Computers & Security

An analysis of cognitive factors in the human decision-making process related to Social Engineering attacks. A Case study of Phishing.

--Manuscript Draft--

Manuscript Number:	COSE-D-21-00594
Article Type:	VSI:Incident Response
Keywords:	Cognition; phishing; Cognitive security; cyberattacks; social engineering
Corresponding Author:	Roberto Andrade Escuela Politecnica Nacional ECUADOR
First Author:	Maria Fernanda Cazares
Order of Authors:	Maria Fernanda Cazares
	Roberto Andrade
	Walter Fuertes
	Julio Proaño
Abstract:	The digital transformation in social and commercial aspects has driven new ways of interacting in the context of health, education, work and personal relationships. While this digital transformation has great benefits for society, it brings cybersecurity concerns. People and organizations are daily exposed to cybersecurity attacks such as phishing, ransomware, spyware, among others. Some of these attacks require the initial interaction of people, so attackers use persuasion techniques to exploit human vulnerabilities in the decision-making process to identify a phishing email. This chapter presents a literature review of Social Engineering attacks, focusing mainly on phishing. To do this, we perform a literature review on social engineering attacks and present a critical review of the cognitive factors, human vulnerabilities, and decision-making process during cybersecurity events. Afterward, we design a study case to simulate attacks on legitimate and infected emails with phishing using a Web Platform. Then, we try to define human vulnerabilities and the cognitive factors involved in these types of attacks with these inputs. The results reveal a cognitive model of users against phishing attacks and legitimate emails.
Suggested Reviewers:	

An analysis of cognitive factors in the human decision-making process related to Social Engineering attacks. A Case study of Phishing.

Abstract
The digital transformation in social and commercial aspects has driven new ways of interacting in the context of health, education, work and personal relationships. While this digital transformation has great benefits for society, it brings cybersecurity concerns. People and organizations are daily exposed to cybersecurity attacks such as phishing, ransomware, spyware, among others. Some of these attacks require the initial interaction of people, so attackers use persuasion techniques to exploit human vulnerabilities in the decision-making process to identify a phishing email. This chapter presents a literature review of Social Engineering attacks, focusing mainly on phishing. To do this, we perform a literature review on social engineering attacks and present a critical review of the cognitive factors, human vulnerabilities, and decision-making process during cybersecurity events. Afterward, we design a study case to simulate attacks on legitimate and infected emails with phishing using a Web Platform. Then, we try to define human vulnerabilities and the cognitive factors involved in these types of attacks with these inputs. The results reveal a cognitive model of users against phishing attacks and legitimate emails.

Keywords: Cognition; Phishing; Cognitive security; cyberattacks; social engineering

Author(s)
1) María Fernanda Cazares[1],
2) Roberto Andrade[2, †]
3) Walter Fuertes[3],
4) Julio Proaño[1],

Author(s) affiliations:
1.UniversidadPolitécnica Salesiana IDEAGEOCA, Quito, EC.
2.Escuela Politécnica Nacional, Quito, EC.
3.Universidad de las Fuerzas Armadas - ESPE, Sangolquí, EC.

Co-author mails:
1. mcazares@ups.edu.ec
2. roberto.andrade@epn.edu.ec
3. wmfuertes@espe.edu.ec
4. jproanoo@ups.edu.ec

Correspondence author detail:
1) roberto.andrade@epn.edu.ec, 593999112440, Quito, Ecuador, Escuela Politécnica Nacional, Quito, EC.

Author's contributions
Conceptualization, R.A.; methodology, M.C., and R.A.; formal analysis, J.P, and M.C.; investigation, R.A., and M.C.; writing–review and editing, M.C., and R.A and W.F; project administration, W.F. All authors have read and agreed to the published version of the manuscript

Funding
Not applicable.

Availability of data and materials
Not applicable.

Competing interests
The authors declare that they have no competing interests and that there is no conflict of interest regarding the publication of this manuscript.

Dear Editor-In-Chief.
Computer & Security.

Please, we send above the details of manuscript entitled: " **An analysis of cognitive factors in the human decision-making process related to Social Engineering attacks**. A Case study of Phishing.", authored by María Fernanda Cazares, Roberto Omar Andrade, Walter Fuertes, Julio Proaño,

We would like to submit our article for a possible publication. This study aims to identify the cognitive process during detection phase of phishing attacks.
Additionally, we ensure that our manuscript has not been submitted simultaneously for publication anywhere else, containing original data. There is no conflict of interest with any party; there is no problem with any kind of ethical standards. Therefore, we are able and willing to sign an agreement for the transfer of copyright to the publisher the moment our manuscript is accepted for publication.

Many thanks in advance for the time devoted to the handling of the review process. We are looking forward to receiving news from the Editorial Board regarding our submission soon.

With our best wishes,
On behalf of the authors.
Roberto Omar Andrade.

An analysis of cognitive factors in the human decision-making process related to Social Engineering attacks. A Case study of Phishing.

María Cazares, Clinical psychologist, Master in Integrative Psychotherapy, University Professor for 10 years, researcher in cognitive psychology in the line of artificial intelligence, applied to the field of psychodiagnosis and cybersecurity, professional practice at a private level for 13 years in childhood behavior disorders, depression, anxiety and personal development.

Roberto Omar Andrade, PhD student in Security Systems at the Faculty of Systems Engineering of the National Polytechnic School (EPN), his master's degree is in Network and Telecommunications Management at the Army Polytechnic School in 2013 and his engineering degree is in Electronics and Telecommunications in the National Polytechnic School (EPN) in 2007. He was the Education Security Officer of the Ministry of Education of Ecuador (MINEDUC) in 2015, Coordinator of Technological Infrastructure in the National Planning Secretariat SENPLADES 2013-2014, Data Center, Security and Administration of Networks in SENPLADES and Sucre Tecnológico 2009-2013 and Technical Engineering for VoIP systems in SERATVoIP 2007-2011. He is a certified CCNA, CCNP and CCNA Security Technical Instructor at EPN from 2010 to 2016. He is certified in Ethical Hacking and Incident Management and Response by Ec-Council.

Walter Fuertes received the Engineering degree in computer systems from School of Computer Science, Universidad de las Fuerzas Armadas (ESPE), Sangolquí, Ecuador, in 1995, the M.Sc. degree in computer networking from the Escuela Politécnica Nacional, Quito, Ecuador, in 1999, and the Ph.D. degree (Hons.) in computer science and telecommunications engineering from the Universidad Autónoma de Madrid, Madrid, Spain, in 2010.,Since 2006, he has actively participated in several research projects focused on the application of virtualization technologies, data sciences, and cyber security. He is currently with ESPE, where he is also a Full Professor (Lecturer–Researcher) with the School of Computer Science. He has authored several technical publications in scientific journals and national and international conferences around the world. His research interests include the management of distributed environments, the network security, cybersecurity, the applied research of virtualization technologies, and serious game.

Julio Proaño Orellana, Doctorado en tecnologias informaticas avanzadas, cuarto nivel -doctor ph.d, universidad de castilla la mancha uclamaster universitario en tecnologias informaticas avanzadas , cuartonivel - magister, universidad de castilla la mancha uclaingeniero electrico , tercer nivel - ingeniero, universidad politecnicasalesiana

Contents lists available at ScienceDirect

Computers & Security

journal homepage: www.elsevier.com/locate/cose

An analysis of cognitive factors in the human decision-making process related to Social Engineering attacks. A Case study of Phishing.

Abstract

The digital transformation in social and commercial aspects has driven new ways of interacting in the context of health, education, work and personal relationships. While this digital transformation has great benefits for society, it brings cybersecurity concerns. People and organizations are daily exposed to cybersecurity attacks such as phishing, ransomware, spyware, among others. Some of these attacks require the initial interaction of people, so attackers use persuasion techniques to exploit human vulnerabilities in the decision-making process to identify a phishing email. This chapter presents a literature review of Social Engineering attacks, focusing mainly on phishing. To do this, we perform a literature review on social engineering attacks and present a critical review of the cognitive factors, human vulnerabilities, and decision-making process during cybersecurity events. Afterward, we design a study case to simulate attacks on legitimate and infected emails with phishing using a Web Platform. Then, we try to define human vulnerabilities and the cognitive factors involved in these types of attacks with these inputs. The results reveal a cognitive model of users against phishing attacks and legitimate emails.

