Approche cognitive des cyberattaques par ingénierie sociale

Bruno Teboul

THE EUROPEAN SCIENTIST

Approche cognitive des cyberattaques par ingénierie sociale.

Bruno Teboul

Paris, le 25 août 2021.

Introduction

Les cyberattaques par ingénierie sociale sont devenues une menace majeure car elles préludent souvent à des cyberattaques beaucoup plus sophistiquées et dévastatrices. Les cyberattaques par ingénierie sociale sont une forme « *d'attaque psychologique* » qui exploitent les faiblesses des fonctions cognitives humaines. Une défense adaptée et résiliente contre les cyberattaques d'ingénierie sociale nécessite une compréhension plus approfondie de la cognition exploitée par les cyberattaquants. Pourquoi sommes-nous plus sensibles aux cyberattaques par ingénierie sociale ? Comment pouvons-nous minimiser, atténuer leurs dommages considérables ? Ces questions ont reçu une certaine attention de la part des chercheurs en sciences cognitives, mais l'analyse de l'état de l'art reste encore superficielle et dispersée dans la littérature. L'objectif de notre article est de clarifier ces enjeux et de définir les bases d'une approche cognitive des cyberattaques par ingénierie sociale.

Les cyberattaques par ingénierie sociale sont l'une des types d'attaques informatiques qui exploitent les failles et les faiblesses psychologiques, bien humaines, en tentant de persuader un individu (une victime) à agir comme prévu, selon un scénario malicieux et efficace à la fois. Ces attaques informatiques exploitent les faiblesses des interactions humaines et des constructions comportementales/culturelles qui se produisent sous de nombreuses formes, y compris le « phishing », « l'escroquerie », les « fraudes », les « spams », le « spear phishing » et les « sock puppets » sur les réseaux sociaux.

1/ Les cyberattaques par ingénierie sociale sous l'angle de la psychologie cognitive.

En 2016, lors des élections américaines, les hackeurs ont utilisé sur les réseaux sociaux des « sock puppets » (également connues sous le nom de « trolls russes ») : il s'agissait d'identités fictives créées uniquement pour influencer (malgré eux) les opinions des citoyens américains (Linvill et al., 2019). L'efficacité relative des technologies de sécurité actuelles a contribué au fait que les attaques d'ingénierie sociale sont devenues les passerelles, les catalyseurs, qui accélèrent les intrusions et les malversations numériques, de plus en plus fréquentes et destructives (sur le plan humain, matériel et financier).

Les cyberattaques les plus sophistiquées et les plus dévastatrices commencent souvent par des cyberattaques d'ingénierie sociale, telles que le spear phishing, où l'attaquant accède à un réseau d'entreprise (Hutchins et al., 2011). En effet, Mitnick et Simon (2003) décrivent de

nombreuses façons d'accéder à des systèmes sécurisés en utilisant des cyberattaques d'ingénierie sociale. Les recherches en ingénierie sociale se sont principalement concentrées sur la compréhension et/ou la détection des attaques d'un point de vue technologique (par exemple, la détection des e-mails de phishing en analysant le contenu des e-mails).

Cependant, il n'y a toujours pas de compréhension systématique des composantes psychologiques de ces attaques, ce qui explique peut-être pourquoi ces attaques sont si répandues et réussies. Cela motive le présent article encore une fois, qui vise à systématiser l'analyse de la cognition humaine à travers le prisme des cyberattaques d'ingénierie sociale.

Du fait, que nous préconisons ici de traiter les cyberattaques par ingénierie sociale comme un type particulier d'attaque « psychologique », nous ouvre de nouvelles perspectives en sciences cognitives et nous permet de dresser les bases d'un domaine que l'on pourrait appeler « l'approche cognitive de la cybersécurité » : elle étend et adapte les principes des sciences cognitives et notamment de la psychologie cognitive au champ d'application de la cybersécurité, tout en embrassant de nouveaux concepts propres au domaine de la cybersécurité.

Cette approche pourrait ouvrir ainsi la voie à la conception de défenses et de stratégies efficaces et résilientes contre les cyberattaques d'ingénierie sociale. Elle garantit qu'elles sont construites sur la base d'hypothèses psychologiquement valides. Par exemple, on peut supposer que les individus sont prêts à participer aux réponses efficientes et résilientes aux cyberattaques d'ingénierie sociale dans la plupart des cas ; ou bien que les victimes sont simplement imprudentes et crédules face aux risques, aux failles ou aux alertes lors d'une cyberattaque. Cependant, ces hypothèses sont discutables car la majorité des cyberattaques d'ingénierie sociale sont conçues pour déclencher des réponses automatiques et inconscientes de la part des victimes tout en déguisant ces attaques en demandes légitimes.

2/ Elargissement du cadre d'étude de la cognition humaine à l'analyse des cyberattaques par ingénierie sociale.

Comme indiqué précédemment, la plupart des études se concentrent sur les aspects techniques des cyberattaques par ingénierie sociale. Par exemple, Gupta et al. (2016) enquêtent sur les défenses technologiques contre les attaques de phishing. Salahdine et Kaabouch (2019) passent en revue les cyberattaques d'ingénierie sociale et les stratégies d'atténuation, mais ils ne discutent pas de facteurs tels que la cognition humaine. Darwich et al. (2012) analysent la relation entre les facteurs humains, les cyberattaques d'ingénierie sociale et les stratégies de cyberdéfenses, mais ils n'examinent pas ce qui rend un individu psychologiquement vulnérable aux cyberattaques d'ingénierie sociale.

