The Human Attack Surface Framework for Phishing

Vahdad, Alireza

Award date:
2020

The Human Attack Surface Framework for Phishing

Alireza Vahdad
Eindhoven University of Technology
Email: a.vahdad@student.tue.nl
Supervisors: Luca Allodi
Nicola Zannone
Collaborators: Pavlo Burda
Simone Pirocca

***Abstract*—Vulnerability to phishing has been studied since the early days of phishing, yet it is still ranked as one of the most effective attack vectors. We provide a structured framework for human attack surface for phishing through an extensive literature review of the effective factors. We conduct an interactive experiment that uses a novel way of real-time customization to exemplify the framework's usage. The result showed the framework's effectiveness, with 68% of the participants preferring the text personalized by using it.**

I. Introduction

Humans are the security's weakest link. Based on Ponemon's latest Study [1], employee mistakes were the most significant threats to the exposure of sensitive data. Due to various improvements and enhancements in technical defensive technologies, such as intrusion detection and prevention systems, directly attacking an organization can be harder than the early days of the Internet [32]. However, human nature remained unchanged.

Social engineering, and phishing in particular, are still among the most effective methods. Verizon has announced phishing as the most successful method in causing data breaches in their 2019 investigation [112]. That shows the potential of phishing, especially when considering their continuation of becoming more targeted and sophisticated, where attackers opt to use more channels, like text messages, postal services, etc. [92]. Based on another recent study by Proofpoint [2], in 2019, more than 55 percent of the respondent organizations had experienced a successful phishing attack, and, in another interesting finding, 39 percent of the global average of the working adult respondents were not able to choose the correct definition of phishing in a multiple-choice question. One of the best exemplars of the importance of the subject can be the recent hack of Twitter's internal systems in which 130 high-profile accounts were compromised caused by successful social engineering attacks on particular Twitter's employees [107]. Regardless of the statistics, it should be accentuated that a successful attack may only need one employee to fall for it [80].

In comparison to phishing, spear-phishing attacks are highly targeted by using the targets' personal information, which makes them more authentic and the users more vulnerable [46], emphasizing the principal role of the customization. Widespread usage of social networking sites, blogs, forums, and other similar online tools and applications, combined with the amount of shared personal information, have turned them into precious resources of Open Source Intelligence (OSINT) [90] and accessible means to deceive users [43]. This can lead to more targeted and better-crafted spear-phishing attacks that, in the end, will result in a higher chance of users succumbing to them. While social network sites are inseparable parts of the people's lives and their popularity increases day by day, they are mostly considered out of bounds in the contracts for the companies' penetration tests due to many factors, including being invasive to the employees' privacy [30].

In this article, we aim to develop a structured framework for the human attack surface for phishing by focusing on the collection and composition of OSINT, leading to the design of an effective phishing attack. Our contributions are as follows:

- We perform an extensive literature review to identify the phishing's effective factors from the extant literature;
- We develop the human attack surface framework for phishing;
- We provide model measurement strategies for each variable of the framework using open source intelligence;
- To exemplify a use-case of the presented framework, and to find out the effect of message customization using the framework, we conduct a real-time experiment where the content of an attack is automatically tailored to the participant's demographic characteristics and personality.
- For the first time, we showcase a method for automated personalization of phishing attacks, which allows the researchers and specialists in the field to conduct more real-life like experiments for spear-phishing.

The rest of the paper is organized as follows. Section II provides background knowledge and related works regarding the development of the phishing attack surface. Section III identifies the effective factors in phishing vulnerability and introduces our developed framework of the human attack surface for phishing. Section IV presents our methodology to show a use-case for the framework. Section V evaluates the results both quantitatively and qualitatively. Section VI presents and discusses our findings. We conclude in section VII.

II. Background and related works

Lastdrager [67] defines phishing as "a scalable act of deception whereby impersonation is used to obtain information from a target".

Due to the phishing's deceptive nature, other related fields to the human study are also of importance. One of the most used concepts in phishing subject is Cialdini's principles of influence [19], that discuss humans' cognitive vulnerabilities, namely Authority, Consistency, Liking, Reciprocity, Scarcity, Social Proof. These principles can have a significant impact on the human's decision making procedure to result in a more favorable result, which their definition can be seen in Table I.

Another influential subject in phishing vulnerability is the difference between heuristic and systematic processing and decision making. Heuristic processing is defined by Eagly and Chaiken [29] as "a limited mode of information processing that requires less cognitive effort and fewer cognitive resources" and is also referred to as "rule of thumb". In this mode, users fail to pay attention to the deception cues and act based on their intuition [114], [117]. Whereas, in systematic mode, users thoroughly examine and investigate the message using their cognitive resources to validate it [36]. Therefore, users are more vulnerable to phishing attempts in heuristic mode and more likely to detect the deception cues in the Systematic mode.

Regarding the effective factors in phishing vulnerability, Mundie et al. [80] categorize the contributing factors to unintentional insider threat, which involves social engineering, into demographic, organizational, and human factors. Their work mostly includes the factors that are effective in increasing the likelihood of the social engineering attacks in a general manner, leaving the other essential factors in phishing attacks, such as pretext and implementation part, out of consideration. Their work also demonstrates the importance of looking at the subject from both personal and professional points of view by dedicating a specific category to the organizational factors.

The study by Darwish et al. [23] encompasses a limited set of variables for profiling the vulnerable users to the phishing attacks that are mostly demographic related. The developed framework by Alseadoon [10] also consists of a limited number of factors that are mainly focused on the personality-related aspects of the target. Furthermore, the social engineering framework by Uebelacker and Quiel [108] is restricted to and focuses on the Big-Five personality of the victim and their mapping to the Cialdini's principles of influence [19]. Correspondingly, Albladi and Weir [4] approach includes socio-psychological, perceptual, habitual, and socio-emotional variables. While their set of variables is more comprehensive than the previous works, other essential factors of phishing attacks are still missing from the framework.

Parrish et al. [88] has developed a framework, specific to the phishing susceptibility, where personal and experiential factors and Big-Five personality profile have been considered. While their approach considers the importance of the implementation part of the phishing attacks, the mentioned factors are too specific by only mentioning the types of lure and hook. In contrast, our approach examines the effective variables from a broader point of view.

TABLE I
Cialdini's principles of influence

Principle	Definition [19], [111]
Authority	Tendency to obey people in authoritative positions, following from the possibility of punishment for not complying with the authoritative requests.
Consistency	Tendency to behave in a way consistent with past decisions and behaviours. After committing to a certain view, company or product, people will act in accordance with those commitments.
Liking	Preference for saying "yes" to the requests of people they know and like. People are programmed to like others who like them back and who are similar to them.
Reciprocity	Tendency to feel obliged to repay favours from others. "I do something for you, you do something for me."
Scarcity	Tendency to assign more value to items and opportunities when their availability is limited, not to waste the opportunity.
Social Proof	Tendency to reference the behaviour of others, by using the majority behaviour to guide their own actions.