Keywords: Cognition; Phishing; Cognitive security; cyberattacks; social engineering

1. Introduction

According to the World Economic Forum in its Global Risk Report («Chapter One - Risks Landscape», s. f.) , Cybersecurity attacks are among the ten threats with the most significant impact in the world. Cybersecurity attacks use different attack vectors such as Phishing, Ransomware, Insider threats, Malware, and DDoS, among others.

Cyberattacks are focused on different targets such as critical infrastructures, organizational systems, or home devices (Joshi et al., 2021). One of the cybersecurity field interests has been attacks based on Social Engineering. The development of specialized security companies focused on protecting infrastructure; however, attackers saw opportunities due to the human vulnerabilities to making errors (Wang et al., 2020) . This fact can be verified with the growth of Phishing attacks during the COVID-19 pandemic. In one four-month period of 2020, the Interpol identified 907,000 spam messages, 737 incidents related to the malware, and 48,000 malicious URLs (*INTERPOL Report Shows Alarming Rate of Cyberattacks during COVID-19*, s. f.).

Cybersecurity Incident Response Team (CSIRT) establishes strategies to face against phishing attacks in two main methods: i) phishing campaigns to warn about phishing attacks and ii) the deployment of automated monitoring systems to detect phishing attacks. The first method is built on basis of development of communication channels via website or email with information to customers and user related to the way phishing attacks works. This first method is development in preparation, remediation and recovery phase of incident response process. The second method is developed on detection phase of incident response process, and it is based on two tasks: i) takedown forms with information give for users or customers when they generate notification of phishing messages; and ii) manual process of classification of phishing messages from spam traps by security analyst. In the Figure 1 is show a summarize incident response process for phishing attacks based on five phases: preparation, identification, containment, remediation and recovery, according to incident response plan proposal by SysAdmin Audit, Networking and Security Institute - SANS.

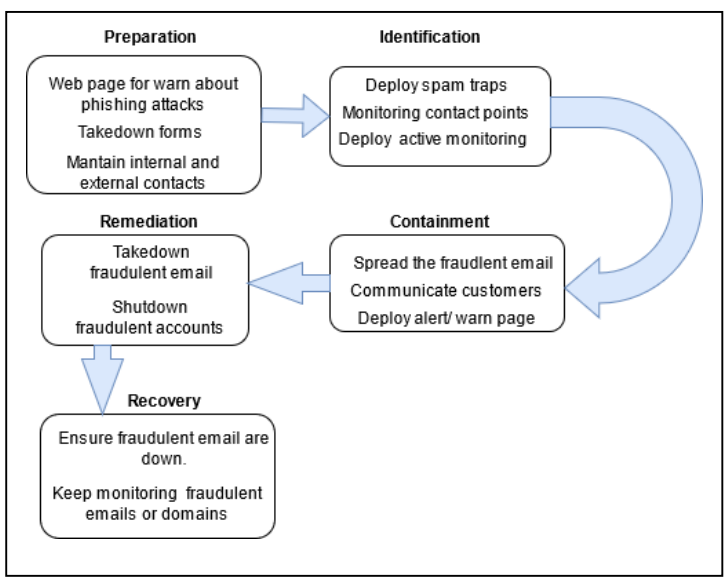

Fig. 1. Incident response for phishing attacks. (Source: own illustration)

The challenge for incident response related with phishing attacks is that social engineering or cognitive hacking attacks seek to take advantage of people's emotional factors such as fear, depression, stress, health, debt, and awareness has been established as a strategy against phishing attacks (Algarni et al., 2014) . On the contrary, some organizations affirm that the training does not avoid that people accept phishing content (January 7 & Minutes, s. f.). People often know about cybersecurity risks when interacting in the digital world; nevertheless, they can make wrong decisions and accept malicious emails or infected pages. This wrong decision is being associated with how humans process information in a systematic and heuristic way. Additionally, other cognitive factors may be involved in decision-making, such as cognitive dissonance or future thinking.

Concerning cognitive dissonance, people may believe what is correct, but external factors such as third-party opinions or trends can make them decide in the opposite direction. For instance, in the study entitled "Leveraging Behavioral Science to Mitigate Cyber Security Risk", Miter discusses aspects like Status Quo Bias or Framing Effect that could influence respect to people do not change their decision, or their decision is based on internal heuristics (Pfleeger et al., s. f.).

Additionally, spontaneous future thoughts could have been based on pre-made memories or experiences (Cole & Kvavilashvili, 2021). However, the cybersecurity domain problem is that these past experiences or memories could be fake (false memories) induced for attackers using fake news (Brown & Reavey, 2017). On the other hand, attackers have perfected their persuasion techniques and seek to take advantage of emotional factors that can alter judgment in decision-making processes (Valaskivi, s. f.).

The research question that we intend to examine in this study is the following: Is it possible to increase the phishing identification ratio using a phishing training website?. Our research hypothesis is as follows: Detection of phishing attacks improves when the mental processes involved are identified. Therefore, this study aims to determine the cognitive factors associated with the Social Engineering attack, focusing mainly on phishing, understood as a fraud that intends to obtain users' private data through the Internet, especially to access their accounts or bank details.

We used Design Science, a results-based information technology research methodology that offers specific guidelines for evaluation and iteration within research projects to achieve this goal. Then, we performed a literature review on social engineering attacks where the human factor and the cognitive process are involved in the face of phishing attacks. Subsequently, we design and implement a Web platform using the cascade Web development methodology for training simulated attacks on legitimate and infected emails with phishing to identify human vulnerabilities and the cognitive factors involved in these types of attacks.

The remainder of this book chapter is structured as follows: Section 1 presents a literature review of Social Engineering attacks and their trends for the following years. Section 2 presents a critical review of the cognitive factors used during cybersecurity events. Section 3 explains the experiment used to identify phishing attacks. Section 4 describes an analysis of the results. Finally, section 5 gives the conclusions and future work lines of the study.

2. Social engineering attacks

International organizations such as the World Economic Forum considered cybersecurity attacks one of the top ten threats to the global economy («Chapter One - Risks Landscape», s. f.). During the last few years, the

increase of cyberattacks has had a considerable variation to be taken into account by individuals, companies, and international organizations.

The types of cyber-attacks used are varied, among which we can mention ransomware, phishing, DDoS, Cross-Site Scripting, among others, each of which has its form of operation and impact. During the COVID-19 context, there was a boost in the use of technology solutions due to the selection of teleworking and Tele-education, which indirectly broadened the attack surface, generating more remarkable growth in the number of cyber-attacks and variability in their form of operation (Andrade et al., 2020). Ransomware attacks focused on hospitals, while phishing attacks targeted organizations and individuals by falsely sending fake work situations or pandemic data.

Resources used to conduct cyberattacks include misconfigured websites, outdated operating systems, IoT devices with default credentials, or DNS systems. While some cyberattacks exploit vulnerabilities in technology systems, others focus on finding human vulnerabilities. Attackers found in social engineering attacks a field of action where they do not need to invest efforts to break algorithms or complex security infrastructures but rather take advantage of people's naivety, lack of knowledge, beliefs, or emotions.

Social Engineering attacks are of various types: phishing, ransomware, spear phishing, vishing, and smishing. Nevertheless, in context, all these attacks focus on persuading the human being to perform a specific action such as giving personal information, accessing a web page, downloading an attached document, or forwarding certain content to their acquaintances (Benavides et al., 2020). Figure 2 illustrates a mind map representing the major types of Social Engineering attacks, their causes, targets, and countermeasures.