Le terme « cognition » peut avoir des significations radicalement différentes selon les contextes. Nous utilisons le terme « cognition » ici au sens le plus large possible, comme terme descriptif pour rendre compte de la faculté humaine de traiter de l'information à des fins de connaissance. Autrement dit, la cognition est le traitement abstrait de l'information mis en œuvre par les neurones du cerveau (Pinker, 2009).

Dans cette perspective, la cognition peut également inclure le traitement de l'information qui résulte de la perception, des sensations, des émotions, des sentiments ainsi que la grande majorité du traitement de l'information neuronale qui ne se reflète pas dans notre conscience (Baars, 1997).

Les psychologues cognitifs considèrent souvent le traitement de l'information comme la fonction de base du cerveau, de la même manière que le foie fonctionne comme un filtre

complexe ou que les artères et les veines sont essentiellement des tuyaux de transport. Les corrélats du traitement de l'information dans le cerveau peuvent être observés directement en utilisant diverses méthodes pour enregistrer l'activité électrique et chimique cérébrales.

On peut distinguer quatre grandes fonctions cognitives de base, elles-mêmes centrées sur quatre composants du traitement de l'information naturelle analogues aux composants logiciels d'un système de traitement de l'information artificiel. Ces quatre composantes sont la perception, la mémoire de travail, la prise de décision et l'action.

La perception convertit les informations contenues dans une phrase, extrait le sens (signification et direction) en codes neuronaux qui peuvent être utilisés pour un comportement intelligent et une expérience consciente (Mesulam, 1998). La mémoire de travail se compose de l'attention et de la mémoire à court terme et coordonne le traitement des informations en donnant la priorité à certaines informations pendant de courtes périodes de temps, souvent pour atteindre un objectif précis (Miyake et Shah, 1999). La prise de décision donne en outre la priorité aux informations provenant de la mémoire de travail et d'autres sources inconscientes et constitue une passerelle vers le comportement (Kahneman, 2011). L'action est la mise en œuvre de calculs issus de la prise de décision et organise également l'activité physique des muscles et des glandes qui sont mesurables en tant que comportement (Franklin et Wolpert, 2011).

La perception, la mémoire de travail, la prise de décision et l'action sont souvent considérées comme étant à peu près séquentielles, mais peuvent s'influencer mutuellement de plusieurs manières. Tous ces processus cognitifs fonctionnent sur une base de connaissances accumulées dans la mémoire, qui informe ces processus, par exemple lors de la perception d'un visage familier.

3/ Charge cognitive, stress et vigilance face aux cyberattaques par ingénierie sociale.

Concentrons-nous à présent, sur trois facteurs cognitifs qui exercent leur influence à court terme : la charge cognitive, le stress et la vigilance. Ces facteurs opèrent sur des échelles de temps relativement courtes (de quelques minutes à quelques heures) et ont fait l'objet de nombreuses études. Nous examinerons comment ces facteurs peuvent être liés à l'influence sur les comportements face à une cyberattaque par ingénierie sociale.

Selon les situation deux tâches peuvent être effectuées en même temps avec peu ou pas de coûts de performance, on parle alors de charge cognitive gérable ; ou bien ces mêmes tâches effectuées ensemble, peuvent être presque impossibles à réaliser tant la charge cognitive nécessaire est élevée. Un bon exemple vient de Shaffer (1975), qui a découvert que les dactylos pouvaient lire et taper très précisément tout en répétant verbalement un message vocal. Les performances, cependant, ont chuté sur les deux tâches lorsqu'elles ont essayé de prendre une dictée (en saisissant un message vocal) tout en essayant de lire un message écrit à haute voix. On pense que les différences reflètent l'utilisation de codes cognitifs phonologiques (basés sur le son) et orthographiques (basés sur les lettres visuelles) différents et difficilement parallélisables.

Dans le premier exemple, un code est utilisé par tâche (phonologique : écoute de la parole ; orthographique : type de lecture), tandis que dans le second, chaque code est utilisé pour les deux tâches (type de parole ; lecture-parole). Pour tenir compte de ces complexités, les psychologues ont développé des théories qui prennent en compte différents types de codes cognitifs (Navon et Gopher, 1979), tels que les entrées sensorielles auditives ou visuelles, les codes verbaux ou spatiaux de niveau supérieur et les codes de sortie pour conduire la parole ou les comportements manuels. (Wickens, 2008).

Le stress aigu peut également influencer la cognition et les comportements des victimes de cyberattaques. Nous distinguons le stress aigu du stress chronique, le stress chronique commençant après une durée de l'ordre de quelques mois, car son impact sur la cognition peut différer et le stress chronique est plutôt catégorisé comme un facteur d'influence sur le long terme. Les réponses neurobiologiques et hormonales à un événement stressant ont été relativement bien analysées dans la littérature spécialisée, tout comme leur impact sur le comportement (Lupien et al., 2009).

Le stress aigu peut influencer l'attention, une composante vitale de la mémoire de travail, de manière aussi bien bénéfique que préjudiciable (Al'Absi et al., 2002). L'effet tunnel attentionnel est l'un de ces effets du stress aigu où l'attention est hyper-focalisée sur les aspects pertinents à la cause du stress, mais elle est moins sensible aux autres informations. Le terme tunneling dérive de l'utilisation de tâches d'attention spatiale, où l'excitation due au stress conduit les sujets à ignorer des choses qui sont plus éloignées de leur centre d'attention (Mather et Sutherland, 2011).