Our approach also emphasizes the importance of considering the users in both their professional roles and personal lives with more comprehensive sets of variables. Moreover, as our framework focuses on the phishing vulnerability, it also considers other crucial factors that make a phishing attempt successful.

Some studies have tackled the issue from a different angle by scrutinizing the brain's physiology and decision-making procedure when facing phishing messages. Researchers in this field claim that this can lead to better user-centered security-related measures, such as education and training. Valecha et al. [110] show that neural responses within specific areas of the brain can demonstrate a person's misidentification of a phishing email and can help to understand the person's phishing susceptibility. In another study Neupane et al. [82] have evaluated the performance of the respondents and measured their neural activities while doing security-related tasks, which led them to find a connection between the user's behavioral performance and the corresponding brain activity.

III. Human attack surface for phishing

In this section, by concentrating on the human attack surface, and through an extensive literature study, the effective factors in phishing vulnerability are identified and further described. Subsequently, by merging these factors, a structured framework for the human attack surface for phishing is introduced, explained, and its usage is discussed.

A. Effective variables in the likelihood

Table II provides an overview of the variables that can affect the likelihood of success in phishing attacks and their descriptions. Additionally, further discussion of the variables is presented in Section A-A in the Appendix.

TABLE II: **Effective variables in phishing success likelihood**

Variable	**Description**
Gender	Females have shown higher susceptibility to phishing [100], [57], [70], [86], [101]. Having greater Internet anxiety [105], less positive attitudes towards the Internet [105], low security self-efficacy [13], and less technical knowledge and training [57] than males are among the mentioned reasons. Another study [95], thinks of gender as a proxy for other effective variables rather than being the true reason for phishing vulnerability by itself.
Age	Younger adults, between 18 to 25 years old, showed a high vulnerability rate [57], [85]. This can be influenced by their lower level of education, fewer years on the Internet, less exposure to training materials, and lower risk aversion. [100] Meanwhile, general cognitive processing capacities and sensitivity to deception and untrustworthy information decline with age, whereas perceived trust increases. [86] Therefore, adults older than 65 years old are among the susceptible groups as well [70], [95],
Level of Internet usage	Higher Internet usage results in a more user exposure to crime or negative experience, and therefore higher threat perception [85], and corresponding risks awareness [45]. Although this variable considers the Internet in a broader view, this is also true for each platform. Users with higher activity levels and more elapsed time since their membership on a platform were more successful in identifying spams or phishing attempts on that specific platform [95], [7], [6]. Moreover, more frequent online shoppers better detected phishing websites [94], and participants who had higher computer usage in general, significantly performed better in phishing susceptibility tests [89], [56].
Education	Education has been proved to have a positive relationship with Internet skills [47], information security awareness [85], lower preference for clickbait [70], and being one of the causes of younger adults' susceptibility to phishing [100].
Computer security literacy	Phishing awareness strikingly influences the user's phishing detection ability [94] and their perceived protective practices and reactions [48]. Also, in different studies, less knowledgeable users about phishing were more vulnerable to it [80], [7], [6], [28].
Previous victimization	Consists of the user's experience of being phished and encountering phishing attempts where a positive relationship has been found between these experiences and the capability of phishing website identification [94], higher overall awareness, and higher risk perception [60], [120].
Information Security Awareness	Information security awareness concerns "individuals' knowledge of what policies and procedures they should follow, their understanding of why they should adhere to them (their attitude) and what they actually do (their behavior)" [76]. Higher information security awareness causes users to judge based on the systematic, deliberate processing rather than heuristic mode [48], have a higher detection rate of spam emails [17], and less inclination towards risky decisions [76]. Information security awareness' education and training for users have been emphasized as effective measures against phishing attacks in different studies [101], [30].
Training	The role of training has been emphasized by different studies, as one of the most effective measures and the key mitigation strategy for the phishing attacks for each organization. These effects are regarding increasing individuals' information security awareness, improving users' protective behaviors, and preventing users from falling for the attacks [80], [74], [85], [101], [30], [48].
Neuroticism	Neuroticism, also known as emotional instability, is the tendency to easily experience negative emotions, such as anger or sadness. This causes individuals with a high level of neuroticism not be able to handle stress appropriately, think clearly, and make decisions; hence they become more vulnerable to phishing [97], [80], [45].
Extraversion	A higher level of extraversion is related to more inclination towards being in the other people's companionship and is associated with characteristics, like sociability or excitement seeking [97], [80]. Different studies have shown relationships between higher extraversion level and higher phishing vulnerability [69], [68], [23], [9].
Openness	Openness is the tendency to try new things and experiences without anxiety accompanied by intellectual curiosity [80]. Studies regarding openness have shown that a high level of openness is related to higher phishing vulnerability [9] and having less strict privacy settings while posting more on Facebook [45].

Table II Continued:

Agreeableness	Agreeableness is the tendency towards altruism, sympathy, and willingness to help rather than being competitive and egocentric [97]. Although it has been found that higher agreeableness is related to a higher information security awareness [76], various research has shown higher agreeableness results in more phishing vulnerability and being at a high rate of security risk for the possessor [88], [23], [80].
Conscientiousness	Conscientiousness consists of traits such as self-control, organizing, and determination [97], which causes the conscious person more likely to obey the security guidelines and training [80], [23]. Higher conscientiousness is related to higher information security awareness and inclination to less risky behavior [76], [91].
Mood	In general, the positive mood causes impulsivity and inertia [55] and gives a sense of security to the person that the environment is safe; therefore, it triggers a low level of cognitive effort by activating the heuristic processing. Whereas, the negative mood is the indicator of an unsafe environment, that needs a higher level of cognitive resources and sets off careful, systematic processing, making the possessor skeptical rather than gullible [24], [102], [34]. However, congruency of the message with target's mood, when it is known, can increase the success likelihood as it 'feels right' for the recipient, especially when other factors leading to thinking are kept at minimum [73], [93].
Work experience	Includes attributes associated with the person's present and previous jobs, such as expertise in the job, years of functioning in each role, gathered skills, and job description. Experience becomes a source of information over time [60], helping the possessor make the right decision, in a way that successful experience is a chief factor of a CISO credibility [65].
Years in the current company	Although a limited numbers of previous studies have considered this variable, two studies have shown that the employees who were hired for a longer period by a company were less likely to fall for the phishing attempts [16], [63].
Stress	Stress has been proved to be related to lower performance, attention or memory deficits, higher task error rate, errors in judgment, narrowing visual attention, and reduced cognitive resources that all make users vulnerable to phishing attacks [80], [103]. Furthermore, stress causes the tunneling effect that results in focusing on the main task and decreasing the attention on peripheral information [102]. This is especially important in phishing that peripheral information plays a crucial role in phishing detection by the user.
Role	Some organizational roles are proved to be more vulnerable to phishing, such as employees from the call center, management, and HR/legal function [101]. Additionally, peers' social pressure to respond quickly is more when a person is higher in the organizational hierarchy, causing cognitive overload for the person [71].
Risk aversion	Causes the person to be knowingly inclined towards choices and decisions that contain less risk. While risk aversion increases by age [3], [27], its lower level is related to higher susceptibility to phishing attacks [100]. Moreover, risk perception triggers systematic processing [114].
Culture	Among the studied cultural factors, lower individualism [17], higher masculinity [49], [33], and higher Power Distance [16] have a higher vulnerability to phishing.
Devices	Devices that a person uses can increase the likelihood of success for phishing attempts by different factors. Some examples are hiding or truncating the complete URL in their browsers or their small screens [83], [12], that makes it hard to investigate the signs of the illegitimacy of a website, simple user interface for entering the credentials in mobile apps, allowing the attacker to develop a similar, believable one easier [39], [99], enhanced habituation caused by the device affordances [113], and the owners' feelings of trust in their mobile devices [12].