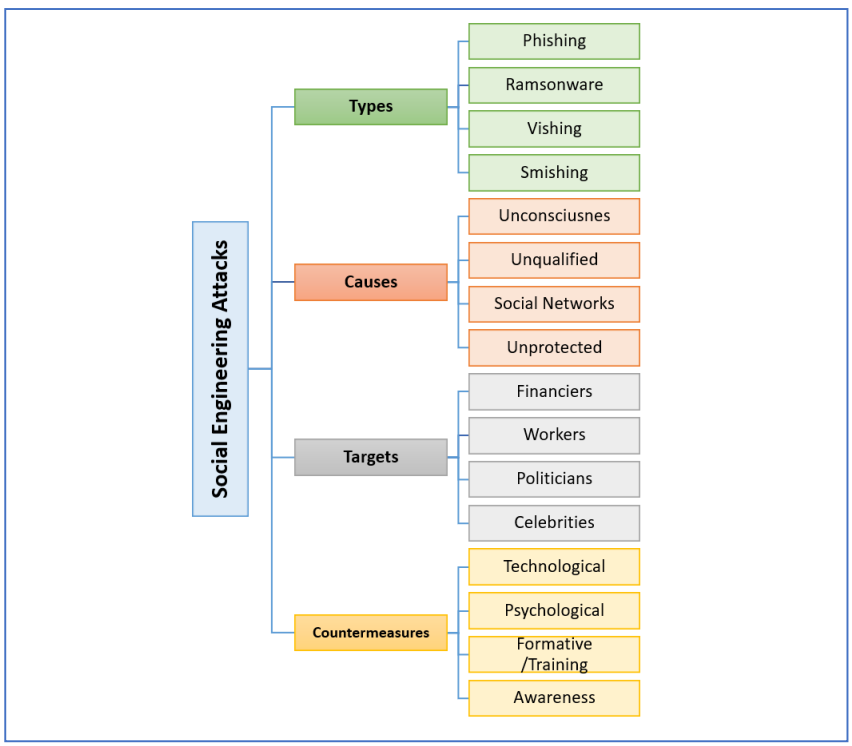

Fig. 2. Mind Map concerning the main aspects of Social Engineering Attacks. (Source: own illustration)

On the other hand, the implementation of countermeasures to minimize the impact has been addressed by international organizations such as the FBI, INTERPOL, EUROPOL, and different companies worldwide. The development of artificial intelligence-based security solutions for the early detection of cybersecurity threats has become the focus of specialized companies and cybersecurity research groups. Awareness campaigns for users about the seriousness of sharing sensitive information when receiving suspicious messages have also been considered strategies (*INTERPOL Report Shows Alarming Rate of Cyberattacks during COVID-19*, s. f.).

PHISHING ATTACKS

Phishing is a malicious technique in which attackers use social engineering mechanisms in an automated fashion with the intention of tricking the victim into revealing confidential data that can later be used to impersonate the victim on a website or in a financial transaction. It is seen as a vital problem that we face on a daily basis. This can become a huge source of risk due to the theft of personal information such as passwords, bank identifiers for criminal use etc.

Nowadays the importance of knowing what phishing is or how to protect yourself from this attack has become indispensable since suffering a deception through a simulated website (banks, administrations, Internet service providers, etc.) is something that anyone could suffer, given that web pages are increasingly more similar to legitimate ones, have fewer errors and the links have few differences to the original. Sometimes even pop-up windows appear where no web address is visible.

The targets of this scam have been evolving. In the beginning they were only looking for bank details, nowadays they are looking for all kinds of information. On the other hand, as phishing and its techniques advance, the measures to combat it are also increasing.

This is why there are several methods to identify phishing (which may or may not work depending on the complexity with which it is made), thanks to advanced methods of cyber-intelligence phishing attack is detected quickly and respond so that the cloned pages are closed, thus controlling the exposure of information, managing to avoid theft of millions of dollars or simply to protect personal data.

But just as these methods develop rapidly, phishing presents new evolutions, having the sending of personal messages and this makes them more credible to the user. Simultaneously, variants of this scam are appearing, for example, the "wishing" or "whale phishing" that is directed to people or small groups that follow a determined criterion (as high government officials).

On the other hand, phishing also uses other forms to spread spyware (spyware) and malware (viruses, Trojans...) such as "hishing" or the latest to be discovered "blow pish" which is a combination of cryptography with computer viruses or malicious code.

The impact of this type of crime is so risky given that the losses are much larger than estimated or expected to exist and this in just clicking on a fake email but this attack has both economic and social impacts and this produces a potential concern. The economic impact can be related to either an individual or a company, while the social impact is on companies because their customers lose confidence in the company.

A. Economic Impact

It is important to know that, since we are talking about a relatively new crime, there are few statistics on its impact, it can be said that, for each successful fraud, the worldwide economic loss is around 1.5 million dollars, while for other attacks the losses have an average of 2400 dollars. This leads us to the conclusion that although virus attacks are more damaging in monetary terms, the economic impact of phishing is not negligible.

B. Social Impact

Especially phishing affects the confidence of users to carry out operations on the Internet, as in the case of bank transfers or transactions. This crime presents a growing phenomenon, as much or as much as technology, is for this reason that the confidence of those interested in using the Internet to perform operations is decreasing. And this is produced by the fear of being a victim of phishing and have a significant monetary loss and this is evident in the platforms of buying and selling online. Producing that more and more people have less reliability to the security systems in the network, and giving us as a result a new impact that is psychological.

C. Psychological Impact

When thinking about this impact we could not give an exact definition since not all people go through the same feelings, but the fear of suffering a scam or fall into the theft of information by seeking a quick solution on the internet and lack of expertis in recognizing certain patterns present in this type of crime makes users present doubts, opting to make any transaction / operation in person. This option was not so bad in the past, but nowadays it is risking your health or your money facing people to a level of stress that exceeds the levels to which we are accustomed daily so what to do how to deal with fear, stress, uncertainty, anxiety and some other feelings that are experienced giving us in conclusion that phishing affects in so many ways, and therefore there is a brake on the evolution of the digital economy.

3. Variables of Human vulnerability in Phishing attacks

Social Engineering attacks continue to be effective. During the pandemic, it has become evident how large technology companies, banks, and organizations of all types continue to be victims of these attacks. This context generates our research questions that guide the construction of this study.

1. What makes social engineering attacks so effective?
2. Why do attackers select this type of attack?
3. How can individuals and companies deal with this problem?

To address these three research questions, we start from the premise that Social Engineering attacks have a different connotation than other cybersecurity attacks because they focus on humans' vulnerabilities. For instance, a study presented by Chang (Chang & Chong, 2010) identified that email frauds are based on the influence of psychological factors through the attackers' use of impersonation of people or organizations, the influence of a sense of authority or urgency, and the provision of a certain level of legitimacy. Social engineering attacks can manage language that shows authority for people and present a formal appearance that appears to be from a chief executive officer (CEO) or chief information officer (CIO). In contrast, others seek to generate a sense of a limited opportunity, such as travel offers (Butavicius et al., 2016). During the pandemic context, attackers adapted to this context, and attacks started to use COVID-19 related topics such as the number of infected, possible treatments, vaccines, and medical equipment.

We can observe that attackers seek to adapt their attack to people's reality to persuade them to act. Jones et al. (Jones et al., 2015) mention three aspects related to psychological influences in the email-making-decision process. First, the persuasiveness of email messages. Second, the cognitive process to judge the legitimacy of the email. Third, the theoretical influence is based on individual differences. People's influence is linked to cultural, ideological, ethical beliefs, social values and biases, and available information factors that determine the mental filters used for information processing (Chai, 2020). During our study, we can discern that there are

two macro factors in the decision-making process: (a) The establishment of mental filters for information processing, and (b) The socio-cultural factors that are related to the generation of mental filters.

In another study, Jones et al. (Jones et al., 2019) indicates that the same individual could judge the email with different results. Factors such as time pressure, the confidence level of an individual, and the detail of an email's content could affect the user decision. Jones mentions that participants in a study about phishing could identify better when the phishing content includes well-known companies. Impulsivity and higher sensation seeking were factors related to persons that commit errors for detect phishing emails.

Lawson et al. (Lawson et al., 2017) mentioned that there is a relationship between personality and persuasion. The experiment proposal by Lawson measures personality measures related to control, trust, neuroticism, extroversion, openness, agreeableness, and conscientiousness and identifies that high extroversion is predictive of increased susceptibility to phishing attacks. Haveli (Halevi et al., 2015) mentions that conscientiousness personality trait is more correlated with phishing response.

The use of influence techniques is intended to make victims use mostly heuristic processing versus systemic thinking. For this goal, attackers use aspects such as a sense of urgency. When individuals using heuristic thinking, they tend not to focus on details (Vishwanath et al., 2011).

The work proposed by Willians (Williams & Joinson, 2020) states that the person with the most confidence in their ability to identify phishing may make mistakes in identifying because they tend to omit details.

Age has been another aspect considered by researchers. Sarro (Sarno et al., 2020) specifies that there is no significant difference between ages in the identification of phishing. However, there is a difference in time-related to the categorization process. Additionally, adults tend to be more cautious and take more time to analyze a message. On the other hand, Tian et al.(Lin et al., 2019) mention that young people tend to be less susceptible to phishing attacks. Also, Tian mentions that older women tend to present more susceptibility to phishing attacks. Grilli et al. (Grilli et al., 2020) mention that older people tend to worse discrimination between genuine and phishing emails according to perceived suspiciousness.