Dans le domaine de la cybersécurité, le tunneling d'attention à partir d'un message de phishing peut conduire à une hyper-concentration sur le texte de l'e-mail mais peut conduire à ignorer une adresse suspecte ou à passer à côté d'alertes périphériques. La mémoire de travail est également vulnérable au stress aigu (Schwabe et Wolf, 2013), notamment en interférant avec la fonction du cortex préfrontal (Elzinga et Roelofs, 2005 ; Arnsten, 2009).

La prise de décision peut être conduite de deux manières fondamentalement différentes (Evans, 2008). La première se produit via des processus relativement automatiques qui sont rapides mais qui peuvent ne pas être un choix optimal dans certains cas (appelés « heuristiques » et « biais ») (Tversky et Kahneman, 1974 ; Gigerenzer, 2008). La deuxième approche consiste à utiliser un raisonnement de traitement conscient et contrôlé, qui est plus lent mais qui peut être plus sensible aux particularités d'une situation donnée. Le stress aigu a une variété d'effets sur la prise de décision et de nombreuses subtilités (Starcke et Brand, 2012), mais il peut, en général, altérer la prise de décision rationnelle. L'un des moyens pour l'endiguer consiste à réduire la probabilité d'une prise de décision contrôlée et à augmenter l'utilisation de traitement automatique.

Vigilance et attention sont deux termes étroitement liés, parfois synonymes, qui décrivent des performances cognitives qui changent systématiquement au fur et à mesure que l'on effectue une tâche donnée. De nombreux travaux ont montré que les performances dans un large éventail de tâches diminuent considérablement au cours de ces périodes de temps relativement courtes, appelées « diminution de la vigilance » (Parasuraman et Rizzo, 2008).

L'impact potentiel de la baisse de vigilance sur le comportement est un facteur important à explorer, car la probabilité d'erreur de l'utilisateur peut varier avec le temps passé à la tâche. Par exemple, la probabilité de télécharger des logiciels malveillants peut augmenter à mesure que les utilisateurs consultent leur boîte de réception, en particulier s'ils disposent de peu de temps. Enfin, nous notons que si les catégories situationnelles de charge cognitive, de stress et de vigilance sont séparément importantes à examiner dans le domaine de la cybersécurité, elles sont également considérées pour interagir les unes avec les autres. Par exemple, une charge mentale élevée et une vigilance prolongée sont stressantes (Parasuraman et Rizzo, 2008). Une autre distinction à garder à l'esprit est que de nombreuses tâches de vigilance sont ennuyeuses et nécessitent une faible charge cognitive. Les études sur la vigilance se généralisent à d'autres environnements. Des salariés en entreprise peuvent avoir des charges de travail élevées et un stress dû à des exigences professionnelles complexes. Cette seule question mériterait d'être examinée dans le cadre d'études appliquées à la cybersécurité.

4/ Psychologie de la « personnalité » des victimes de cyberattaques par ingénierie sociale : le framework « Big 5 ».

Contrairement aux facteurs à court terme qui reflètent la situation actuelle et peuvent changer rapidement, les « facteurs à long terme » couvrent les attributs plus stables d'une personne et de ses expériences, qui elles ne changent que progressivement. Nous tenons compte des facteurs de personnalité, d'expertise, d'âge, de sexe et de culture. Ces facteurs offrent une certaine prévisibilité du comportement individuel dans une situation donnée. Dans le contexte de la cybersécurité, des facteurs psychologiques à long terme peuvent avoir un impact sur la façon dont un individu réagit aux attaques informatiques par ingénierie sociale.

Pour les psychologues, la « personnalité » est un terme technique qui diffère quelque peu de l'usage ordinaire. Il fait référence aux différences individuelles affectant les pensées, les sentiments et les comportements qui sont relativement cohérents au fil du temps et des situations. Nous disons « relativement » parce que, comme indiqué ci-dessus, les pensées, les sentiments et les comportements dépendent fortement de la situation, et les approches axées sur la durée de vie ont prouvé des changements notables de la personnalité avec l'âge (Donnellan et Robins, 2009).

Les recherches en psychologie cognitive sur la personnalité sont dominées par le « *Framework Big 5* » de la personnalité. Ce cadre théorique et méthodologique a été développé durant une grande partie du vingtième siècle sous diverses formes (Digman, 1997). Le Framework Big 5 est basé sur des méthodes statistiques (analyse factorielle) qui identifient des dimensions abstraites qui peuvent économiquement expliquer une grande partie de la variance dans les mesures de la personnalité. Les facteurs sont analysés sont : conscience, amabilité, neuroticisme, ouverture à l'expérience et extraversion. De nombreuses études sur la relation entre l'ingénierie sociale et la personnalité se concentrent sur l'ouverture, la conscience et le neuroticisme, qui sont considérés comme ayant le plus d'impact sur la vulnérabilité et la sensibilité aux attaques par ingénierie sociale.

Les facteurs qui composent le Framework Big 5 sont les suivants :

1. Ouverture : la volonté d'expérimenter de nouvelles choses.
2. Conscience : favorise les normes, fait preuve de maîtrise de soi et l'autodiscipline et la compétence.
3. Extraversion : être plus convivial, extraverti et interactif avec plus de monde.
4. Agréabilité : être coopératif, désireux d'aider les autres et croire à la réciprocité.
5. Neuroticisme : tendance à ressentir des sentiments négatifs, de la culpabilité, le dégoût, la colère, la peur et la tristesse.