B. Effective variables in personalization

Table III, contains the variables from the related literature, that can be used in the personalization of the message and points to consider for using them, their descriptions, and examples for them. These variables give the opportunity of crafting a more believable message to increase the success rate. Further discussion for the variables is presented in Section A-B of the Appendix.

C. The framework

To integrate the mentioned effective variables into a structural design, the framework, which can be seen in Figure 1, is presented in this section.

Due to each variable's nature, three main vertical classifications of Static, Time dependant, and Environmental dependant were introduced to embrace and represent the variables inside them. Additionally, each class is split into personal and professional, considering the two most important parts of every person's life.

- Static variables are the target attributes that are unlikely to change during the period of gaining intelligence and the actual attack.
- Time dependant variables are highly likely to change in the mentioned time frame and, as the name suggests, are reliant on the time of the attack.
- Environmental variables are also dependant on the environment where the target is present when the attack is conducted.

Moreover, the framework is divided horizontally into three categories related to a phishing attack, namely likelihood, pretext, and implementation.

- Variables positioned in the likelihood section of the framework can indicate the target vulnerability to phishing attacks. Knowledge of these variables, when properly used, can significantly heighten the chance of phishing success.
- Definition of the pretext, as stated by Workman [122] is when "an imposter creates a setting designed to influence an intended victim to release sensitive information, pay money, or perform actions that compromise the confidentiality of information.". In other words, pretext is the story that the attacker chooses to deceive the target. The more the message is personalized to the target and fits their expectations, the higher the chance of succeeding in the attempt as the high level of personalization results in more trust in the message [58]. The compelling role of trust, as an influential factor in security behavior, has been investigated in different studies [63], [117]. Correspondingly, this category's variables can give a wide range of possibilities for crafting an authentic, believable pretext.
- Implementation part of the framework includes the variables that are influential in the attack's implementation phase. Accurate implementation of the attack plays a decisive role in the phishing success rate.

Furthermore, knowing the audience and target them in a proper context is a vital task in phishing. Context is the glue that holds the whole story and phishing components together. Dhamija et al. [25] discuss when the look and feel of the phishing sites are the same as the real targeted sites, context, or nature of the requested personal information is the only cue to the user to differentiate them. Greene et al. [41] study showed that the alignment of the user's context and phishing context is a crucial element for phishing susceptibility, that clickers and non-clickers interpret the same cues differently, based on the alignment of the message with their work context. This has been emphasized upon by another study [103], that as the context relevancy goes higher, the likelihood of the target paying attention becomes lower. Also, in a study by Benenson et al. [14] fitting of the message context with the user's expectations was the second most prevalent reason for clicking on the spam and also unfit between the situation context and life context of the message with user's expectation were responsible for 38.8% and 11.6% of reasons for not clicking on them respectively.

The importance of the match between the context and the variables is emphasized by the rectangle embracing the whole variables in the framework. Every variable should be used in the right context, and the context should be suitable for each of the variables.

D. Proxy variables

Additionally, some variables are hard to measure. Figure 2 illustrates the proxy variables that can be used to facilitate the measurement of the main variables. For each of the Big Five personality traits, two facets, among the available six facets, are chosen from the Big Five Inventory (BFI) [62], and the Revised NEO Personality Inventory (NEO PI-R) [21]. The higher level of each facet means that the main trait is stronger in the target's personality. Other proxy variables are selected from the existing literature (discussed in the Appendix A-A), where they are chosen based on their availability in OSINT and the higher likelihood of their inference from the available sources.

Correspondingly, Table VIII provides model measurement strategies for each variable of the framework by using OSINT, that can be helpful in giving more insight regarding the information gathering phase.

E. Framework's usage

Gathering intelligence and having the target's information for the presented variables in the human attack surface framework will increase the likelihood of success. Subject variables should be seen as interconnected ones, having direct effects on each other, and need to be examined and evaluated jointly.

Furthermore, the mere presence of a certain variable might be a helpful solution for the general phishing attempts, it cannot be a reliable method for highly targeted attacks. As one sample scenario, it may be the situation that many gathered variables, except for one, indicate the vulnerability to phishing for a target. Nevertheless, that one variable can change the