Bailey et al. (Bailey et al., 2020) consider that older adults may be more vulnerable than young adults to novice advice that may not be as beneficial as expert advice. This appears because older adults discriminate less than young adults between different sources (novice vs. expert) of advice. This may be attributed to how older adults with a reduced preference for autonomy and more deficient working. Ebner et al. (*Uncovering Susceptibility Risk to Online Deception in Aging | The Journals of Gerontology: Series B | Oxford Academic*, s. f.) remark that higher susceptibility was associated with lower short-term episodic memory in middle-old users and lower positive affect in young-old and middle-old users. Greater susceptibility awareness was associated with better verbal fluency in middle-old users and a more significant positive effect in young and middle-old users.

Related to the influence of gender in phishing judgment, Haveli (Halevi et al., 2015) identifies that women are more susceptible to spear-phishing attacks. Verkijika (Verkijika, 2019) mentions that self-efficacy tends to increase more in women than men in the same vein. Self-efficacy is useful for anti-phishing techniques. Finally, regarding the phishing frequency, Singh (Singh et al., 2019) specifies that people exposed to phishing mail improve t hit rate and increase false alerts.

The points covered in scientific research about human vulnerability variables that influenced judgment in phishing events can summarize in Table 1.

Table 1. Predisposing factors that facilitate predisposition to the probability of being a victim of Phishing.

Demographic factors	Personality factors	Environmental factors
Age	Impulsivity	Time pressure
Gender	Higher sensation seeking	Work context
	Extroversion	Urgency sense
	Susceptibility	
	Confidence	
	Conscientiousness	

4. Decision making in Phishing events

The predisposing factors have a direct relationship with the cognitive processes associated with the decision-making process from the literature reviewed. Attackers seek to influence the cognitive process that drives the selection of an incorrect or harmful decision on the phishing victim.

Decision-making has been investigated in several areas of knowledge, such as psychology, cognitive science, and neuroscience. The research proposal by Trueblood et al. (Trueblood et al., 2018) in the medical image review field mentions that people's decision-making is based on sensory information. These decisions will be affected by the amount of information accumulated over time. This accumulated information is related to neural activity in multiple cortical and subcortical brains. Additionally, Trueblood mentions that different factors can affect the decision-making process, such as time pressure, quantified by the amount of information needed to decide and prior expectations.

A relevant aspect to consider is that in the execution of repetitive tasks such as reviewing medical images or in the context of this study reviewing e-mails, information can be processed heuristically. Heuristic processing may be more efficient for frequent tasks and is a simplified form of thinking.

The identification of phishing attacks process requires the ability to classify and interpret the characteristics and information in e-mail and web-pages. Training and education on cybersecurity attacks and the diffusion of the risk associated with phishing attacks are developed inside organizations. However, people still fall into phishing attacks. In order to improve decision-making based on the identification of phishing attacks, it is critical to understand the cognitive processes involved in these decisions.

Figure 3 shows a macro representation of cognitive processes related to decision-making based on information processing. Individuals acquired information for sensory perception; then the brain correlated the new information with accumulated information in the brain; in this process, multiple cortical and subcortical brain areas are working. Finally, an individual makes a decision.

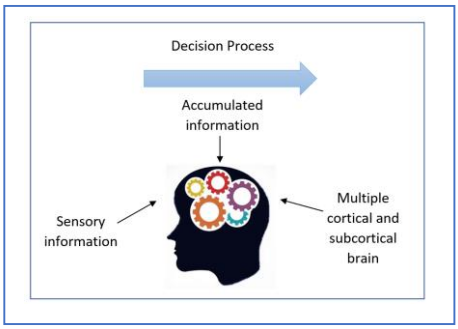

Fig. 3. Brain process for decision-making. (Source: own illustration)

In this decision-making process, maybe factors could be influenced or modified for an attacker to try people select phishing messages. For instance, an attacker could manipulate the information for trying people do not to perceive details in the sensory perception process. The attacker could send information that presents a conflict with previous information accumulated and bias in the decision process.

4.1 Mental Filters

Learning required the sensory elements that are based on the principles of causality are converted into knowledge. People's experiences are also converted into cognitive elements that later allow us to make decisions. Attackers can try to manipulate the cognitive process to carry out an attack. Based on the premise that mental models are closely related to social values, attackers may focus the content on raising a social issue of people who are in urgent need of money and appeal to the feelings of the victims to get a voluntary donation.

Under the same context, if the person establishes a mental filters to establish a slight suspicion that the email is malicious, there will be the possibility that the person executes the action suggested by the attacker if the person observes that several co-workers make the donation; this cognitive event is called "Cognitive Dissonance". Mental filters are build trhough of existing mental models based on human tendency to process information for making decision. For instance, the anchoring mental model is based on the human tendency to use the initial phishing of information (anchor) to make decision.

Table 2. Description of some mental models.

Mental Model	Description
Anchoring	Human tendency to use the initial piece of information to make decision.
Game theory	Decision is based on consequence of actions form a set of players
Illusion of control	Human tendency to overestimate their ability to *control* events
Incentive	Motivation to do something by promising to reward

Inversion principle	Thinking in opposite way to use it to resolve a problem
Loss aversion	Human feel that losses as more painful than equivalent gains
Margin of safety	Human tendency for dealing with the reality that if things can go wrong
Scarcity	Human tendency to want something the is limited
Signaling theory	Human tendency to broadcasting aspects of your identity, beliefs, personality, and lifestyle at all times, in all situations.

4.1.1 Cognitive dissonance

Cognitive dissonance can influence a person who has a decision considered correct to make an opposite decision influenced by third parties or tendencies. Social psychological theories mention that people's uncertainty about their decisions' correctness may lead them to compare with other people's opinions and abilities (Festinger, 1962). Dissonance is greater when it is related to self-image and affects feelings of foolishness, immorality, and others. Cognitive dissonance could be produced by the following aspects (*Password Cognitive Dissonance: America's Cyber-Security Problem | Information Security at UVA, U.Va.*, s. f.):

- Compliance behavior is a tendency to rationalize malice. For example, a user downloads files or movies from malicious sites, assuming it is normal and safe;
- Decision Making, people experience cognitive dissonance when they are in conflict with the decision they make. Cognitive dissonance can help to make better decisions by people when resolving the conflict;
- If a person receives an email notifying him that he has received an inheritance but decides to ignore it as false. In that case, he is likely to experience dissatisfaction with the decision and think about the possibility that he made a mistake and thinking that the message was valid;
- The effort, people experience cognitive dissonance when they select an alternative that they know is not correct, but that involves less effort. For instance, a person could select a password easy to remember, although it is not secure.

Cognitive dissonance influence decision-making in any one of three ways:

- Change behavior,
- Justify the behavior by changing the conflicting attitude,
- Alternatively, justify the behavior by adding new attitudes.

The study entitled "Lever-aging Behavioral Science to Mitigate Cyber Security Risk" presented by MITRE (n.d.) discusses aspects like Status Quo Bias or Framing Effect. These aspects could change their decision to influence internal heuristics.

Social and cultural perspective generates cognition. Any experience in social life can therefore modify our cognitive approach. This aspect is taken for attackers that try to modify the decision process. From the cybersecurity, perspective is needed to establish strategies for people to manage cognitive dissonance positively.

4.1.3 Bias

Bias describes a person's tendency to view something from a particular perspective. This perspective prevents the person from being objective and impartial. The following are a type of Bias:

- Optimism bias, individuals overestimate their abilities, and underestimate the risk of information security attacks, for instance, think that phishing attacks could be lower than others (Warkentin et al., 2013);
- Status quo bias describes people's tendency not to change an established behavior without a compelling incentive to do so;
- Fatalistic thinking, individuals perceive their risk to information security attacks, such as phishing, to be lower than others;
- The framing effect is a type of cognitive Bias related to how people react depending on the information presented if it represents a loss or gain. The framing effect is mainly related to automatic responses to stimuli, and it results from the intuitive and deliberative systems. Risk-averse people tend to use primarily intuitive systems. The intuitive system is responsible for fast processes that are effective, emotional, and automatic. Guo et al. (*Thinking Fast Increases Framing Effects in Risky Decision Making - Lisa Guo, Jennifer S. Trueblood, Adele Diederich, 2017*, s. f.) mention that time pressure might alter the attention process and result in behavior changes. In contrast, Fan (*Frontiers | Education and Decision-Making: An Experimental Study on the Framing Effect in China | Psychology*, s. f.) mentions that risk preferences and choice strategies are sensitive to a decision-maker's biological conditions. For instance, Guo mentions that women have greater sensitivity to negatively framed information than men.