L'expertise est généralement limitée à un domaine relativement étroit et ne se transfère pas à d'autres domaines (aussi facilement que nous aurions tendance à le croire). Ce point théorique est aussi appelé « *problème de transfert* » dans la littérature (Kimball et Holyoak, 2000). Le transfert limité d'expertise peut être aggravé par des illusions cognitives telles que l'effet Dunning-Kruger. L'effet Dunning-Kruger montre empiriquement que les individus surestiment souvent leur compétence par rapport à leur performance objective (Kruger et Dunning, 1999). De même, « *l'illusion du savoir* » montre que les gens en savent généralement beaucoup moins sur un sujet qu'ils ne le croient, comme le révèle l'article de Keil publié en 2003.

Dans le domaine de la cybersécurité, ces phénomènes empiriques renforcent la confiance des utilisateurs. Une expertise restreinte en matière de cybersécurité peut être bénéfique, mais une expertise informatique très générale ne suffit pas à conférer des avantages ou des bénéfices réels en matière de sécurité.

5/ Les cyberattaques par ingénierie sociale définies comme « attaques psychologiques ».

Comme nous l'avons dit plus haut, les cyberattaques d'ingénierie sociale sont un type d'attaque psychologique qui exploite les fonctions cognitives humaines pour persuader un individu (c'est-à-dire une victime) de se conformer à la demande d'un attaquant (Anderson, 2008). Ces attaques sont centrées sur un message d'ingénierie sociale élaboré par un hackeur dans le but de persuader une victime d'agir comme il le souhaite. Ces attaques s'appuient souvent sur des constructions comportementales et culturelles pour manipuler une victime afin qu'elle prenne une décision basée sur la satisfaction (gratification), plutôt que sur la base du meilleur résultat (optimisation) (Kahneman, 2011 ; Indrajit, 2017). A titre d'exemple, la plupart des individus échangeraient des informations sur leur vie privée pour des raisons de commodité, ou négocieraient la divulgation d'informations contre une récompense (Acquisti et Grossklags, 2005).

A l'ère du déluge informationnel et de nos communications numériques continues, la charge cognitive peut affecter la capacité d'un individu à traiter des messages d'ingénierie sociale. Pfleeger et Caputo (2012) observent que la charge cognitivee pourrait amener les individus à négliger des éléments qui ne sont pas associés à la tâche principale. Cet effet, appelé « *cécité d'inattention* », affecte la capacité d'un individu à remarquer les événements périphériques en se concentrant sur la tâche principale (Simons, 2000).

Dans la plupart des cas, la sécurité est une tâche périphérique ou secondaire. Par exemple, lorsqu'un employé tente de gérer plusieurs tâches simultanément : répondre à des centaines d'e-mails, tout en répondant à des appels téléphoniques et à une demande occasionnelle d'un supérieur, l'employé est plus susceptible d'ignorer les indices contenus dans les messages d'hameçonnage qui pourraient l'alerter sur une arnaque, une escroquerie. Une étude de 2020 a examiné le comportement d'hameçonnage réel en envoyant aux employés un e-mail d'hameçonnage inoffensif : les résultats de l'étude ont révélé que plus la surcharge mentale (auto-perçue) était élevée, plus elle était associée à la probabilité de cliquer sur le lien d'hameçonnage (Jalali et al.).

Vishwanath et al. (2011) tirent parti de deux attaques de phishing qui ciblent une université, et interrogent les étudiants de premier cycle sur leurs souvenirs et leur réponse aux e-mails de phishing. Ils constatent qu'en présence d'un e-mail perçu comme pertinent, les individus se concentrent davantage sur les indices d'urgence, tout en négligeant les indices d'escroquerie dans le message, tels que l'adresse e-mail de l'expéditeur ou la grammaire/l'orthographe des e-mails. Ils constatent également que les personnes qui gèrent régulièrement de gros volumes d'e-mails font preuve d'une grande inattention lors de l'évaluation des e-mails, ce qui les rend plus vulnérables aux attaques de type phishing. Ils constatent également qu'une charge élevée de courrier électronique déclenche une réponse automatique, ce qui signifie que la charge cognitve augmente considérablement la vulnérabilité d'une victime, face aux attaques de type phishing.

6/ La vulnérabilité des victimes de cyberattaques et le niveau d'expertise en cybersécurité : conscience, auto-efficacité et formation.

S'agissant de l'usage d'internet et des navigateurs Web : Kumaraguru et al. ont montré dès 2006 comment la façon d'afficher les informations sur un site web et notamment la présence de certificats de sécurité sont déterminantes et varient en fonction des individus et de leur niveau d'expertise: (i) les individus non experts considèrent moins d'indicateurs de sécurité que les experts ; (ii) des individus non experts utilisent des règles simples pour déterminer la

légitimité d'une demande, tandis que les experts prennent également en compte d'autres informations utiles (par exemple, le contexte) qui permet de révéler la présence de failles, ou d'anomalies de sécurité; (iii) les individus non experts prennent des décisions sur la base de leurs émotions, tandis que les experts prennent leurs décisions sur la base du raisonnement ; et (iv) les individus non experts s'appuient davantage sur des éléments visuels (falsifiables) pour prendre des décisions car ils ne savent pas que ces indicateurs de sécurité peuvent être compromis, tandis que les experts sont plus efficaces pour identifier les éléments suspects dans un message. Par exemple, ils observent qu'un individu non expert peut décider de télécharger un logiciel en fonction de son envie, si le site de téléchargement est juste disponible et attractif ; alors qu'un expert pourrait déterminer à quel point il en a réellement besoin et si le site de téléchargement est une source fiable.