TABLE III
EFFECTIVE VARIABLES IN MESSAGE PERSONALIZATION

Variable	Description	Examples
Web platform	Each website and social media has a different nature and essence. Users expect more to see content related to the image of those media they have in mind. Therefore, knowing the platforms that the user has activity on can be truly helpful in creating a more suitable and more believable phishing message.	In a study on the Facebook platform [95], the number of clicks on sales spam was twice as the number of media spam. Sales spam, that were analogous to Facebook's advertisements fitted the Facebook platform more than media spam, that were mostly related to porn or violent content, which users are not expecting to see in such a platform.
Contacts network	The chance of the victims succumb to a targeted attack is four times higher if the sender is a known acquaintance, resulting in a higher chance of ignoring the critical clues by the recipient [57]. Also, being friends of friends will raise the chance of being accepted by the user due to the networking and connecting nature of the social medias [101].	Recipients of a study [95] had a higher probability of falling for the spam coming from friends of friends or pages, and spam re-shared by the friends of the recipients had a higher chance of succeeding than when they were re-shared by an unknown source.
Communities membership	Membership in either a virtual or physical community. Members of a community generally have similar traits or interests that can be informative about the target. Furthermore, members are more open to giving out personal information, especially in communities for social causes, where members want to help others [36]. One of the notable aspects of most of these communities is their easy access.	Target's interests, skills, beliefs, personalities, or time of presence in a location are a few examples. In a study [101], regarding the mentioned easy to access feature, authors could access a private discussion forum of a company consisted of 1200 employees without any verification.
Residence	Knowledge of user's residence and surroundings are considered as high value and can lead to more personalized attacks [98], [30], [90], [37], that some researchers have used geographic contexts to improve the believability of their emails [86].	Country of living, home address, and organized events in the neighborhood.
Work place	Just as the residency, knowledge about the user's workplace is valuable to the attacker [37], [80], [30].	The company name, location, working hours, or colleagues.
Life events	These events consist of all the happenings in the person's life, such as a newborn baby, attending a seminar, and having a real unpaid invoice among their tasks [41]. Such events can temporarily affect contextual relevancy [103], which causes the recipients to ignore some important cues for the phishing detection, such as email's source [115].	As a recent example, in April 2020, with the Coronavirus pandemic outbreak, Google company stated that it was blocking 18 million scam emails related to the Coronavirus everyday [106]. Also, a sudden growth of 667% was reported by Barracuda Networks security firm for the relevant phishing attacks by the end of February [36].
Likes and interests	Individuals tend to spend more time with the person that they think is more similar to them [22]. Knowledge of the targets' likes and interests, emphasized in different studies [104], [90], [109], [30], especially when customized based on the life domains of the target [87] making the target perceives the source as "like me".	Hobbies of the target are among the best examples of this category that can help to know about the person's interests [104].
Communication norms	Hadnagy in "The Art of Human Hacking" [44] indicates communication style as one of the decisive factors to successful elicitation. The communication norm consists of relevant information regarding the impersonated entity's communication aspects, which can enhance the message credibility.	In a study [120], receiving emails contradictory to the company's communication norm was emphasized by the respondents as a sign for the email illegitimacy, such as inappropriate day and time and receiving an external email where the employee typically only receives internal emails. Other examples include the company means of communication, like email, text messaging, or social media, tone, and language of the messages.
Visual cues	Visual cues are another staple factor for increasing the believability of the message [8], that the correct implementation of the visual cues can even fool the most sophisticated users [25]. These cues are linked to the perceived user trust from the brand, and perfectly imitated ones in a phishing email can stimulate that trust [79]. Moreover, users perceive appropriate visual cues as a sign of legitimacy and emails containing them as more trustworthy and persuasive [121], that users were more likely to fall for phishing emails containing logos [14].	Examples include logos, images, copyright statements, slogans, fonts, and margins.

whole equation, when, for example, the target is a Cyber Security specialist.

The proposed framework can be used on both the offensive and defensive side.

a) Offensive:

- **Attack development:** The framework assists the attacker to concentrate the efforts on the right place. This is especially important in the OSINT gathering phase, where the attacker can use the framework to identify the most effective variables in different phases of the attack (likelihood, pretext, and implementation) and concentrate on gathering those among all the available information.
- **Target identification:** Knowledge of the more vulnerable people leads to aiming for the targets that have a higher chance of getting phished. Specifically, it can help the attacker to spot the weak points of entry into the target organization by directing the attempts on the most vulnerable employees, which not only increases the likelihood of success but also attracts less attention in the whole organization.