Figure 4 shows a phishing message based on the criteria of gain. The message shows to people that they win something valuable with the goal to generate a bias on people and influence in the decision of select the message for not lose the opportunity of this reward.

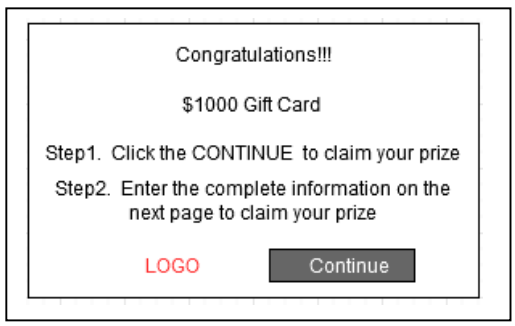

Fig. 4. Phishing message based on gain criteria.

4.2 Heuristic Process

Concerning the heuristic process, decision-making strategies can be developed with little information, a low level of detail, readily available information. The heuristic can be cognitive and practical.

According to Shah et al., heuristics rely on one or more of the following methods for effort-reduction

- Examining fewer cues;
- Reducing the difficulty associated with retrieving and storing cue values;
- Simplifying the weighting principles for cues;
- Integrating less information;
- Examining fewer alternatives.

4.3Future thoughts

Thinking about the future influences our behavior. Oettingen et al. (*The Motivating Function of Thinking about the Future*, s. f.) mention that beliefs and judgments and the predictive aspects of thinking about the future. Hershfield et al. (Hershfield et al., 2011) mention that taking time to simulate and enjoy a positive experience in advance can allow you to derive benefits for the experience twice.

Episodic future thinking is defined as the ability to project oneself into the future to pre-experience the future consequences of present actions. For instance, Segovia et al. (Segovia et al., 2020) take advantage of episodic future thinking to probe that for providing health-related information has a positive impact on the food choices. Figure 5 shows phishing email focus on generate a positive experience based on the need for people to know their intelligence level.

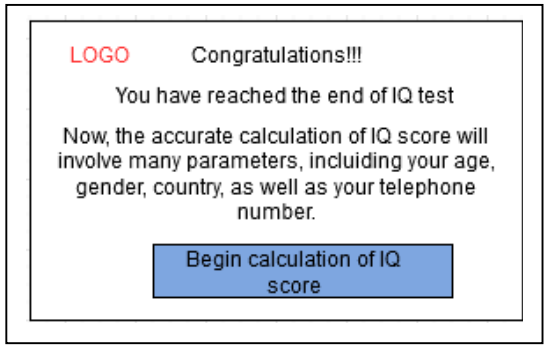

Fig. 5. Phishing message based on people's needs.

5. Case study (Phishing)

This section describes the case study used, which focuses on studying the cognitive process and human behavior victims of a Social Engineering attack.

5.1 Participants

We recruited both novice and systems professionals to complete the experiment, i.e., approximately seventy undergraduate students of both respectful universities near Quito and Sangolquí cities from Ecuador. Also, ten professionals from the security information specialized organizations participated. We targeted about equal numbers of "experienced" and "inexperienced" practitioners. We had 10 "experienced" and 70 "inexperienced" participants.

5.2 Materials

To create the stimuli, we development a bank of 20 digital images of phishing messages based on real phishing emails received in our institutional emails and put them in a web software. Additionality, we replied legitime email messages.

We classify the messages into one of the two types (phishing and legitime) based on their characteristics. Two cybersecurity professionals confirm the classification of each phishing.

5.3 Procedure

First, participants completed a training stage to familiarize themselves with phishing (i.e., both novices and experts completed the training for consistency). Then, the training focused on teaching participants to identify phishing. The training activities had no sequence. The phishing and legitimate emails were randomly chosen. Choice reaction time tasks that allow for human cognition.

5.4 Phishing experiment setup

It is critical to understand the cognitive processes underlying decision-making from phishing messages to improve training and minimize the occurrence of being the victim of the attack.

This study aims to use experimental methods and computational tools developed in the area of decision-making to probe the cognitive processes involved in phishing decisions in novices, experts, and people without a technical background.

To examine this process experimentally, we passively development a set of digital images of both phishing and non-phishing. A panel of expert cybersecurity classified each of these images, providing a fully annotated data set consisting of images of varying types and levels of difficulty. Using this image bank, we developed a perceptual decision-making experiment to investigate how time pressure and externally imposed bias influence individuals' behavior.

5.5 Data analysis

Figure 6 shows the results of the decision-making evaluation process in simulated phishing exercises. The cognitive tasks that are described as those used in each of the exercises that were presented are analysis, reading, doubting, verifying or reviewing.

of the exercises are analysis, reading, doubting, verifying, and reviewing.

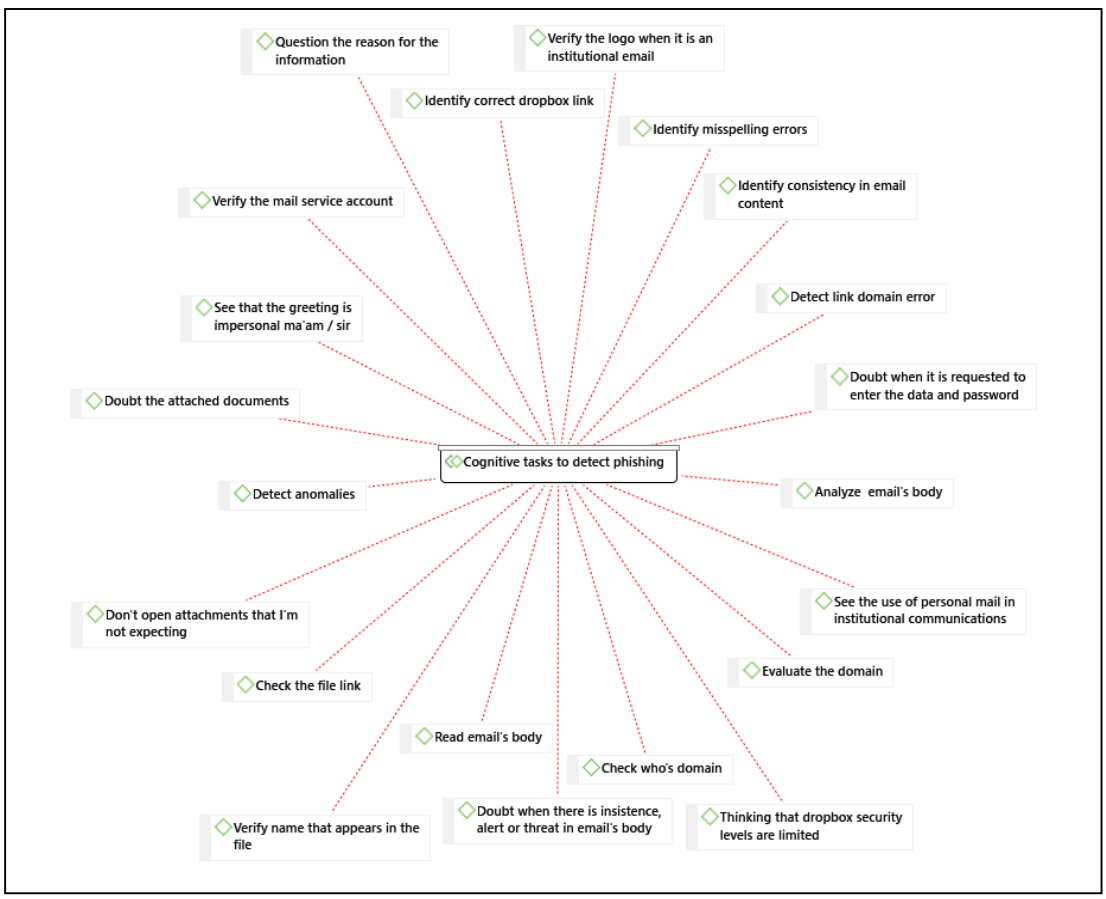

Fig. 6. Cognitive tasks to detect phishing. (Source: own illustration)

The analysis of the mistakes that inexperienced people can make when they have to detect phishing is shown

in Figure 7.