La conscience est généralement associée à la maîtrise de soi, qui diminue le comportement impulsif (Cho et al., 2016). Pattinson et al. (2012) constatent que les individus moins impulsifs gèrent mieux les messages de phishing. Halevi et al. (2015) montrent que les personnes ayant une conscience élevée et une perception du risque plus faible sont davantage susceptibles d'être victimes de messages de cyberattaque par ingénierie sociale. Lawson et al. (2018) constatent que l'extraversion diminue la précision de la détection du phishing tandis qu'une conscience élevée augmente la précision de la détection, et que l'ouverture est associée à une plus grande précision dans la détection des messages légitimes.

Darwich et al. (2012) constatent que les individus plutôt extravertis et dotés d'amabilité, posent un risque plus élevé pour la sécurité informatique. McBride et al. (2012) constatent que la conscience est associée à une faible auto-efficacité et à une faible appréciation de la menace cyber. En règle générale, la formation des non-experts met souvent l'accent sur leur sensibilisation.

Dans une étude sur les victimes de fraudes de phishing et de logiciels malveillants, Jansen et Leukfeldt (2016) avancent que la plupart des sujets déclarent avoir des connaissances en matière de cybersécurité, mais il s'avère que seuls quelques-uns d'entre eux ont effectivement les connaissances revendiquées. Downs et al. (2006) constatent que la prise de conscience des failles de sécurité dans les messages de phishing, ne se traduit pas par des comportements responsables et conscients : en effet, la plupart des participants sont incapables d'assumer leurs actions et les conséquences néfastes qu'elles peuvent parfois entrainer.

Il peut être intuitif de penser que les personnes ayant reçu une formation informatique sont beaucoup moins exposées et victimes de cyberattaques par ingénierie sociale. C'est tout le contraire, Ovelgönne et al. (2017) constatent que les développeurs de logiciels sont impliqués dans plus d'incidents de cyberattaque que les non informaticiens. Purkait et al. (2014) relatent qu'il n'y a aucune relation entre la capacité d'une personne à détecter les sites d'hameçonnage et son niveau d'éducation et/ou son backround technique (Halevi et al. 2013).

Harrison et al. (2016) constatent que les connaissances sur les attaques de phishing augmentent l'attention et l'élaboration lorsqu'elles sont combinées à des connaissances et à une expérience subjective ; elles réduisent donc la susceptibilité d'être victime de messages de cyberattaque par ingénierie sociale. Wang et al. (2012) constatent que la connaissance des attaques de phishing augmente l'attention pour détecter les indices, les signes « avant-coureurs ».

Pattinson et al. (2012) indiquent que plus la maitrise de l'informatique est élevée, plus la capacité à faire face aux messages de phishing est importante, déterminante. Wright et Marett (2010) constatent (i) qu'une combinaison de connaissances et de formation en cybersécurité est efficace contre les attaques de phishing ; (ii) les personnes dont l'auto-efficacité (c'est-à-dire la capacité à gérer des événements inattendus) et l'expérience Web sont moindres, sont

donc beaucoup plus susceptibles d'être victimes de cyberattaques d'ingénierie sociale ; et (iii) les personnes ayant une auto-efficacité élevée sont moins susceptibles de se conformer aux demandes d'informations suggérées lors d'attaques par phishing. Halevi et al. (2016) disent qu'une auto-efficacité élevée est corrélée à une meilleure capacité à répondre aux incidents de sécurité informatique.

Arachchilage et Love (2014) précisent que l'auto-efficacité, lorsqu'elle est combinée à des connaissances sur les attaques de type phishing, peut conduire à des stratégies efficaces pour lutter contre ce type de cyberattaques. Wright et Marett (2010) constatent que les facteurs expérientiels (Van Schaik et al. 2017) associés à une perception du risque plus élevée des menaces est pertinente et efficace dans la prévention des menaces.

Redmiles et al. (2018) montrent qu'avec l'essor du e-commerce, les individus qui achètent en ligne sont compétents pour identifier les spams et moins susceptibles de cliquer sur les liens avec contenus malicieux. Gavett et al. (2017) constatent que l'éducation et l'expérience antérieure avec les attaques de phishing ont augmenté la suspicion sur les sites de phishing. Caïn et al. (2018) observent que les incidents de sécurité passés n'affectent pas de manière significative les comportements sécurisés.

Abbasi et al. (2016) constatent que (i) les femmes et les hommes plus âgés et instruits qui ont été victimes d'attaques de phishing dans le passé sont moins susceptibles d'être à nouveau victimes d'attaques de phishing ; (ii) les jeunes femmes peu sensibilisées au phishing et ayant déjà subi de légers dommages (matériels et/ou financiers) causées par des attaques de type phishing n'ont pas nécessairement une sensibilité plus faible, aux futures attaques de phishing ; et (iii) les jeunes hommes ayant une grande auto-efficacité et une conscience élevée du phishing et des expériences antérieures d'attaques de phishing n'ont pas non plus nécessairement une plus faible sensibilité aux futures attaques de phishing.