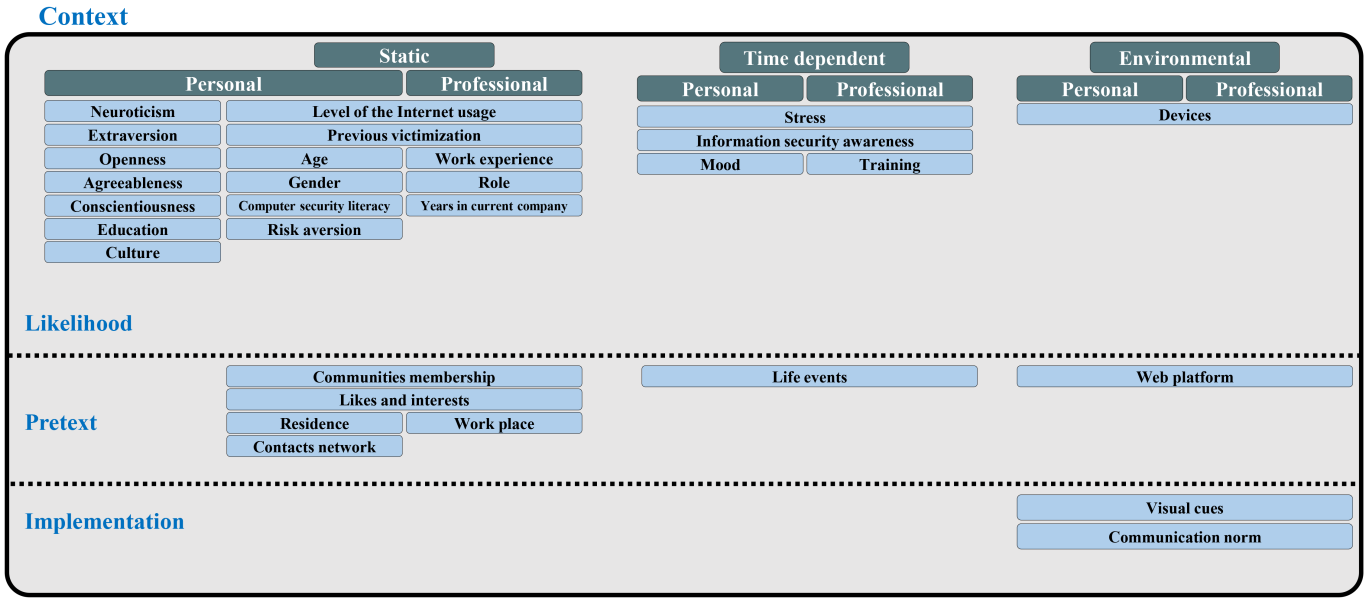

Fig. 1. The Human Attack Surface Framework for Phishing

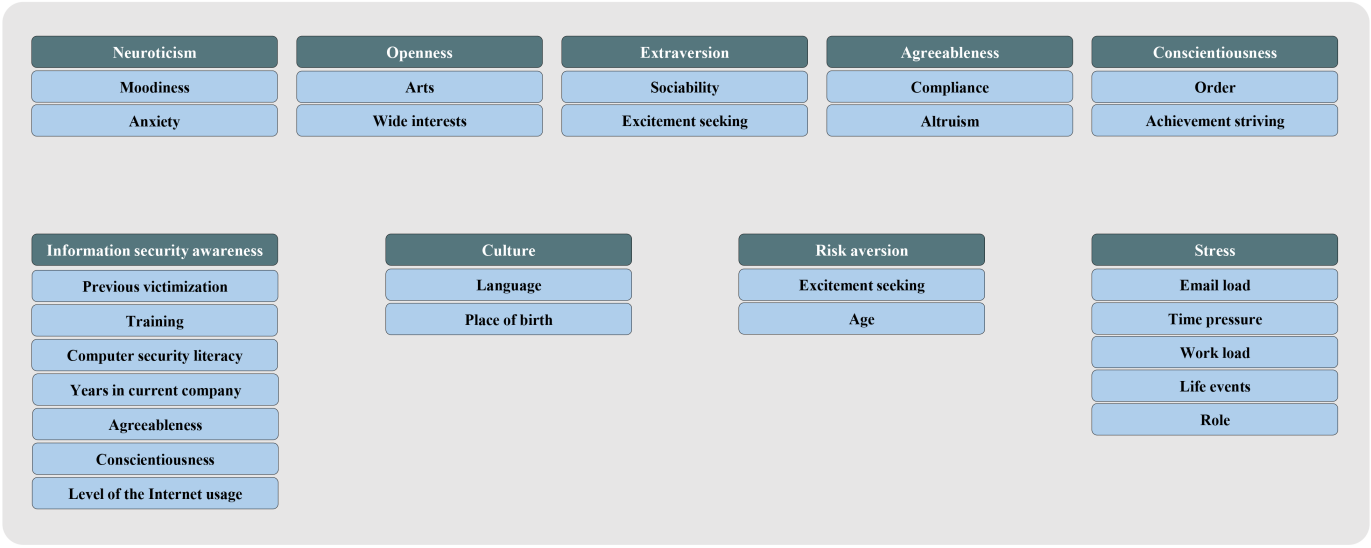

Fig. 2. Proxy variables to facilitate the measurement of the main variables

b) Defensive:

- **Attack surface evaluation:** Organizations can use the presented framework to measure their phishing attack surface. Subsequently, they can identify the vulnerabilities and reduce the attack surface by focusing on the weak spots.
- **Design of training activities:** Organizations can use the framework to create customized and tailored educational material and training, specific for each employee to train them against the targeted attacks. Consequently, this makes the training more effective and better increases the employees' awareness, for example, by using each employee's personal information in the internal phishing campaigns' pretexts.

IV. Experimental Evaluation

We performed an experiment to showcase the introduced framework in action and how it can be used for security measurements and training. Our experiment was based on two studies. First, the study by Hirsh et al. [51] that showed the higher effectiveness and influence of the messages that are congruent with the recipient's personality. Second, the social engineering framework developed by Uebelacker and Quiel [108] that maps each of the Big Five personality traits to Cialdini's principles of influence [19] that have the most persuasiveness on that specific type of personality.

The experiment's focus is to determine the impact of the message personalization on participants' preferences in general and the effect of personalization when tailored to their personalities. To achieve this, We selected four out of the Big Five personality traits from the introduced framework and used the corresponding principle of influence from the Uebelacker and Quiel framework [108] to personalize the message along with other personalization elements from the framework (see IV-A3). The personalized message was shown to the participants, adjacent to another baseline text. The experiment's success is measured in terms of rates of the users that preferred the personalized message.

A. Experiment design

The experiment was designed in a way to simulate the attacker collecting information from OSINT derived out of the targets' social media. This was to imitate the situation in a real attack that there is limited access to intelligence regarding the target in the information-gathering phase, and the attacker can only decide based on the available information. Accordingly, all the treatments were performed only based on the information that can be inferred from the target's social media to replicate the real-world scenario as much as possible.

To achieve this goal, we conducted an interactive, real-time experiment in the form of a Human Intelligence Task (HIT) in the Amazon Mechanical Turk (MTurk) platform to recruit the needed participants. The overview of the experimental procedure can be seen in Figure 17 in the Appendix. Participants first answered an online survey (see IV-A2), to determine their demographic characteristics and their personality traits that can be inferred from their social media. Next, they were faced with two emails, the baseline version, and the treated version created from the participants' answers to the survey (see IV-A3) to choose their preferred version.

1) Participants: Two hundred participants (mean age 35 years old) were recruited from Amazon Mechanical Turk (MTurk) website. All participants were at least 18 years old and informed about the essence of the study prior to its beginning via the MTurk platform.

2) Survey: The first part of the experiment contained an online survey (see Appendix B), with 7 to 10 minutes expected completion time, consisting of two sections. The first section contains general demographic questions about the respondent's age, educational degree, and work experience in Amazon Mechanical Turk. These variables are both influential in phishing vulnerability, discussed in section III, and used in the personalized version of the text. The second section includes personality questions and two questions regarding mood and stress level, focusing on the respondent's social media activities. Four out of five of the Big Five personality traits that have shown higher susceptibility to phishing in the literature, namely Neuroticism, Agreeableness, Extraversion, and Openness, were used in the experiment and evaluated using the corresponding two facets, introduced in Figure 2.

Each personality trait's facet contained two questions in the survey. First, the respondents were asked about their self-image concerning that specific trait, and its presence in their personalities (e.g., I consider myself as an anxious person). Answers to all the traits questions ranged from 'Strongly Disagree' to 'Strongly Agree'. If their answers were either 'Agree' or 'Strongly Agree', the next question for that trait would be shown to them; otherwise, the question remained invisible. This question asked about the possibility of inferring the presence of that specific trait in their personalities by a third person looking at their social media feeds (e.g., A person looking at my social media feed(s) can infer that I am an anxious person.). The latter question was also asked for their mood and stress level for the current period, in the time of answering the survey.

The questions relevant to the person's self-image will be referred to as type G questions in the rest of this article, and the questions that consider the inference of the trait from social media will be referred to as type S. The importance of the type S questions is twofold; first, depending on the participants' answers, they show the possibility of inference by an attacker for that trait from the social media profile of the intended victim. Secondly, it is an indicator of the presence of that trait in the respondent's personality, or at least the presence of the self-image, since the respondent had answered whether Agree or Strongly Agree to the previous type G question.

3) Experimental treatment: Our team crafted the baseline email by using Amazon Mechanical Turk as the pretext. In this regard, we examined some of the emails, messages, forums, and newsletters, sent from Amazon Mechanical Turk, to become familiar with their visual cues and communication norm. We further explored the MTurk platform to become aware of similar tasks and the ruling norms surrounding them. Moreover, we investigated workers' reviews, messages, and attitudes in relevant forums and groups to find the desirable features that they seek in each MTurk's HIT. The baseline email (Fig. 3) invites workers, in a general manner, with no customization, to subscribe for a HIT related to the MTurk's platform.