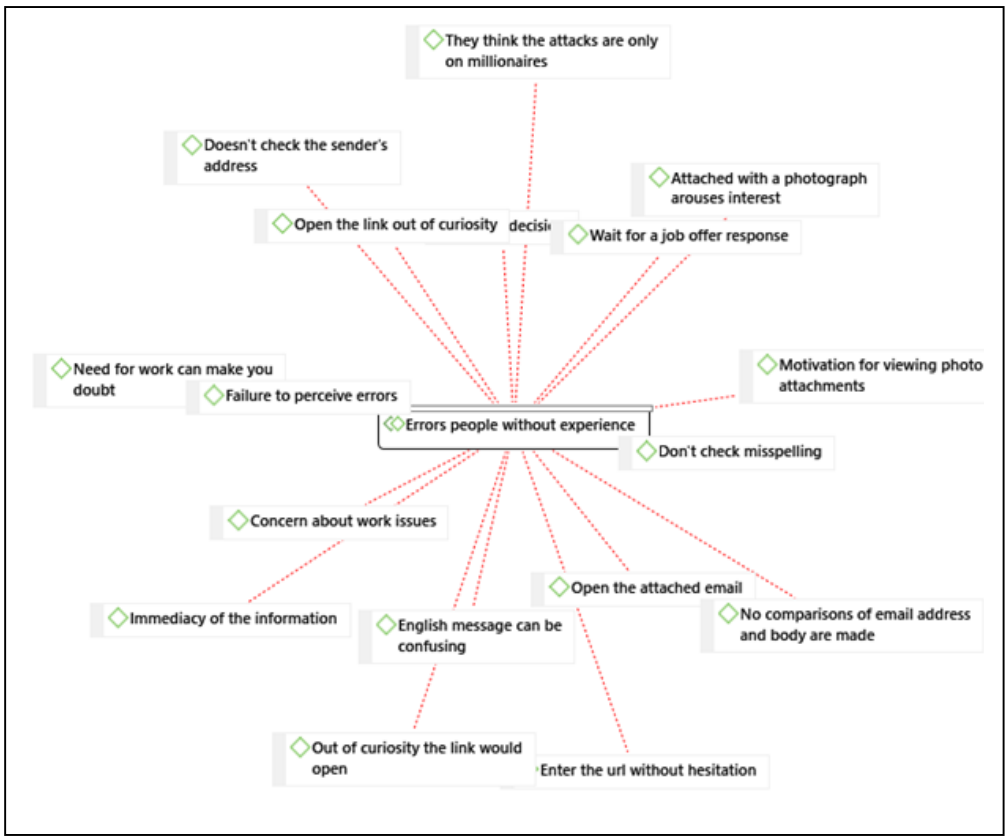

Fig. 7. Errors in detection phishing

The details and information that the participants evaluated in each email exercise are represented in Figure 8 , which shows the aspects of the cognitive assessment that influenced the selection of the two proposed alternatives (legitimate or phishing).

.

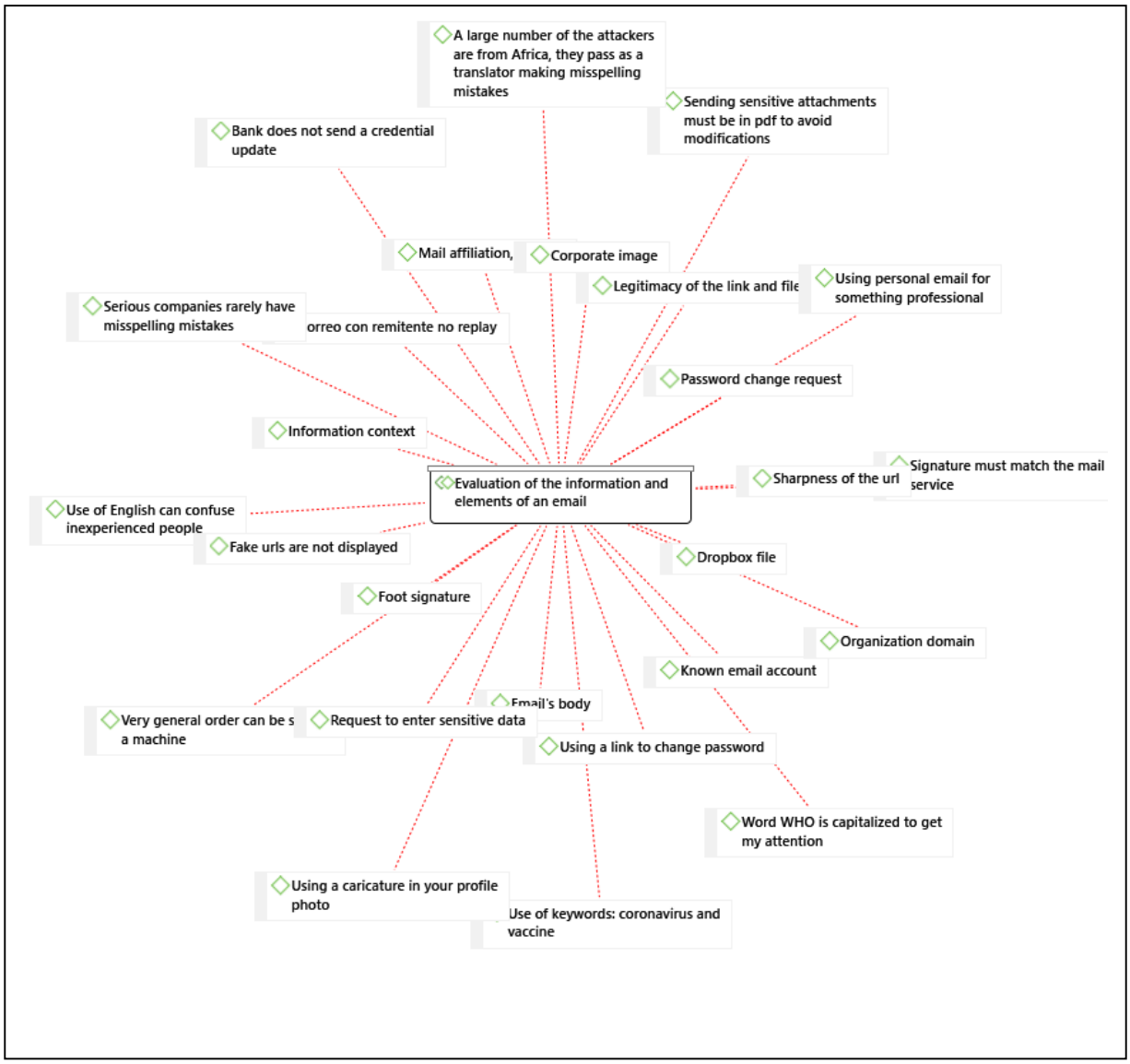

Fig. 8 . Cognitive evaluations of an email to detect phishing.

Table 3 shows the characteristics that were used most frequently in the phishing detection of the mail exercises shown on the platform and that guided the decision-making towards one of the two alternatives proposed.

Table3. Characteristics that were considered in making decisions

Phishing	**Freq.**	**Legitimate**	**Freq.**
Greeting and generic content	8	Correspondence with the domain	6
Persuasive content	7	Bank notification email	6
Attachment domain does not correspond to google drive	7	Correct URLs	4
Legitimacy of the URL, the link is not from google	6	No personal data is requested in the content	4
Use of personal mail for institutional communication	6	There is no link or attachment	2
There is no match of the mail domain with the attachment or link daemon	6	Good writing, no misspelling mistakes	2
Seeks to get information	6	Email accounts if applicable	2
Suspicious or altered link	5	Addressed mail with first name last name	1
The mail domain does not correspond to the name	5	Use https in the domain	1
Banks do not send links to recover keys	4	Domains and emails belong to google	1
Attachment domain is not dropbox	4	Attachment link is real	1
Google does not send links to change keys	3		
Grammatical mistakes, writing errors	3		
Link send to another site	2		
Sender of the mail has another identity	2		
Signature does not match the name of the mail sender	1		
Emails with attachments from unknown people	1		

6 Discussion

6.1 Characterizing phishing attacks

To understand how people execute decision-making process when need to evaluate if an email message is phishing or not. We development a web training platform based on a set of simulated email exercises built from real phishing and legitime emails. For the experiment we had three types of participants: Full-insight (professionals that are involve in cybersecurity fields with more than 5 years of experience), Partial-insight (students of the last levels of computer science degree) and No-insight (person that don`t have a technical background). Each participant had to decide between 10 exercises if they were phishing or legitime. In each exercise the participants provide us a feedback of the reason because they decide if the message were phishing or not.