Conclusion

Le succès des cyberattaques d'ingénierie sociale est inversement lié à leur prévalence, cela pose ainsi un dilemme : lorsque les défenses automatisées sont efficaces pour détecter et filtrer la plupart des cyberattaques d'ingénierie sociale, les attaques restantes ont plus de chances d'aboutir. Une approche pour faire face à ce dilemme est de recourir aux principes de la psychologie cognitive comme nous l'avons fait ici. On sait que la majeure partie du traitement de l'information neuronal est isolé de la conscience (Nisbett et Wilson, 1977). Certaines informations pourraient être consciemment utilisées, mais elles peuvent aussi ne pas être conscientes à un moment donné. Notre système visuel, par exemple, calcule la profondeur 3-D à partir d'entrées rétiniennes 2-D (DeValois et DeValois, 1990). Nous n'expérimentons pas consciemment les calculs nécessaires pour transformer l'entrée 2-D en une perception 3-D. Au lieu de cela, nous sommes conscients du résultat produit (c'est-à-dire de voir un monde en 3D) mais pas du processus qui a conduit au même résultat final. Les influences du traitement subconscient sont bien connues pour avoir un impact sur le comportement humain (Kahneman, 2011 ; Nosek et al., 2011).

En synthèse, la plupart des informations mobilisées dans le cadre de notre revue de littérature nous permet de dresser le constat suivant: (i) la charge cognitive et le stress augmentent la vulnérabilité aux cyberattaques d'ingénierie sociale ; (ii) l'effet de la vigilance, de la personnalité, de la conscience et de la culture reste à étudier pour être concluant ; (iii) la connaissance du domaine, (certains types d'expériences) et l'âge (ainsi que certains autres facteurs) réduisent la vulnérabilité aux cyberattaques d'ingénierie sociale ; et (iv) le genre peut ne pas avoir d'effet significatif sur la vulnérabilité d'une personne aux cyberattaques d'ingénierie sociale.

Nous pensons que l'impact du niveau d'expertise/connaissance sur la vulnérabilité aux cyberattaques d'ingénierie sociale est crucial. Effectivement, le niveau expertise/connaissance peut réduire la vulnérabilité d'une personne, car l'expertise permet de détecter les indices trompeurs utilisés par les cyberattaquants utilisant l'ingénierie sociale. Il est plus difficile pour un expert d'être victime de cyberattaques d'ingénierie sociale, tant la vigilance et la connaissance du domaine « cyber » ont un impact significatif, sur la réduction de la vulnérabilité.

Nous formulons que le stress réduit la capacité d'une personne à détecter les indices d'attaques ou de tentatives d'escroqueries ou de compromissions, mais l'effort d'attaque aurait un impact encore plus important sur l'augmentation de la vulnérabilité. Nous pensons que la charge cognitive peut avoir un plus grand impact sur la vulnérabilité d'une personne, car elle réduirait considérablement sa capacité à détecter les indices malveillants.

Enfin, nous considérons que l'importance de tous ces facteurs peut être spécifique aux scénarios d'attaque. Ceci est corroboré par deux études récentes : van der Heijden et Allodi (2019) ont montré que certains facteurs comme la charge cognitive peuvent être exploités pour mener des attaques de phishing, car les e-mails malveillants peuvent coïncider avec un volume d'e-mails élevé ; et Jalali et al. (2020) ont confirmé que certains facteurs à court et à long terme (par exemple, une charge cognitive élevée et un manque d'expertise) sont deux facteurs impactants particulièrement les personnels travaillant le secteur médical.
La sensibilisation et les connaissances techniques générales ne réduisent pas nécessairement la vulnérabilité aux cyberattaques d'ingénierie sociale.

Nous voyons combien la formation continue et initiale (au sens universitaire) à la cybersécurité sous l'angle de la psychologie cognitive sont fondamentales et urgentes en France. Nous devons mettre en place les cursus idoines et proposer aux entreprises des sessions de formations à tous les salariés. Il est nécessaire de construire en partenariat avec l'ANSSI, des universités, des industriels des programmes de formations dédiées, si l'on veut réduire la prolifération et l'impact dans toutes les organisations publiques et privées des cyberattaques par ingénierie sociale. L'approche cognitive de la cybersécurité est sans aucun doute la clef de voute d'une stratégie préventive, fiable et résiliente, pour toutes les entreprises désireuses de se protéger en amont des cyberattaques.

Formation et éducation à l'approche cognitive de la cybersécurité doivent devenir des priorités nationales. Former les salariés du public et du privé à la compréhension des mécanismes psychologiques qui structurent et sous-tendent notre exposition et notre vulnérabilité à la cyberdélinquance et à la cybercriminalité est vital dans une société numérique hyperconnectée.

Bruno Teboul

Bibliographie

Abbasi, A., Zahedi, F. M., and Chen, Y. (2016). “Phishing susceptibility: the good, the bad, and the ugly,” in 2016 IEEE Conference on Intelligence and Security Informatics (ISI) (Tucson: IEEE), 169–174. doi: 10.1109/ISI.2016.77 45462

Acquisti, A., and Grossklags, J. (2005). Privacy and rationality in individual decision making. IEEE Secur. Privacy 3, 26–33. doi: 10.1109/MSP.2005.22

Al’Absi, M., Hugdahl, K., and Lovallo, W. R. (2002). Adrenocortical stress responses and altered working memory performance. Psychophysiology 39, 95–99. doi: 10.1111/1469-8986.3910095

Anderson, R. J. (2008). Security Engineering: A Guide to Building Dependable Distributed Systems, 2nd Edn. New York, NY: Wiley Publishing.