The basis of the treated version is the same as the baseline version; however, every treated version consists of extra customized variables. The specific age range and work experience, that fit the participants' corresponding attributes are the variables that were used in every customized text. Depends on the fit with the respondent's personality, one to four number of other variables, based on Cialdini's principles of influence [19], were added to the treated version basis (Fig 4). The texts for the variables were selected from the Oliveira's weapons of influence and life domains email samples at https://github.com/danielaoliveira/Counter-Balanced-Emails−−−Weapons-of-Influence-and-Life-Domains. Variables and the corresponding principles of influence and added texts can be seen in table IV. Also, all the texts contained no logo, general greetings, and the same signature used by MTurk, to comply with the visual cues and the communication norm of the actual emails from MTurk.

TABLE IV
VARIABLES AND CORRESPONDING VALUES FOR THE TREATED EMAIL CUSTOMIZATION.

Treated variable	Corresponding principle of influence [19]	Customized value	Default value
Age	-	and in the age group XY	
Work experience	-	(1)workers with at least one year of experience (2)workers with less than one year of experience	all workers
Agreeableness	Liking	to be its eyes and ears	
Extraversion	Social Proof	586 members of your community have already joined.	
Neuroticism	Authority	the Amazon MT Research Team asks you to	you should
Openness	Scarcity	the next 24 hours	2 weeks

Greetings from Amazon Mechanical Turk,

Amazon Mechanical Turk wants you in its research regarding the new platform for Mturk.

Our research will take place over the course of the next two weeks, and will include all workers. We need to recruit workers to join our HIT to respond to a survey related to their experience with the present platform. Participants are eligible for a reward in the amount of $10 per assignment.

To add your worker ID to the qualification list, you should click on the link below and fill out the information within 2 weeks.

<Subscription link>

Sincerely,

Amazon Mechanical Turk

https://www.mturk.com

Fig. 3. Baseline email text.

Greetings from Amazon Mechanical Turk,

Amazon Mechanical Turk wants you$AGREEABLENESS in its research regarding the new platform for Mturk.

Our research will take place over the course of the next two weeks, and will include $WORK_EXPERIENCE$AGE. We need to recruit workers to join our HIT to respond to a survey related to their experience with the present platform. Participants are eligible for a reward in the amount of $10 per assignment.

$EXTRAVERSIONTo add your worker ID to the qualification list, $NEUROTICISM click on the link below and fill out the information within $OPENNESS.

<Subscription link>

Sincerely,

Amazon Mechanical Turk

https://www.mturk.com

Fig. 4. Treated email text basis.

B. Procedure

The schematic view of the experiment setting can be seen in Figure 17. In the first part of the experiment, participants were redirected to the experiment page after reading the instruction and accepting the HIT in the MTurk platform.

On the first page, they answered the survey hosted on our web server, discussed in IV-A2. After the submission of the survey, the data was received by our server, and the respondent was redirected to the next page of the experiment. Simultaneously, on the server side, an automated routine analyzed the submitted data, and the available personality traits in the respondents were identified. A trait was considered available if the respondents answered 'Agree' or 'Strongly Agree' to the possibility of inferring the facet of that trait from their social media for any of its two facets. This is to simulate an attacker that could infer that information from the victim's social media profile. Subsequently, the customized version of the text was crafted by the calculation and placement of the respondent's age range and work experience in the template's dedicated fields, and the corresponding texts of the bold personality traits if the criteria were met.

On the second page of the experiment, the texts for the baseline email, and the generated treated email were presented to the respondents in a side-by-side manner. To remove bias, the baseline version was shown first (on the left side of the screen) to the first half of the participants, and the treated version was shown first to the second half. Then, the respondents were asked to answer which of the presented emails were they more likely to follow up to by clicking on the link. Another question was also asked about the likelihood of clicking on the subscription link if they had really received that email in their mailboxes. Each of the mentioned questions had a comment box that asked the participants for the reasoning behind their choice. After the submission of the answers, participants received their unique codes to finish the HIT in MTurk.

C. Ethical considerations

All participants agreed to participate voluntarily after being informed about the experiment. A fixed $1 compensation was received by each participant for contribution to the study after its completion, based on the average spent time of seven minutes. No personally identifiable information was collected or stored for this experiment, and the collected information was only used for scholarly purposes.

TABLE V
DEMOGRAPHIC CHARACTERISTICS OF PARTICIPANTS

Variable	Values	Number (%)
Age	18-25	23 (11.5)
	26-30	56 (28)
	31-35	49 (24.5)
	36-40	31 (15.5)
	41-45	10 (5)
	46-50	10 (5)
	51-55	8 (4)
	56 or older	13 (6.5)
Education	Primary school or lower	4 (2)
	Secondary school	43 (21.5)
	University degree	153 (76.5)
Work experience in MTurk	Less than 12 months	43 (21.5)
	12-24 months	48 (24)
	More than 24 months	109 (54.5)

V. RESULT EVALUATION

Due to the focus on the qualitative evaluation of the results, we have not performed statistical data analysis. Table V shows the participant's demographic characteristics of 200 participants. The majority of the participants were young, well-educated, and experienced with the MTurk platform.

Table VI summarizes the participants' answers to the survey's personality questions per each personality trait. Type G questions (relevant to the person's self-image), are shown by "_G" at the end, and type S questions (relevant to the inference possibility of the trait from person's social media) by "_S". Presence of a type S answer means that the participant had answered with Agree or Strongly Agree to the corresponding G type question. Also, Table VII summarizes the participants' answers to `Mood_S` and `Stress_S` questions. Only one participant answered Strongly Disagree to all the type S questions, and no information can be inferred to the `Mood_S` and `Stress_S` questions, which can be the sign of having no social media.

All personality facets, except for `Anxiety_S` with 34, had at least 60 answers of whether Agree or Strongly Agree to the type S questions (Table VI), indicating that attackers may be able to infer some of this information from the respondents' social media. `New things_S` with 96 total positive answers had the highest number among the traits. By calculating the percentage of the positive answers to the type S questions for the Big Five facets based on the number of participants who had answered that question, `Sociability_S` with 71.4% and `Excitement seeking_S` with 68.5%, as facets of Extraversion, ranked highest, and `Moodiness_S` with 40.5% and `Anxiety_S` with 44.7%, as facets of Neuroticism, ranked the lowest. The average percentage was 57%.

Furthermore, the number of answers indicating the possibility of inference for `Mood_S` and `Stress_S` was 142 and 141, respectively (Table VII). Additionally, answers to `Mood_S` and `Stress_S` followed a similar pattern that, for example, participants with a positive mood were also in the relaxed state, or when no information can be inferred from their social media about their stress level, no information can be inferred regarding their mood as well.