From the web training platform to evaluate how people analyze if one message is phishing or not, we development two mental models. The first one show in the Figure 9 is focus on phishing. The mental model for decision of phishing defines a set of nodes that represent indicators that are used in the moment to executed the cognitive task to analyze message. For instance, people based their decision in the analyze of URL, if it had low legitimacy (domain for unknow source). Table 4 shows indicators and their features used for people to decide that email is phishing.

Table 4. Indicators used for people to decide if email is phishing attack.

Indicators	***features***
URL	Low legitimacy
Content	Grammatical mistakes
	Persuasive content
	Get personal information
Attachment	Domain not equal to URL
Link	Domain not equal to URL
Identity Sender	Suspicious

The second one show in the Figure 10 is focus on select legitime messages. The mental model for decision of legitime defines a set of nodes that represent indicators that are used in the moment to executed the cognitive task to analyze message. For instance, people based their decision in the analyze of URL, if it used https. Table 5 shows indicators and their features used for people to decide that email is legitime.

TABLE I. TABLE TYPE STYLES

Indicators	***features***
URL	High legitimacy

Indicators	***features***
	Use https
	From well-known organization
Content	Not grammatical mistakes
	No get personal information
Attachment	Real domain
Link	Not send in the email.

We can observe in Figures 9 and 10 that the number of indicators is more in phishing model that in legitime model. In contrast, the number of features in each indicator is more in legitime model that in phishing model. To establish the weight in each link between nodes in the both mental models were based on relative frequency analyze of the words used for 10 full-insight participants in the feedback process according to label of indicators and features show in the Table 4 and Table 5.

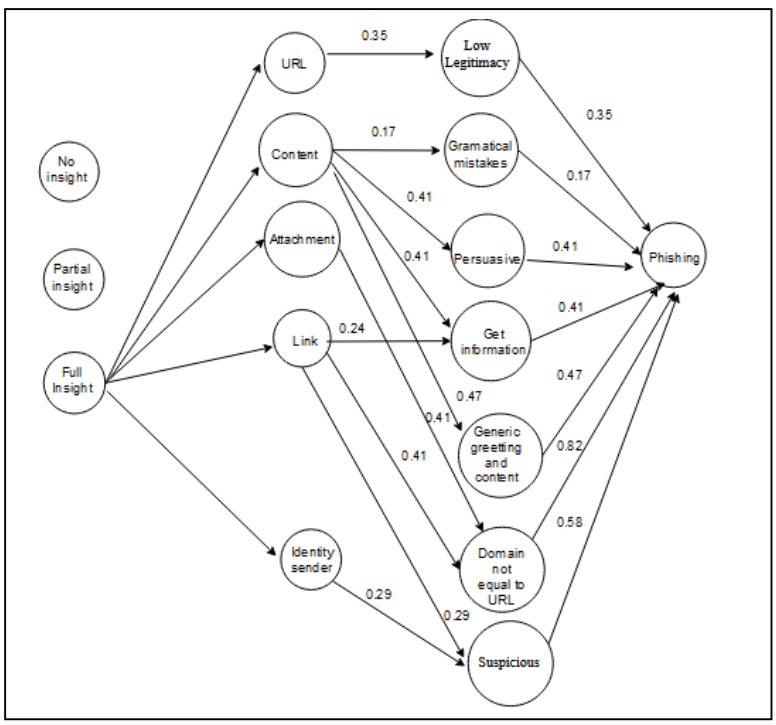

Fig. 9. Phishing message based on people's needs.

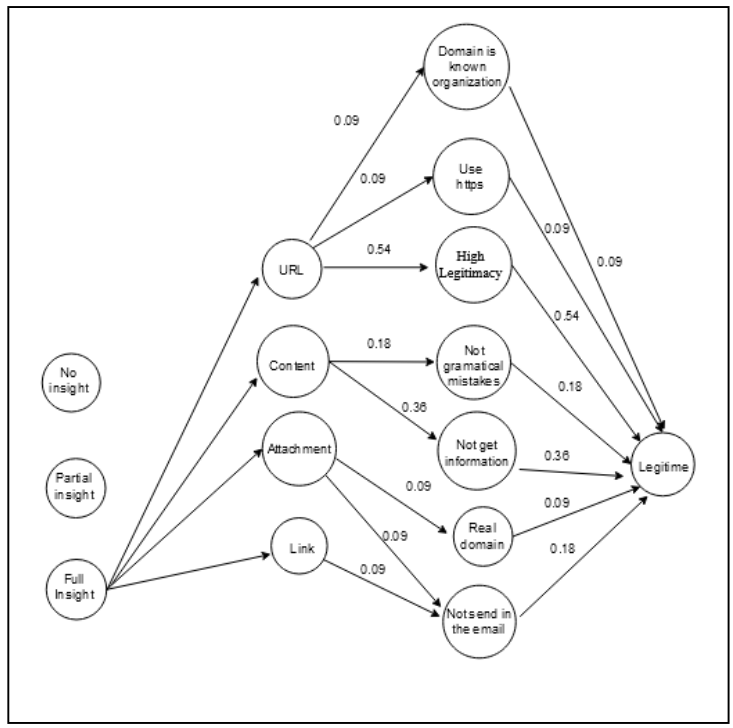

Fig.10. Phishing message based on people's needs.

6.2 User-centric Incident Response

From a user-centric cybersecurity defense perspective, an alternative would be to strengthen the human aspects that attackers use. However, modifying cultural, ideological, or social value factors might not be a feasible option due to different factors associated with time, ethics, and freedom of thought, among others beyond the scope of this paper. The next alternative is to focus on the mental filters for information processing, to understand how all these cognitive elements are developed and executed by the mind when faced with the decision to respond to a cybersecurity attack. It is essential to mention that these cognitive processes have a bias-related, as we have mentioned by the aspects raised by Parenti. For example, in specific religious contexts, a person who receives an email suggests a chain to be sent to ten of his contacts. Even if his critical judgment allows him to identify a possible phishing attack, his religious beliefs could be more robust and push him to resend the message.

As we can observe in a macro way, cybersecurity defense processes concerning social engineering attacks require the implementation of technological solutions and depend on the actions performed by people.

Understanding the human aspects of cybersecurity attacks will enable the development of training programs that focus on the technical aspects and seek to generate more effective decision-making processes that consider the influence of the cognitive process.

Cognitive dissonance is central to many forms of persuasion to change beliefs, values, attitudes, and behaviors. To get users to change their cyber behavior, we can first change their attitudes about cybersecurity. For example, a system could emphasize a user's sense of foolishness concerning the cyber risks he is taking, enabling dissonant tension to be injected suddenly or allowed to build up over time. Then, the system can offer the user ways to relieve the tension by changing his behavior.

Under this context, people's mental filters may be more focused on establishing ways to help than on possibly identifying whether the message is legitimate.

The learning process can be limited if it does not include experiences that allow the cognitive elements to be strengthened. Improving perception allows users to strengthen the process of abstraction of content.

The learning process could focus on that person to understand the problem of security attacks. They base their cognitive process on their experiential experience. At this point, people should not always face phishing situations because they would be exposed to risks that can affect them economically or personally. It is here where simulation scenarios can contribute to generating this experiential experience that generates the cognitive.

We can observe the mental model use for decide if the message was legitime or phishing is different. Additionality, the main elements evaluated were different for phishing and legitime. In the case of phishing participants analyze in more detail the content. In contrast for decide if the email was legitime participants decide to analyze in more detail the URL.

Generally, human tendency for some persons could be consider that any email that they receive is legitime and consider the evaluation of mental model focus more in analyze the URL. So, if the URL of domain for sender uses https and it is for one known organization, participant could decide the email is good. In recent days, many people fall in phishing attacks because attackers used Microsoft email server to execute the attacks.

7 Conclusions and future work

The purpose of this book's chapter was to determine the cognitive factors associated with social engineering, especially during a phishing attack. We conducted a literature review. Then, a case study was carried out for understanding the behavior of potential victims. Using a simulation through a web platform, it was possible to understand the cognitive and behavioral process of users to recognize legitimate and phishing-infected emails. After the data analysis, we obtained a cognitive model of users against phishing attacks and legitimate emails.

As future work, we plan to create a training web platform that includes the proposed cognitive model to minimize social engineering attacks from the enterprise, the industry, and the family.