Arachchilage, N. A. G., and Love, S. (2014). Security awareness of computer users: a phishing threat avoidance perspective. Comput. Hum. Behav. 38, 304–312. doi: 10.1016/j.chb.2014.05.046

Arnsten, A. F. (2009). Stress signalling pathways that impair prefrontal cortex structure and function. Nat. Rev. Neurosci. 10:410. doi: 10.1038/nrn2648

Baars, B. J. (1997). In the theatre of consciousness. global workspace theory, a rigorous scientific theory of consciousness. J. Conscious. Stud. 4, 292–309. doi: 10.1093/acprof:oso/9780195102659.001.1

Cho, J.-H., Cam, H., and Oltramari, A. (2016). “Effect of personality traits on trust and risk to phishing vulnerability: modeling and analysis,” in 2016 IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA) (San Diego, CA: IEEE), 7–13.

Darwish, A., El Zarka, A., and Aloul, F. (2012). “Towards understanding phishing victims’ profile,” in 2012 International Conference on Computer Systems and Industrial Informatics (Sharjah: IEEE), 1–5. doi: 10.1109/ICCSII.2012.6454454

DeValois, R. L., and DeValois, K. K. (1990). Spatial Vision, Vol. 14. Oxford University Press. doi: 10.1093/acprof:oso/9780195066579.001.0001

Digman, J. M. (1997). Higher-order factors of the big five. J. Pers. Soc. Psychol. 73:1246. doi: 10.1037/0022-3514.73.6.1246

Donnellan, M. B., and Robins, R. W. (2009). “The development of personality across the lifespan,” in The Cambridge Handbook of Personality Psychology,

Downs, J. S., Holbrook, M. B., and Cranor, L. F. (2006). “Decision strategies and susceptibility to phishing,” in Proceedings of the Second Symposium on Usable Privacy and Security (Pittsburgh, PA: ACM), 79–90. doi: 10.1145/1143120.1143131

Elzinga, B. M., and Roelofs, K. (2005). Cortisol-induced impairments of working memory require acute sympathetic activation. Behav. Neurosci. 119:98. doi: 10.1037/0735-7044.119.1.98

Evans, J. S. B. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. Annu. Rev. Psychol. 59, 255–278. doi: 10.1146/annurev.psych.59.103006.093629

Franklin, D. W., and Wolpert, D. M. (2011). Computational mechanisms of sensorimotor control. Neuron 72, 425–442. doi: 10.1016/j.neuron.2011.10.006

Gavett, B. E., Zhao, R., John, S. E., Bussell, C. A., Roberts, J. R., and Yue, C. (2017). Phishing suspiciousness in older and younger adults: the role of executive functioning. PLoS ONE 12:e0171620. doi: 10.1371/journal.pone.0171620

Gupta, S., Singhal, A., and Kapoor, A. (2016). "A literature survey on social engineering attacks: phishing attack," in 2016 International Conference on Computing, Communication and Automation (ICCCA) (Greater Noida: IEEE), 537–540. doi: 10.1109/CCAA.2016.7813778

Halevi, T., Lewis, J., and Memon, N. (2013). "A pilot study of cyber security and privacy related behavior and personality traits," in Proceedings of the 22nd International Conference on World Wide Web (Singapore: ACM), 737–744. doi: 10.1145/2487788.2488034

Halevi, T., Memon, N., and Nov, O. (2015). Spear-phishing in the wild: a real-world study of personality, phishing self-efficacy and vulnerability to spear-phishing attacks. SSRN Electron. J. doi: 10.2139/ssrn.2544742.

Harrison, B., Svetieva, E., and Vishwanath, A. (2016). Individual processing of phishing emails: how attention and elaboration protect against phishing. Online Inform. Rev. 40, 265–281. doi: 10.1108/OIR-04-2015-0106

Hutchins, E. M., Cloppert, M. J., and Amin, R. M. (2011). Intelligence-driven computer network defense informed by analysis of adversary campaigns and intrusion kill chains. Lead. Issues Inform. Warfare Secur. Res. 1:80.

Indrajit, R. E. (2017). Social engineering framework: Understanding the deception approach to human element of security. Int. J. Comput. Sci. Issues 14, 8–16. doi: 10.20943/01201702.816

Jalali, M. S., Bruckes, M., Westmattelmann, D., and Schewe, G. (2020). Why employees (still) click on phishing links: investigation in hospitals. J. Med. Internet Res. 22:e16775. doi: 10.2196/16775

Jansen, J., and Leukfeldt, R. (2016). Phishing and malware attacks on online banking customers in the Netherlands: a qualitative analysis of factors leading to victimization. Int. J. Cyber Criminol. 10:79. doi: 10.5281/zenodo.58523

Kahneman, D. (2011). Thinking, Fast and Slow. New York, NY: Farrar, Straus and Giroux.

Kimball, D. R., and Holyoak, K. J. (2000). "Transfer and expertise," in The Oxford Handbook of Memory, eds E. Tulving, and F. I. M. Craik (New York, NY: Oxford University Press), 109–122.