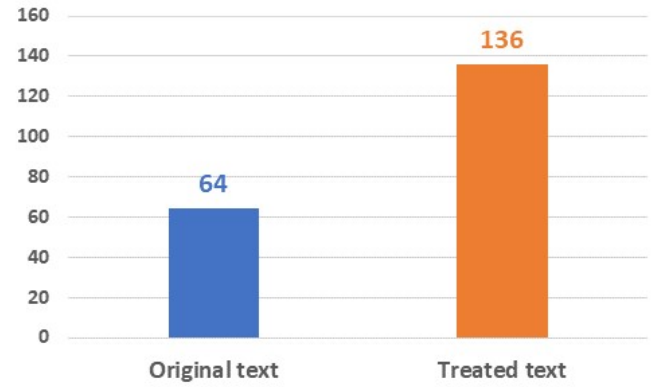

Fig. 5. Overall preferences for each version of the text

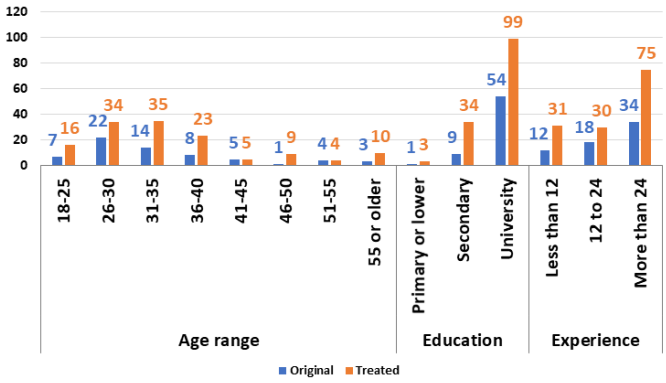

Fig. 6. Outcome preferences by participants demographics characteristics

A. Evaluation of outcome preferences

Regarding the text preference, 136 participants preferred the treated version, and 64 voted for the original version (Figure 5). Figure 6 illustrates the outcome preferences by participants' demographic characteristics. For the age groups of 41 to 45 and 51 to 55, half of the participants chose the original version, and half of them selected the treated one. For other age groups, the treated version had the highest share. The age group of 46 to 50 showed the highest difference, with most participants (9 out of 10) voted for the treated version.

Figures 7, 8, and 9 present the outcome preferences for the personality questions by categorizing them into Type G, Type S, and `Mood` and `Stress`, respectively.

In general, considering all answers to personality questions, the treated version had a higher number compared to the original version, except for `Excitement Seeking_S` and `Altruism_S` with the chosen answer of Strongly Agree that the numbers were equal, and `Moodiness_S` and `Anxiety_S` that the chosen answer of Strongly Agree resulted in a higher preference for the original version (Fig-

TABLE VI
ANSWERS TO THE PERSONALITY QUESTIONS FOR EACH TRAIT

Question	Strongly Disagree	Disagree	Neither	Agree	Strongly Agree
Moodiness_S	34	60	25	69	12
Anxiety_G	42	59	23	48	28
Anxiety_S	7	22	13	27	7
Art_G	14	23	28	86	49
Art_S	6	17	28	59	25
New Things_G	5	16	29	105	45
New Things_S	9	16	29	70	26
Sociability_G	22	42	31	74	31
Sociability_S	1	13	16	52	23
Excitement Seeking_G	25	55	28	64	28
Excitement Seeking_S	1	6	22	39	24
Compliance_G	12	28	34	86	40
Compliance_S	7	26	32	43	18
Altruism_G	13	23	58	82	24
Altruism_S	9	17	20	46	14

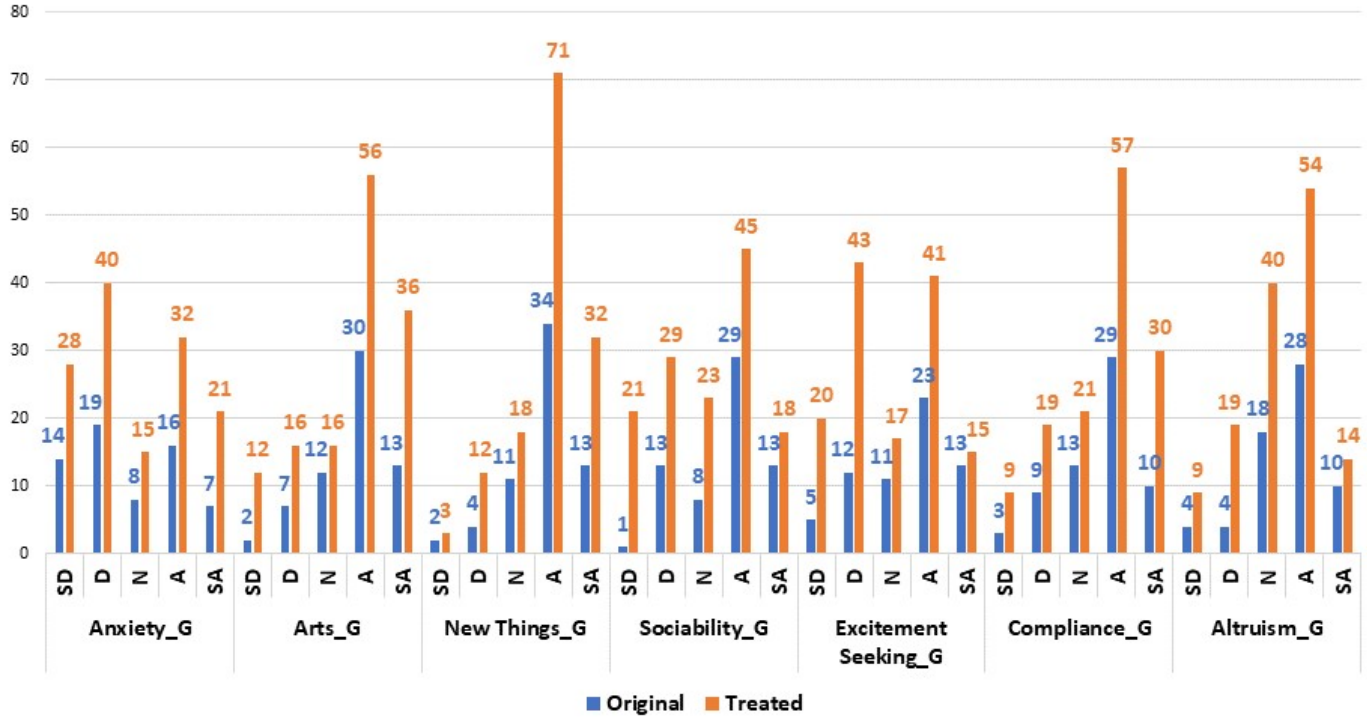

Fig. 7. Outcome preferences based on personality traits Type G questions

TABLE VII
ANSWERS TO THE PERSONALITY QUESTIONS FOR MOOD AND STRESS

Question	Negative	Neutral	Positive	No info can be inferred
Mood_S	4	36	102	58
	Stressed	Neither	Relaxed	No info can be inferred
Stress_S	7	32	102	59

ure 8). Additionally, from the 22 participants who answered Strongly Disagree to `Sociability_G`, 21 preferred the treated version (Figure 7). Furthermore, `Mood` and `Stress` showed a comparable pattern regarding the preference for each version. For example, 66 participants with the Positive mood, and 64 participants in Relaxed state, chose the treated version (Figure 9).