References

Algarni, A., Xu, Y., & Chan, T. (2014). Social Engineering in Social Networking Sites: The Art of Impersonation. *2014 IEEE International Conference on Services Computing*, 797-804.

https://doi.org/10.1109/SCC.2014.108

Andrade, R. O., Ortiz-Garcés, I., & Cazares, M. (2020). Cybersecurity Attacks on Smart Home During Covid-19 Pandemic. *2020 Fourth World Conference on Smart Trends in Systems, Security and Sustainability (WorldS4)*, 398-404. https://doi.org/10.1109/WorldS450073.2020.9210363

Bailey, P. E., Ebner, N. C., Moustafa, A. A., Phillips, J. R., Leon, T., & Weidemann, G. (2020). The weight of advice in older age. *Decision*. https://doi.org/10.1037/dec0000138

Benavides, E., Fuertes, W., & Sanchez, S. (2020). Caracterización de los ataques de phishing y técnicas para mitigarlos. Ataques: Una revisión sistemática de la literatura. *Ciencia y Tecnología*, *13*(1), 97-104.

Brown, S. D., & Reavey, P. (2017). False memories and real epistemic problems. *Culture & Psychology*, *23*(2), 171-185. https://doi.org/10.1177/1354067X17695764

Chai, S. (2020). Does Cultural Difference Matter on Social Media? An Examination of the Ethical Culture and Information Privacy Concerns. *Sustainability*, *12*(19), 8286. https://doi.org/10.3390/su12198286

Chang, J. J. S., & Chong, M. D. (2010). Psychological influences in e- mail fraud. *Journal of Financial Crime*, *17*(3), 337-350. https://doi.org/10.1108/13590791011056309

Chapter One—Risks Landscape. (s. f.). *Global Risks Report 2020*. Recuperado 8 de abril de 2021, de https://wef.ch/2RowGnF

Cole, S., & Kvavilashvili, L. (2021). Spontaneous and deliberate future thinking: A dual process account. *Psychological Research*, *85*(2), 464-479. https://doi.org/10.1007/s00426-019-01262-7

Festinger, L. (1962). Cognitive Dissonance. *Scientific American*, *207*(4), 93-106.

Frontiers | Education and Decision-Making: An Experimental Study on the Framing Effect in China | Psychology. (s. f.). Recuperado 8 de abril de 2021, de https://www.frontiersin.org/articles/10.3389/fpsyg.2017.00744/full

Grilli, M. D., McVeigh, K. S., Hakim, Z. M., Wank, A. A., Getz, S. J., Levin, B. E., Ebner, N. C., & Wilson, R. C. (2020). Is This Phishing? Older Age Is Associated With Greater Difficulty Discriminating

Between Safe and Malicious Emails. *The Journals of Gerontology: Series B*, *gbaa228*. https://doi.org/10.1093/geronb/gbaa228

Halevi, T., Memon, N., & Nov, O. (2015). *Spear-Phishing in the Wild: A Real-World Study of Personality, Phishing Self-Efficacy and Vulnerability to Spear-Phishing Attacks* (SSRN Scholarly Paper ID 2544742). Social Science Research Network. https://doi.org/10.2139/ssrn.2544742

Hershfield, H. E., Goldstein, D. G., Sharpe, W. F., Fox, J., Yeykelis, L., Carstensen, L. L., & Bailenson, J. N. (2011). Increasing Saving Behavior Through Age-Progressed Renderings of the Future Self. *Journal of Marketing Research*, *48*(SPL), S23-S37. https://doi.org/10.1509/jmkr.48.SPL.S23

January 7, E. C., & Minutes, 2014 10. (s. f.). *Why Training Doesn't Mitigate Phishing*. Recuperado 8 de abril de 2021, de https://www.bankinfosecurity.com/interviews/training-doesnt-mitigate-phishing-i-2148

Jones, H. S., Towse, J. N., & Race, N. (2015). Susceptibility to Email Fraud: A Review of Psychological Perspectives, Data-Collection Methods, and Ethical Considerations. *International Journal of Cyber Behavior, Psychology and Learning (IJCBPL)*, *5*(3), 13-29. https://doi.org/10.4018/IJCBPL.2015070102

Jones, H. S., Towse, J. N., Race, N., & Harrison, T. (2019). Email fraud: The search for psychological predictors of susceptibility. *PLOS ONE*, *14*(1), e0209684. https://doi.org/10.1371/journal.pone.0209684

Joshi, C., Aliaga, J. R., & Insua, D. R. (2021). Insider Threat Modeling: An Adversarial Risk Analysis Approach. *IEEE Transactions on Information Forensics and Security*, *16*, 1131-1142. https://doi.org/10.1109/TIFS.2020.3029898

Lawson, P., Zielinska, O., Pearson, C., & Mayhorn, C. B. (2017). Interaction of Personality and Persuasion Tactics in Email Phishing Attacks. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *61*(1), 1331-1333. https://doi.org/10.1177/1541931213601815

Lin, T., Capecci, D. E., Ellis, D. M., Rocha, H. A., Dommaraju, S., Oliveira, D. S., & Ebner, N. C. (2019).

Susceptibility to Spear-Phishing Emails: Effects of Internet User Demographics and Email Content. *ACM Transactions on Computer-Human Interaction*, *26*(5), 32:1-32:28. https://doi.org/10.1145/3336141

Password Cognitive Dissonance: America's Cyber-Security Problem | Information Security at UVA, U.Va. (s. f.). Recuperado 8 de abril de 2021, de https://security.virginia.edu/password-cognitive-dissonance

Sarno, D. M., Lewis, J. E., Bohil, C. J., & Neider, M. B. (2020). Which Phish Is on the Hook? Phishing Vulnerability for Older Versus Younger Adults. *Human Factors*, *62*(5), 704-717. https://doi.org/10.1177/0018720819855570

Segovia, M. S., Palma, M. A., & Nayga, R. M. (2020). Can episodic future thinking affect food choices? *Journal of Economic Behavior & Organization*, *177*, 371-389. https://doi.org/10.1016/j.jebo.2020.06.019

Singh, K., Aggarwal, P., Rajivan, P., & Gonzalez, C. (2019). Training to Detect Phishing Emails: Effects of the Frequency of Experienced Phishing Emails. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *63*(1), 453-457. https://doi.org/10.1177/1071181319631355

The motivating function of thinking about the future: Expectations versus fantasies. - PsycNET. (s. f.). Recuperado 8 de abril de 2021, de /doiLanding?doi=10.1037%2F0022-3514.83.5.1198

Thinking Fast Increases Framing Effects in Risky Decision Making—Lisa Guo, Jennifer S. Trueblood, Adele Diederich, 2017. (s. f.). Recuperado 8 de abril de 2021, de https://journals.sagepub.com/doi/full/10.1177/0956797616689092

Trueblood, J. S., Holmes, W. R., Seegmiller, A. C., Douds, J., Compton, M., Szentirmai, E., Woodruff, M., Huang, W., Stratton, C., & Eichbaum, Q. (2018). The impact of speed and bias on the cognitive processes of experts and novices in medical image decision-making. *Cognitive Research: Principles and Implications*, *3*(1), 28. https://doi.org/10.1186/s41235-018-0119-2

Uncovering Susceptibility Risk to Online Deception in Aging | The Journals of Gerontology: Series B | Oxford

Academic. (s. f.). Recuperado 8 de abril de 2021, de https://academic.oup.com/psychsocgerontology/article/75/3/522/4969910?login=true

Valaskivi, K. (s. f.). *Beyond Fake News: Content confusion and understanding the dynamics of the contemporary media environment*. 8.

Verkijika, S. F. (2019). "If you know what to do, will you take action to avoid mobile phishing attacks": Self-efficacy, anticipated regret, and gender. *Computers in Human Behavior*, *101*, 286-296. https://doi.org/10.1016/j.chb.2019.07.034

Vishwanath, A., Herath, T., Chen, R., Wang, J., & Rao, H. R. (2011). Why do people get phished? Testing individual differences in phishing vulnerability within an integrated, information processing model. *Decision Support Systems*, *51*(3), 576-586. https://doi.org/10.1016/j.dss.2011.03.002

Warkentin, M., Xu, Z., & Mutchler, L. (2013). *I'm Safer than You: The Role of Optimism Bias in Personal IT Risk Assessments*.

Williams, E. J., & Joinson, A. N. (2020). Developing a measure of information seeking about phishing. *Journal of Cybersecurity*, *6*(tyaa001). https://doi.org/10.1093/cybsec/tyaa001

Declaration of interests

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☒The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: