (como verificação de vírus, limpeza de cache, verificar sites seguros antes de inserir informações de cartão de crédito ou prestar atenção ao spear phishing) quando sentem que têm controle sobre os riscos de segurança.

Viés de Confirmação

Quando alguém toma uma posição sobre um assunto, é mais provável que perceba ou dê crédito às evidências de que apoia essa posição do que evidências que a desacreditam. Esse viés de confirmação (ou seja, a busca por evidências para confirmar uma posição) resulta em situações em que as pessoas não estão tão abertas a novas ideias quanto pensam que estão são. Muitas vezes reforçam as suas atitudes existentes através da recolha selectiva de novas provas, interpretando evidências de forma tendenciosa ou recuperando seletivamente informações da memória. Por exemplo, um analista encontrar um padrão percebido em uma série de falhas tenderá a deixar de procurar outras explicações e em vez disso, busque evidências que confirmem sua hipótese. (Exemplo: M. Lewicka, "Confirmation Bias: Erro cognitivo ou estratégia adaptativa de controle de ação?" em M. Kofta, G. Weary e G. Sedek,*Pessoal Controle em Ação: Mecanismos Cognitivos e Motivacionais*. Nova Iorque: Springer. 1998, pp. 233–255.)

Implicações: Os usuários podem ter impressões iniciais sobre o quão protegidas (ou não) as informações a infraestrutura que eles estão usando. Para superar seu viés de confirmação, o sistema deve fornecer aos usuários com um arsenal de evidências para encorajá-los a mudar suas crenças atuais ou a mitigar seus excessos

confiança.

Efeito Dotação

O efeito dotação descreve o fato de que as pessoas geralmente atribuem um valor maior aos objetos que possuem do que objetos que não possuem. Um efeito relacionado é que as pessoas reagem mais fortemente à perda do que ao ganho; isto é, eles tomarão medidas mais fortes para evitar perder algo do que para ganhar algo. (Exemplo: R. Thaler, "Rumo a uma teoria positiva da escolha do consumidor"*Revista de Comportamento Econômico e Organização*, 1, 1980, págs. 39-60.)**Implicações:**Os usuários podem pagar mais (tanto figurativa quanto literalmente) pela segurança quando se trata de permite que eles mantenham algo que já possuem, em vez de ganhar algo novo. Este efeito, juntamente com uma efeito de enquadramento, pode ter impacto particular na privacidade. Quando uma ação é expressa como uma perda de privacidade (em vez de um ganho de capacidade), as pessoas podem reagir negativamente.

6.3 Heurística

Em psicologia, uma heurística é uma regra simples inerente à natureza humana ou aprendida para reduzir a deficiência cognitiva. carga. Portanto, os consideramos atraentes para abordar as questões de carga cognitiva descritas anteriormente.
as regras da heurística são usadas para explicar como as pessoas fazem julgamentos, decidem questões e resolvem problemas;
As heurísticas são particularmente úteis para explicar como as pessoas lidam com problemas complexos ou incompletos informação. Quando as heurísticas falham, elas podem levar a erros sistemáticos ou vieses cognitivos.

Heurística de Afeto

A heurística do afeto permite que alguém tome uma decisão com base em um afeto (ou seja, um sentimento) em vez de deliberação racional. Se alguém tem um bom pressentimento sobre uma situação, pode perceber que ela apresenta baixo risco; da mesma forma, um mau pressentimento pode levar a uma maior percepção de risco. (Exemplo: M. Finucane, E. Peters e DG MacGregor, "A Heurística do Afeto", em T. Gilovich, D. Griffin e D. Kahneman,*Heurística e Vieses: A Psicologia do Julgamento Intuitivo*. Cambridge University Press, 2002, pp. 397–420.)**Implicações:**Se os usuários percebem pouco risco, o sistema pode precisar de um projeto que crie um efeito mais crítico em relação ao computador

segurança que os incentivará a tomar medidas de proteção. O sistema também deve recompensar o sistema administrador que analisa atentamente um registro de auditoria do sistema porque algo simplesmente não "parece" certo.

Heurística de disponibilidade

A heurística da disponibilidade refere-se à relação entre facilidade de recordação e probabilidade. Em outras palavras, devido à heurística de disponibilidade, alguém irá prever a probabilidade ou frequência de um evento em uma população com base na facilidade com que os eventos vêm à mente. Quanto mais recente, emocional, ou vívido for um evento, maior a probabilidade de ele vir à mente. (Exemplo: A. Tversky e D. Kahneman, "Disponibilidade: Uma Heurística para Julgar Frequência e Probabilidade"*Psicologia Cognitiva*5, 1973, págs.207-232.)**Implicações:**Os usuários serão mais persuadidos a agir de forma responsável se o sistema for projetado para usar eventos pessoais vívidos como exemplos, em vez de estatísticas e fatos. Além disso, se o sistema relatar eventos recentes eventos cibernéticos, pode ser mais eficaz para encorajar os utilizadores a tomarem medidas para prevenir futuros eventos adversos eventos. As escolhas dos usuários também podem ser fortemente influenciadas pela primeira coisa que lhes vem à mente. Portanto, Exercícios frequentes de segurança podem encorajar um comportamento de segurança mais desejável. Por outro lado, um sistema que já passou algum tempo sem um grande incidente cibernético pode induzir os administradores a uma falsa sensação de segurança devido à baixa frequência de eventos. Os administradores podem então tornar-se negligentes na aplicação atualizações de segurança devido ao longo período de operação sem incidentes.

6.4 Modelos Comportamentais Relacionados à Saúde

Na segurança cibernética, enquadramos muitas questões usando metáforas relacionadas com a saúde porque elas são, de muitas maneiras, análogo. Por exemplo, falamos de vírus e infecções ao descrever ataques. Da mesma forma, discutir o aumento da imunidade a intrusões ou o aumento da resiliência após um ataque bem-sucedido. Para isso razão pela qual acreditamos que as estratégias de design de segurança podem alavancar a pesquisa significativa em questões relacionadas à saúde modelos comportamentais. Discutimos vários modelos candidatos aqui.

Modelo de Crenças em Saúde

O Modelo de Crenças em Saúde, desenvolvido na década de 1950 após o fracasso de um programa gratuito de triagem de tuberculose programa, ajudou o Serviço de Saúde Pública dos EUA ao tentar explicar e prever comportamentos de saúde. focado em atitudes e crenças. Seis constructos descrevem as crenças fundamentais de um indivíduo com base em suas percepções de: suscetibilidade, gravidade, benefícios, barreiras, pistas para ação e autoeficácia na execução de uma determinado comportamento de saúde. Os benefícios percebidos devem superar as barreiras ou custos. (Exemplo: I. Rosenstock, "Origens históricas do modelo de crença em saúde",*Monografias de Educação em Saúde,*2(4), 1974.)

Implicações: Os modelos de educação em saúde e segurança são semelhantes. Se o Modelo de Crenças em Saúde se traduz para a conscientização sobre segurança cibernética, um usuário tomará medidas de segurança de proteção se sentir que algo negativo condição pode ser evitada (por exemplo, vírus de computador podem ser evitados), tem uma expectativa positiva de que, ao tomar uma ação recomendada evitará uma condição negativa (por exemplo, fazer uma verificação de vírus impedirá uma infecção viral infecção) e acredita que pode executar com sucesso a ação recomendada (por exemplo, está confiante de que ele sabe como instalar arquivos de proteção contra vírus). O modelo sugere sucesso apenas se os benefícios (por exemplo, manter a si mesmo, sua organização e a nação seguros) superam os custos (por exemplo, tempo de download, perda de trabalhar).

Modelo de Processo Paralelo Estendido

O Modelo de Processo Paralelo Estendido (EPPM) é uma extensão do Modelo de Crenças em Saúde que tenta melhorar a eficácia da mensagem usando ameaças. Com base na estrutura de controle de perigo/controle de medo de Leventhal, O EPPM, que tem múltiplos componentes, explica por que muitos apelos ao medo falham e incorpora o medo como um elemento-chave variável e descreve a relação entre medo e eficácia. Leventhal define o controle do perigo processo como um indivíduo que busca reduzir o risco apresentado, tomando medidas diretas e fazendo adaptações mudanças, mas o processo de controle do medo se concentra em mudanças desadaptativas na percepção, suscetibilidade e gravidade do risco. O EPPM fornece orientações sobre como construir mensagens eficazes de apelo ao medo: Enquanto as percepções de eficácia forem mais fortes do que as percepções de ameaça, o usuário entrará em controle de perigo

modo (aceitar a mensagem e tomar as medidas recomendadas para evitar que o perigo aconteça). (Exemplos: K. Witte, “Colocando o medo de volta nos apelos ao medo: o processo paralelo estendido Modelo,"*Monografias de Comunicação*, 59, 1992, pp. 329-349; H. Leventhal, “Descobertas e Teoria no Estudo das Comunicações de Medo”, em L. Berkowitz, ed.,*Avanços em Psicologia Social Experimental*, Vol. 5, Nova Iorque: Academic Press, 1970, pp. 119-186.)**Implicações**:Quando usadas adequadamente, ameaças e O medo pode ser útil para incentivar os usuários a cumprir as normas de segurança. No entanto, as mensagens não podem ser muito forte, e os usuários devem acreditar que são capazes de cumprir com sucesso as recomendações de segurança. Isso O modelo pode explicar como incentivar os usuários a aplicar patches de segurança e desempenho, usar e manter ferramentas antivírus e evite comportamentos arriscados online.

Representações de Doenças

A comunidade de saúde tem muita experiência na representação da natureza e da gravidade de doença aos pacientes, para que possam tomar decisões informadas sobre as opções de tratamento e saúde. em particular, há lições a serem aprendidas com a forma como as mensagens de medo são usadas em situações relativamente agudas. situações para encorajar as pessoas a tomarem medidas de promoção da saúde, como usar cinto de segurança ou desistir fumar. Pesquisadores de saúde (Leventhal, Meyer e Nerenz, 1980) descobriram que diferentes tipos de informações são necessárias para influenciar atitudes e reações a uma ameaça percebida à saúde e ao bem-estar sendo, e que as mudanças de comportamento duram apenas por curtos períodos de tempo. Ao estender seu modelo inicial, os pesquisadores buscaram adaptações e esforços de enfrentamento para os pacientes que vivenciam doenças crônicas. as representações da doença resultante integram os mecanismos de enfrentamento com os esquemas existentes (ou seja, diretrizes normativas que as pessoas seguem), permitindo que os pacientes compreendam seus sintomas e orientando quaisquer ações de enfrentamento. As representações da doença têm cinco componentes: identidade, linha do tempo, consequências, controle/cura e coerência da doença. (Exemplos: H. Leventhal, D. Meyer e DR Nerenz, “The Common “Representação sensorial do perigo da doença”, em S. Rachman, ed.,*Contribuições para a Psicologia Médica*, Novo York: Pergamon Press, 1980, pp. 17–30; H. Leventhal, I. Brissette e EA Leventhal, “The Common- “Modelo de autorregulação da saúde e da doença”, em LD Cameron e H. Leventhal, eds.,*O Eu-*

Regulação do comportamento de saúde e doença, Londres: Routledge, 2003, pp. 42–65.)**Implicações**:Em um sistema bem projetado, os usuários preocupados em confiar em um site, pessoa ou documento podem obter novos informações sobre sua postura de segurança e avaliar suas tentativas de lidar (por exemplo, moderar, curar ou lidar) com seus efeitos. Então, os usuários formam novas representações com base em suas experiências. Essas as representações tendem a ser cumulativas, com informações de segurança sendo adotadas, descartadas ou adaptadas conforme necessário. Assim, as representações provavelmente estarão vinculadas à seleção de procedimentos de enfrentamento, ações planos e resultados. Esses resultados podem ser importantes para o desenvolvimento de estratégias de resposta a incidentes.

Teoria da Ação Raciocinada/Teoria do Comportamento Planejado

A Teoria da Ação Raciocinada e a Teoria do Comportamento Planejado baseiam-se em duas noções: (1) pessoas são razoáveis e fazem bom uso das informações ao decidir entre comportamentos e (2) as pessoas consideram as implicações de seu comportamento. O comportamento é direcionado a objetivos ou resultados, e as pessoas escolher livremente os comportamentos que os levarão em direção a esses objetivos. Eles também podem escolher não agir se Eles acreditam que agir os afastará de seus objetivos. As teorias levam em consideração quatro conceitos: intenção comportamental, atitude, normas sociais e controle comportamental percebido. A intenção de se comportar tem um influência direta no comportamento real em função da atitude e das normas subjetivas. A atitude é uma função tanto das consequências pessoais esperadas do comportamento quanto do valor afetivo atribuído a essas consequências. (Exemplo: I. Ajzen, "Das Intenções às Ações: Uma Teoria do Comportamento Planejado", em J. Kuhl e J. Beckmann, orgs.,*Controle de Ação: Da Cognição ao Comportamento*. Berlim, Heidelberg, Nova Iorque: Springer-Verlag, 1985.)**Implicações**: Para incentivar os usuários a mudarem seu comportamento de segurança, o o sistema deve criar mensagens que afetem as intenções dos usuários; por sua vez, as intenções são alteradas por influenciando as atitudes dos usuários por meio da identificação de normas sociais e controle comportamental. Os usuários devem percebem que podem controlar a conclusão bem-sucedida de suas tarefas de forma segura e protegida.

Modelo de Estágios de Mudança

O Modelo de Estágios de Mudança avalia a prontidão de uma pessoa para iniciar um novo comportamento, fornecendo estratégias

ou processos de mudança para guiá-la através dos estágios de mudança para ação e manutenção. A mudança é uma processo envolvendo progressão através de seis estágios: pré-contemplação, contemplação (pensamentos), preparação (pensamentos e ações), ação (mudança real de comportamento), manutenção e término. Portanto, as intervenções para mudar comportamentos devem corresponder e afetar o estágio apropriado. Para progredir Nos estágios iniciais, as pessoas aplicam processos cognitivos, afetivos e avaliativos. À medida que as pessoas se movem para manutenção ou término, eles dependem mais de compromissos e condicionamentos, (Exemplos: JO Prochaska, JC Norcross e CC DiClemente,*Mudando para sempre: o programa revolucionário que Explica os seis estágios da mudança e ensina como se libertar de maus hábitos*. Nova Iorque: W. Morrow; 1994; JO Prochaska e CC DiClemente, "A Abordagem Transteórica", em JC Norcross e MR Goldfried, orgs.*Manual de Integração em Psicoterapia*, 2ed., Nova Iorque: Oxford Imprensa Universitária, 2005. pp. 147-171.)**Implicações**:Para mudar comportamentos relacionados à segurança, é necessário primeiro avaliar o estágio dos usuários antes de desenvolver processos para provocar mudanças de comportamento. Por exemplo, obter desenvolvedores de software para implementar a segurança no ciclo de vida de desenvolvimento do código e, especialmente, ao longo o ciclo de vida é notoriamente difícil. Atualmente, muitos esforços são direcionados para mover os desenvolvedores diretamente para estágio quatro (ação), sem a devida atenção à importância dos estágios anteriores.

Modelo de Processo de Precaução-Adoção

Teorias que tentam explicar o comportamento examinando os custos e benefícios percebidos da mudança de comportamento só funciona se a pessoa tiver conhecimento ou experiência suficiente para ter formado uma crença. A Precaução-O Modelo de Processo de Adoção busca compreender e explicar o comportamento observando sete estágios: inconsciente; não envolvido; decidindo sobre agir; decidiu não agir; decidiu agir; agindo; e Manutenção. As pessoas devem responder melhor a intervenções adequadas ao estágio em que se encontram. (Exemplos: ND Weinstein, "O Processo de Adoção por Precaução",*Psicologia da Saúde*, 7(4), 1988, págs. 355-386; ND Weinstein e PM Sandman, "Um modelo do processo de adoção de precaução: evidências de Teste de radônio em casa"*Psicologia da Saúde*, 11(3), 1992, págs. 170-180.)**Implicações**: Ações de segurança pode estar relacionado aos sete estágios. Pode ser necessário avaliar o estágio de um usuário antes de desenvolver um

processo para provocar a mudança de comportamento desejada.

7 Aplicando as descobertas da ciência comportamental: o caminho a seguir

Apresentamos alguns resultados iniciais que mostram por que esta abordagem multidisciplinar provavelmente produzirá insights úteis. Nesta seção final, descrevemos as próximas etapas para determinar as melhores maneiras de combinar ciência comportamental com ciência da computação para gerar segurança cibernética aprimorada. As etapas recomendadas envolver o incentivo a workshops multidisciplinares, a realização de estudos empíricos em todas as disciplinas e construindo um repositório acessível de descobertas multidisciplinares.

7.1 Workshops que unem comunidades

O trabalho multidisciplinar pode ser desafiador por vários motivos. Primeiro, como observado pelos participantes de um evento nacional Workshop da Academia de Ciências (2010), existem terminologias e definições inconsistentes em disciplinas. Particularmente para palavras como “confiança” ou “risco”, duas disciplinas diferentes podem usar a mesma palavra mas com significados e pressupostos muito diferentes. Em segundo lugar, existem poucos incentivos para publicar descobertas entre disciplinas, muitos pesquisadores trabalham em áreas distintas e separadas que normalmente não compartilham informação. Por esta razão, recomendamos a criação de workshops que conectem as comunidades para que que o conhecimento de cada comunidade pode beneficiar as outras.

Em Julho de 2010, o Instituto para a Protecção da Infra-estrutura da Informação (I3P) realizou um workshop de dois dias para reunir membros da comunidade de ciências comportamentais e da comunidade de segurança cibernética, examinar como colocar em prática as descobertas avaliadas com sucesso e estabelecer grupos de pesquisadores dispostos a avaliar empiricamente descobertas promissoras e avaliar sua aplicabilidade à segurança cibernética. O workshop criou uma oportunidade para a formação de grupos de pesquisadores e profissionais ansiosos por avaliar e adotar formas mais eficazes de integrar a ciência comportamental com a segurança cibernética. Ou seja, o workshop é o primeiro passo no que esperamos ser uma parceria contínua entre a ciência da computação e a ciência comportamental

ciência que melhorará a eficácia da segurança cibernética.

O resultado do workshop incluiu:

- Identificação de descobertas existentes que podem melhorar a segurança cibernética no curto prazo.
- Identificação de potenciais descobertas da ciência comportamental que poderiam ser aplicadas, mas necessitam avaliações empíricas de seus efeitos na segurança cibernética.
- Identificação de áreas e problemas de segurança cibernética onde a aplicação de conceitos de a ciência comportamental pode ter um impacto positivo.
- Estabelecimento de um repositório inicial de informações sobre ciência comportamental e segurança cibernética.

Como resultado deste workshop, vários estudos de spear phishing foram conduzidos em universidades e indústrias configurações e um estudo de incentivos, para demonstrar empiricamente que tipos de incentivos (ou seja, dinheiro, vagas de estacionamento convenientes, reconhecimento público, etc.) motivariam mais os usuários a ter uma boa experiência cibernética higiene, foi projetado para administração futura. Um segundo workshop foi realizado em outubro de 2011 para relatar sobre as descobertas dos estudos e organizar estudos futuros.

Workshops deste tipo não só podem funcionar como catalisadores para o início de novas pesquisas, mas também podem incentivar a interação e a cooperação contínuas entre disciplinas. Esforços semelhantes estão sendo incentivados em diversas áreas da segurança cibernética, particularmente na segurança utilizável (Pfleeger, 2011).

7.2 Avaliação empírica entre disciplinas

Esperamos expandir o corpo de conhecimento sobre as interações entre o comportamento humano e o ciberespaço. segurança por meio de investigações que produzirão tanto projetos experimentais inovadores quanto dados que podem formar a base da replicação experimental e da adaptação de aplicações a situações particulares. No entanto, existem são desafios para realizar este tipo de pesquisa, especialmente quando os recursos são limitados. Para Por exemplo, normalmente não é possível construir o mesmo sistema duas vezes (um como controle, um como tratamento) e

comparar os resultados, portanto, um bom planejamento experimental é crucial para produzir resultados sólidos e confiáveis com níveis suficientes de validade externa.

A avaliação empírica dos efeitos da mudança na segurança cibernética envolve muitas coisas, incluindo identificar variáveis, controlar os efeitos de viés e interação e determinar o grau em que os resultados podem ser generalizados. Estes são princípios fundamentais do método empírico, mas muitas vezes não são compreendidos ou não aplicados adequadamente. Esperamos produzir diretrizes mais abrangentes para projeto experimental, com o objetivo de auxiliar profissionais de segurança cibernética e cientistas comportamentais na concepção avaliações que produzirão os resultados mais significativos. Estas diretrizes destacarão diversas questões:

- A necessidade de projetar um estudo de modo que as variáveis de confusão e o viés sejam reduzidos ao máximo possível.
- A necessidade de declarar a hipótese experimental e identificar variáveis dependentes e independentes.
- A necessidade de identificar os participantes da pesquisa e determinar qual população está sob escrutínio.
- A necessidade de procedimentos de amostragem claros e completos, para que a amostra represente o objeto identificado população.
- A necessidade de descrever as condições experimentais com detalhes suficientes para que o leitor possa compreender o estudo e também replicá-lo.
- A necessidade de realizar um debriefing pós-experimento eficaz, especialmente para estudos onde o real a intenção do estudo não é revelada até que o estudo seja concluído.

Existem vários exemplos de bons projetos experimentais para estudos na intersecção de comportamento ciência e segurança cibernética. Por exemplo, muitas lições foram aprendidas em um experimento focado em informações privilegiadas ameaça (Caputo, Maloof e Stephens, 2009). Neste estudo, os pesquisadores encontraram vários desafios na seleção da melhor amostra e seguindo procedimentos empíricos rigorosos. Eles documentaram a importância de testar o projeto experimental antes de envolver os participantes alvo. Em particular, foi

difícil fazer com que os participantes corporativos realizem as tarefas experimentais com a mesma motivação que os os usuários médios têm ao fazer seus trabalhos regulares. Portanto, os pesquisadores usaram testes piloto para determinar o que motivaria os participantes. Em seguida, a motivação foi incorporada ao desenho do estudo. Embora este estudo tenha utilizado funcionários corporativos, redes reais e tarefas plausíveis para realizar a pesquisa ambiente o mais realista possível, gerando conjuntos de dados em qualquer situação controlada reduziu a capacidade dos pesquisadores de generalizar as descobertas para situações complexas.

Há muitos estudos que podem se beneficiar de uma melhor coleta de dados e de um melhor desenho de estudo. Pfleeger et al. (2006) sugerem um roteiro para melhorar a coleta e análise de dados de informações de segurança cibernética. Além disso, Cook e Pfleeger (2010) descrevem como construir melhorias em conjuntos de dados existentes e descobertas.

7.3 Repositório de Resultados

Estamos a construir um repositório de descobertas relevantes, incluindo conjuntos de dados onde disponíveis, para servir pelo menos dois propósitos. Primeiro, fornecerá a base para a tomada de decisões sobre quando e como incluir comportamento considerações na especificação, projeto, construção e uso de produtos e processos de segurança cibernética. Em segundo lugar, permitirá aos investigadores e aos profissionais replicar estudos nos seus próprios contextos, para confirmar ou refutar descobertas anteriores e adaptar os métodos às necessidades e restrições específicas. Essas informações estabelecerão a base para a segurança cibernética baseada em evidências.

Este artigo relata as descobertas de nossa incursão inicial na combinação de ciência comportamental e cibernética segurança. Nos últimos anos, tem-se falado muito em convidar ambas as disciplinas a colaborar, mas pouco se tem feito trabalho foi feito para abrir amplamente a discussão para ambas as comunidades. Nossos workshops tiveram uma abordagem ousada e ampla etapas, e espera-se que as atividades aqui relatadas, construídas sobre os ombros do trabalho realizado em ambos comunidades nas últimas duas décadas, encorajará outros a se juntarem a nós para pensar de forma mais expansiva sobre problemas de segurança cibernética e possíveis soluções. Em particular, encorajamos outros envolvidos em pesquisas em todas as disciplinas para nos contatar, para que possamos estabelecer vínculos virtuais e reais que nos movam

para a compreensão e implementação de uma segurança cibernética melhorada.

8 Referências

Amos, Deborah, “Desafio: Triagem Aeroportuária Sem Discriminação”, Edição Matinal, National Rádio Pública, 14 de janeiro de 2010, disponível em http://www.npr.org/templates/story/story.php?storyId=122556071

Baier, Annette, “Confiança e Antitruste”, *Ética,* Vol. 96, No. 2, 1986, pp. 231-260.

Baker, Peter e Carl Hulse, “Os EUA tiveram sinais precoces de conspiração terrorista, diz Obama”, *New York Times*, 30 Dezembro de 2009, página 1.

Bell, David E. e Leonard J. La Padula, “Sistemas de Computação Seguros: Fundamentos Matemáticos”, Relatório técnico MITRE MTR-2547, The MITRE Corporation, Bedford, MA, 1973.

Biba, Kenneth J., “Considerações de integridade para sistemas de computadores seguros”, Relatório técnico do MITRE MTR-3153, The MITRE Corporation, Bedford, MA, abril de 1977.

Biddle, Robert, Sonia Chiasson, PC van Oorschot, “Senhas gráficas: aprendendo com o primeiro Geração”, Relatório Técnico 09-09, Escola de Ciência da Computação, Universidade Carleton, Ottawa, Canadá, 2009.

Bond, Michael, “Comentários sobre a autenticação GrIDSure”, 28 de março de 2008, disponível em http://www.cl.cam.ac.uk/~mkb23/research/GridsureComments.pdf

Brostoff, Sacha e M. Angela Sasse, “As faces de acesso são mais úteis do que as senhas? Um teste de campo investigação”, em S. McDonald et al. (Eds) “Pessoas e Computadores XIV - Usabilidade ou Então”, *Procedimentos do IHC 2000,* Sunderland, Reino Unido, Springer, 2000, pp. 405-424.

Brostoff, Sacha e M. Angela Sasse, “Dez strikes e você está fora: aumentando o número de logins

Tentativas podem melhorar a usabilidade da senha,*Anais do Workshop CHI 2003 sobre Humano-Computador Sistemas de Interação e Segurança*, Ft. Lauderdale, Flórida, 2003.

Burke, Cody, "A coleta de informações encontra a sobrecarga de informações",*Basex TechWatch*, 14 de janeiro de 2010, disponível emhttp://www.basexblog.com/2010/01/14/intelligence-gathering-meets-io/

Caputo, Deanna, Marcus Maloof e Gregory Stephens, "Detecção de roubo interno de segredos comerciais"*IEEE Segurança e Privacidade*7(6), novembro/dezembro de 2009, pp. 14-21.

Castelfranchi, Cristiano e Rino Falcone, "Princípios de Confiança para MAS: Anatomia Cognitiva, Importância e Quantificação,*Anais da Terceira Conferência Internacional sobre Multiagentes Sistemas,*1998.

Castelfranchi, Cristiano e Rino Falcone, "Confiança Social: Uma Abordagem Cognitiva", em Cristano Castelfranchi e Yao-Hua Tan, editores.,*Confiança e Engano em Sociedades Virtuais*, Kluwer Acadêmico Editora, Amsterdã, 2002.

Chase, WG e HA Simon, "Percepção no Xadrez",*Psicologia Cognitiva*4(1), 1973, págs. 55–81.

Cheshire, Coye e Karen Cook, "O surgimento de redes de confiança sob incerteza: implicações para Interações na Internet",*Analisar e Crítica*26, 2004, págs. 220-240.

Chiasson, Sonia, Alain Forget, Elizabeth Stobert, Paul C. van Oorschot e Robert Biddle, "Múltiplos interferência de senha em senhas de texto e senhas gráficas baseadas em cliques",*ACM Computador e Segurança das Comunicações*(CCS), novembro de 2009, pp. 500-511.

Clements, Paul e Linda Northrup,*Linhas de produtos de software: práticas e padrões*, Addison-Wesley, Leitura, MA, 2001.

Consumer's Union, "Vazamentos de identidade: uma fonte surpreendente é o seu governo em ação"*Relatórios do consumidor*, Setembro de 2008, disponível emhttp://www.consumerreports.org/cro/money/credit-loan/identity-roubo/vazamentos-de-identidade-governamental/visão-geral/vazamentos-de-identidade-governamental-ov.htm

Cook, Ian P. e Shari Lawrence Pfleeger, “Desafios de suporte à decisão de segurança na coleta e Usar,"*Segurança e Privacidade IEEE*, 8(3), maio-junho de 2010, pp. 28-35.

Datta, Pratim e Sutirtha Chatterjee, “A economia e a psicologia da confiança do consumidor em Intermediários em Mercados Eletrônicos: o Quadro EM-Trust,*Revista Europeia de Informação Sistemas*, 17(1), fevereiro de 2008, pp. 12-28.

Dhamija, Rachna e Adrian Perrig, “Déjà Vu: Um estudo de usuário usando imagens para autenticação”, *Anais do 9º Simpósio de Segurança USENIX*, Denver, CO, agosto de 2000.

Ellsberg, Daniel J., Risco,*Ambiguidade e Decisão*, Relatório RAND D-12995, RAND Corporation, Santa Mônica, CA, 1964.

Everitt, Katherine, Tanya Bragin, James Fogarty e Tadayoshi Kohno, “Um estudo abrangente de frequência, interferência e treinamento de múltiplas senhas gráficas”,*Conferência ACM sobre Direitos Humanos Fatores em Sistemas de Computação*(CHI), abril de 2009. Jhawar, Ravi, Philip Inglesant, Martina Angela Sasse e Nicolas Courtois, “Make Mine a Quadruple: Reforçando a Segurança do PIN Gráfico Único Autenticação,"*Anais da Quinta Conferência Internacional sobre Segurança de Redes e Sistemas*, 6 a 8 de setembro de 2011, Milão, Itália.

Klein, GA e R. Calderwood, “Modelos de decisão: algumas lições do campo”,*Transações IEEE sobre Sistemas, Homem e Cibernética*21(5), setembro/outubro de 1991, pp. 1018-1026.

Klein, Gary A.,*Fontes de Poder: Como as Pessoas Tomam Decisões*, MIT Press, Cambridge, MA, 1998.

Klein, Gary A. e Eduardo Salas, orgs.,*Ligando a experiência e a tomada de decisão naturalista,*Erlbaum, 2001.

Klein, Gary A.,*Luzes de rua e sombras: em busca das chaves para a tomada de decisão adaptativa*, MIT Imprensa, Cambridge, MA, 2009.

Lamandé, Emmanuelle, “GrIDSure autentica a mais recente plataforma de aplicativos remotos da Microsoft”,

Revista Global Security, 27 de abril de 2010, disponível emhttp://www.globalsecuritymag.com/GrIDsure-autentica-Microsoft-s,20100427,17307.html

Lerner, JS e LZ Tiedens, "Retrato do tomador de decisões irritado: como as tendências de avaliação moldam A influência da raiva na cognição"*Revista de Tomada de Decisão Comportamental*(Edição especial sobre emoção e Tomada de decisão), 19, 2006, pp. 115-137.

Libicki, Martin C. e Shari Lawrence Pfleeger, "Coletando os pontos: formulação e solução de problemas Elementos", RAND Occasional Paper OP-103-RC, RAND Corporation, Santa Monica, CA, 2004.

Mack, A. e I. Rock,*Cegueira por desatenção.*MIT Press, Cambridge, MA, 1998.

Mayo, Deborah e Rachelle Hollander, orgs.,*Evidências aceitáveis: ciência e valores em risco Gerenciamento*, Oxford University Press, 1991.

Miller, George A., "O número mágico sete mais ou menos dois: alguns limites em nossa capacidade de Informações do Processo",*Revisão Psicológica*63, 1956, págs. 81–97.

Academia Nacional de Ciências,*Rumo a uma melhor usabilidade, segurança e privacidade da tecnologia da informação: Relatório de um Workshop*, National Academies Press, Washington, DC, 2010.

Ofsted (Escritório do Reino Unido para Padrões em Educação, Serviços e Habilidades para Crianças), "O uso seguro de novos Technologies", Relatório OFSTED 090231, Manchester, Reino Unido, fevereiro de 2010.

Pfleeger, Shari Lawrence, "Draft Report on the NIST Workshop", março de 2011, disponível em http://www.thei3p.org/docs/publications/436.pdf

Pfleeger, Shari Lawrence, Joel Predd, Jeffrey Hunker e Carla Bulford, "Insiders se comportando mal: Abordando os maus atores e suas ações"*Transações IEEE sobre Forense e Segurança da Informação*, 5(2), março de 2010.

Pfleeger, Shari Lawrence, Rachel Rue, Jay Horwitz e Aruna Balakrishnan, Investindo em Segurança Cibernética: O Caminho para Boas Práticas,*Cutter IT Journal*, 19(1), janeiro de 2006, pp. 11-18.