Kruger, J., and Dunning, D. (1999). Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments. J. Pers. Soc. Psychol. 77:1121. doi: 10.1037/0022-3514.77.6.1121

Kumaraguru, P., Acquisti, A., and Cranor, L. F. (2006). "Trust modelling for online transactions: a phishing scenario," in Proceedings of the 2006 International Conference on Privacy, Security and Trust: Bridge the Gap Between PST Technologies and Business Services (Markham, ON: ACM), 11. doi: 10.1145/1501434.1501448

Linvill, D. L., Boatwright, B. C., Grant, W. J., and Warren, P. L. (2019). “The Russians are hacking my brain!” investigating Russia’s internet research agency twitter tactics during the 2016 United States presidential campaign. Comput. Hum. Behav. 99, 292–300. doi: 10.1016/j.chb.2019. 05.027

Lupien, S. J., McEwen, B. S., Gunnar, M. R., and Heim, C. (2009). Effects of stress throughout the lifespan on the brain, behaviour and cognition. Nat. Rev. Neurosci. 10:434. doi: 10.1038/nrn2639

Mather, M., and Sutherland, M. R. (2011). Arousal-biased competition in perception and memory. Perspect. Psychol. Sci. 6, 114–133. doi: 10.1177/1745691611400234

McBride, M., Carter, L., and Warkentin, M. (2012). Exploring the role of individual employee characteristics and personality on employee compliance with cybersecurity policies. RTI Int. Instit. Homeland Secur. Solut. 5:1.

Mesulam, M.-M. (1998). From sensation to cognition. Brain 121, 1013–1052. doi: 10.1093/brain/121.6.1013

Mitnick, K., and Simon, W. L. (2003). The Art of Deception: Controlling the Human Element of Security. Indianapolis, IN: Wiley Publishing.

Miyake, A., and Shah, P. (1999). Models of Working Memory: Mechanisms of Active Maintenance and Executive Control. Cambridge, UK: Cambridge University Press. doi: 10.1017/CBO9781139174909

Navon, D., and Gopher, D. (1979). On the economy of the human-processing system. Psychol. Rev. 86:214. doi: 10.1037/0033-295X.86.3.214

Nisbett, R. E., and Wilson, T. D. (1977). Telling more than we can know: verbal reports on mental processes. Psychol. Rev. 84:231. doi: 10.1037/0033-295X.84.3.231

Nosek, B. A., Hawkins, C. B., and Frazier, R. S. (2011). Implicit social cognition: from measures to mechanisms. Trends Cogn. Sci. 15, 152–159. doi: 10.1016/j.tics.2011.01.005

Ovelgönne, M., Dumitras, T., Prakash, B. A., Subrahmanian, V. S., and Wang, B. (2017). Understanding the relationship between human behavior and susceptibility to cyber attacks: a data-driven approach. ACM Trans. Intell. Syst. Technol. 8:51:1–51:25. doi: 10.1145/2890509

Parasuraman, R., and Rizzo, M. (2008). Neuroergonomics: The Brain at Work, Vol. 3. New York, NY: Oxford University Press.

Pattinson, M., Jerram, C., Parsons, K., McCormac, A., and Butavicius, M. (2012). Why do some people manage phishing e-mails better than others? Inform. Manage. Comput. Secur. 20, 18–28. doi: 10.1108/09685221211219173

Pendleton, M., Garcia-Lebron, R., Cho, J.-H., and Xu, S. (2016). A survey on systems security metrics. ACM Comput. Surv. 49, 1–35. doi: 10.1145/3005714

Pinker, S. (2009). How the Mind Works (1997/2009). New York, NY: Norton & Company.

Purkait, S., Kumar De, S., and Suar, D. (2014). An empirical investigation
of the factors that influence internet user’s ability to correctly identify a phishing website. Inform. Manage. Comput. Secur. 22, 194–234. doi: 10.1108/IMCS-05-2013-0032

Redmiles, E. M., Chachra, N., and Waismeyer, B. (2018). “Examining the demand for spam: who clicks?” in Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (Montreal, ON: ACM), 212. doi: 10.1145/3173574.3173786

Salahdine, F., and Kaabouch, N. (2019). Social engineering attacks: a survey. Future Internet 11:89. doi: 10.3390/fi11040089

Schwabe, L., and Wolf, O. T. (2013). Stress and multiple memory systems: from ‘thinking? to ‘doing’. Trends Cogn. Sci. 17, 60–68. doi: 10.1016/j.tics.2012.12.001 Shaffer, L. (1975). Control processes in typing. Q. J. Exp. Psychol. 27, 419–432. doi: 10.1080/14640747508400502

Starcke, K., and Brand, M. (2012). Decision making under stress: a selective review. Neurosci. Biobehav. Rev. 36, 1228–1248. doi: 10.1016/j.neubiorev.2012. 02.003

Van Schaik, P., Jeske, D., Onibokun, J., Coventry, L., Jansen, J., and Kusev, P. (2017). Risk perceptions of cyber-security and precautionary behaviour. Comput. Hum. Behav. 75, 547–559. doi: 10.1016/j.chb.2017.05.038

Vishwanath, A., Harrison, B., and Ng, Y. J. (2018). Suspicion, cognition, and automaticity model of phishing susceptibility. Commun. Res. 45, 1146–1166. doi: 10.1177/0093650215627483

Wang, J., Herath, T., Chen, R., Vishwanath, A., and Rao, H. R. (2012). Phishing susceptibility: An investigation into the processing of a targeted spear phishing email. IEEE Trans. Profess. Commun. 55, 345–362. doi: 10.1109/TPC.2012.2208392

Wickens, C. D. (2008). Multiple resources and mental workload. Human Factors 50, 449–455. doi: 10.1518/001872008X288394

Wright, R. T., and Marett, K. (2010). The influence of experiential and dispositional factors in phishing: an empirical investigation of the deceived. J. Manage. Inform. Syst. 27, 273–303. doi: 10.2753/MIS0742-12222 70111