Predd, Joel, Shari Lawrence Pfleeger, Jeffrey Hunker e Carla Bulford, “Insiders se comportando mal”, *Segurança e Privacidade IEEE*6(4), julho/agosto de 2008, pp. 66-70.

Riegelsberger, Jens, M. Angela Sasse e John D. McCarthy, “O dilema do pesquisador: avaliar Confiança na Comunicação Mediada por Computador”,*Revista Internacional de Estudos Humano-Computador,*Volume. 58, No. 6, 2003, pp. 759-781.

Riegelsberger, Jens, M. Angela Sasse e John D. McCarthy, “A Mecânica da Confiança: Uma Estrutura para Pesquisa e Design”,*Revista Internacional de Estudos Humano-Computador,*Vol. 62, No. 3, 2005, pp. 381-422.

Rock, I. e P. Engelstein, “Um estudo de memória para a forma visual”.*Revista Americana de Psicologia*, 72, 1959, págs. 221-229.

Sasse, M. Angela, “GrIDsure Usability Trials,” 2007, disponível em http://www.gridsure.com/uploads/UCL%20Report%20Summary%20.pdf

Sasse, M. Angela, Sacha Brostoff e Dirk Weirich, “Transformando o 'elo mais fraco: um ser humano-Abordagem de interação computacional para segurança utilizável e eficaz”, em R. Temple e J. Regnault. eds., *Segurança de Internet e Wireless*, IEE Press, Londres, 2002, pp. 243-258.

Sasse, M. Angela e Ivan Flechais, “Segurança Utilizável: Por que Precisamos Dela? Como A Obtemos?”, em Lorrie Faith Cranor e Simson Garfinkel, orgs.,*Segurança e Usabilidade*, O'Reilly Publishing, Sebastopol, CA, 2005, págs. 13-30.

Scandura, JM “Teorização Determinística na Aprendizagem Estrutural: Três Níveis de Empirismo,”*Jornal de Aprendizagem Estrutural*3, 1971, págs. 21–53.

Schneier, Bruce, “Ataques Semânticos: A Terceira Onda de Ataques à Rede”, em*Boletim informativo Crypto-Gram*, 15 de outubro de 2000, disponível emhttp://www.schneier.com/crypto-gram-0010.html

Simons, Daniel J. e CF Chabris, “Gorilas em nosso meio: cegueira desatenta sustentada por

Eventos dinâmicos."*Percepção*, 28, 1999, págs. 1059-1074.

Simons, Daniel J. e Melinda S. Jensen, "Os efeitos das diferenças individuais e da dificuldade das tarefas na "Cegueira por desatenção"*Boletim e Revista Psiconômicos*16(2), 2009, págs. 398-403.

Slovic, Paul, ed.*A Percepção do Risco*, Earthscan Ltd., Londres, 2000.

Smith, Walter, Becky Hill, John Long e Andy Whitefield, Andy, "Uma estrutura orientada para o design para Modelagem do Planejamento e Controle de Tarefas Múltiplas na Administração de Secretariado", *Comportamento e Tecnologia da Informação*, 16(3), 1997, págs. 161-183.

Spira, Jonathan B., "A conspiração terrorista do dia de Natal: como a sobrecarga de informações prevaleceu e "A partilha de conhecimentos sobre contraterrorismo falhou"*Basex TechWatch*, 4 de janeiro de 2010, disponível em http://www.basexblog.com/category/analysts/jonathan-b-spira/

Standing, L. "Aprendendo 10.000 Imagens"*Revista trimestral de psicologia experimental*, 27, 1973, págs. 207-222.

Tenner, Eduardo,*Por que as coisas reagem: tecnologia e a vingança das consequências não intencionais*, Imprensa Vintage, 1991.

Underwood, BJ, "Interferência e Esquecimento",*Revisão Psicológica*64, 1957, págs. 49-60.

Virginia Tech, "Quando os usuários resistem: como a gestão de mudanças e a resistência dos usuários à senha segurança,"*Pamplin*, Outono de 2011, disponível em http://www.magazine.pamplin.vt.edu/fall11/passwordsecurity.html

Wixted, John T., "A psicologia e a neurociência do esquecimento",*Revisão Anual de Psicologia*, 55, 2004, págs. 235-269.

Yamagishi, T. e M. Matsuda, "O papel da reputação em sociedades abertas e fechadas: um estudo experimental Estudo de Comércio Online", Centro de Estudos de Fundamentos Culturais e Ecológicos da Mente, Trabalhando Série de artigos 8, 2003.

Zviran, Moshe e William J. Haga, “Senhas cognitivas: a chave para um controle de acesso fácil”, *Computadores e Segurança* 8(9), 1990, págs. 723-736.

Este trabalho foi patrocinado por bolsas do Instituto de Proteção de Infraestrutura de Informação da Dartmouth College, sob o número de prêmio 2006-CS-001-000001 do Departamento de Segurança Interna, Diretoria Nacional de Segurança Cibernética.