Moreover, considering both type G and S questions (Figures 7 and 8), `New things_G` had the highest number for treated version preference among the personality facets, with 103 participants, and the lowest number was dedicated to `Anxiety_S` with 53 participants. However, as shown in the

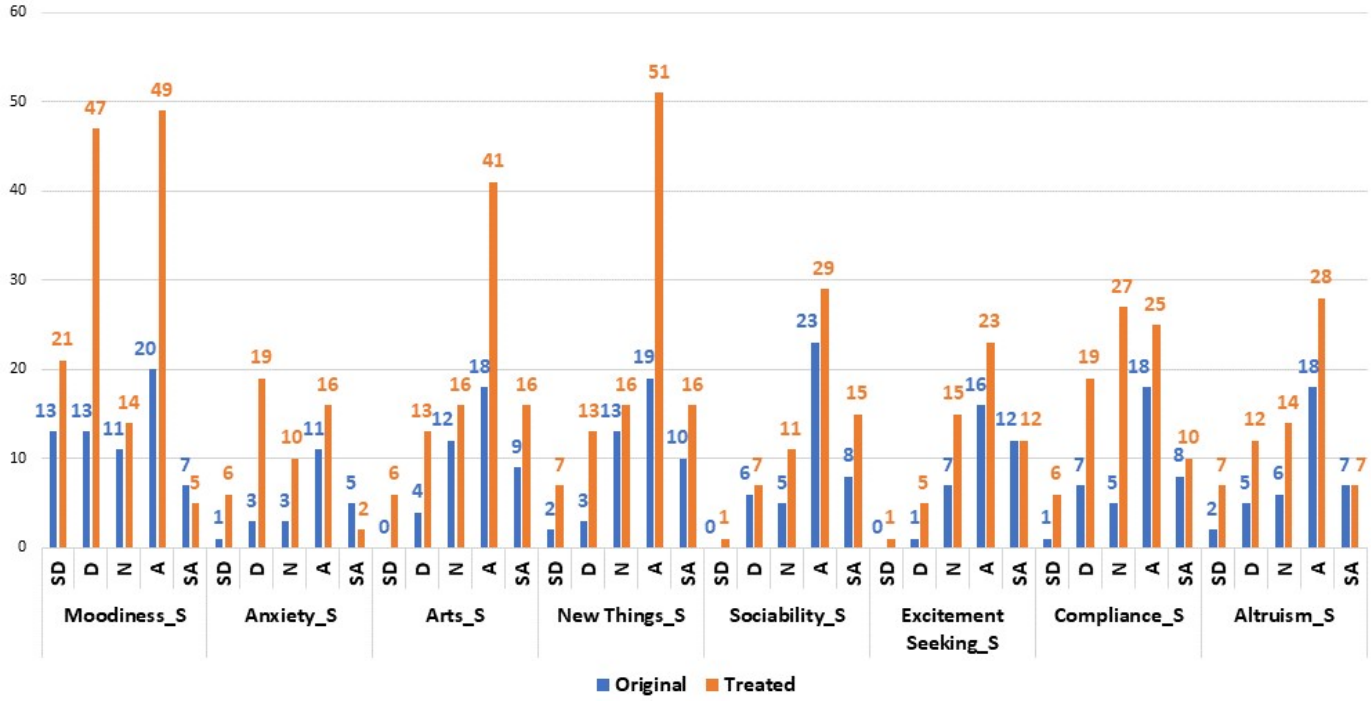

Fig. 8. outcome preferences based on personality traits Type S questions

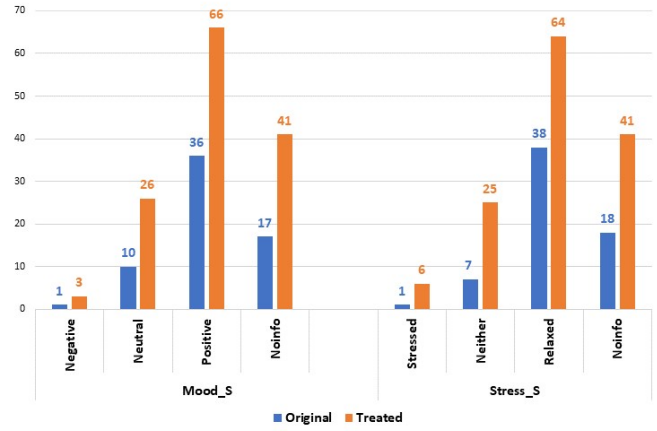

Fig. 9. Outcome preferences based on Mood and Stress Type S questions

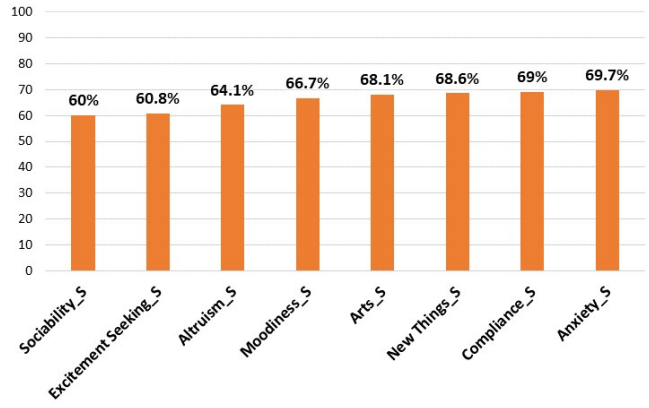

Fig. 10. Percentage of the preference for the treated version per each personality facet, including all the answers to the type S questions.

Figure 10, by taking the number of participants that had answered type S questions for each facet, the highest percentage of treated version preference was for `Anxiety` with 69.7%, and the lowest percentage belonged to `Sociability` with 60%. This can suggest the facets that were more affected by the personalization.

To investigate the effect of personality-related treatments, Figure 11, divides the answers to the type S questions into two categories of 'Neither or lower' and 'Agree or higher' for each trait. For both mentioned categories, the percentage of the preferred version, based on the number of the answers in that category, is presented. This shows the difference in the text preference when the corresponding treatment was introduced to the text for that trait (Agree or higher category), compared to when it was not. The treated version had a higher percentage in every category compared to the original version. However, for all the traits, the percentage for the treated version preference was lowered when the trait-specific treatment was added, except for `New Things_S` that raised by 3%. This was most striking for `Anxiety_S` and `Compliance_S` with 30.4% and 22.7% decrease in the treated version preference percentage, respectively. Therefore, for the majority of the traits, the introduced trait-specific treatments resulted in a lower preference for the treated version.

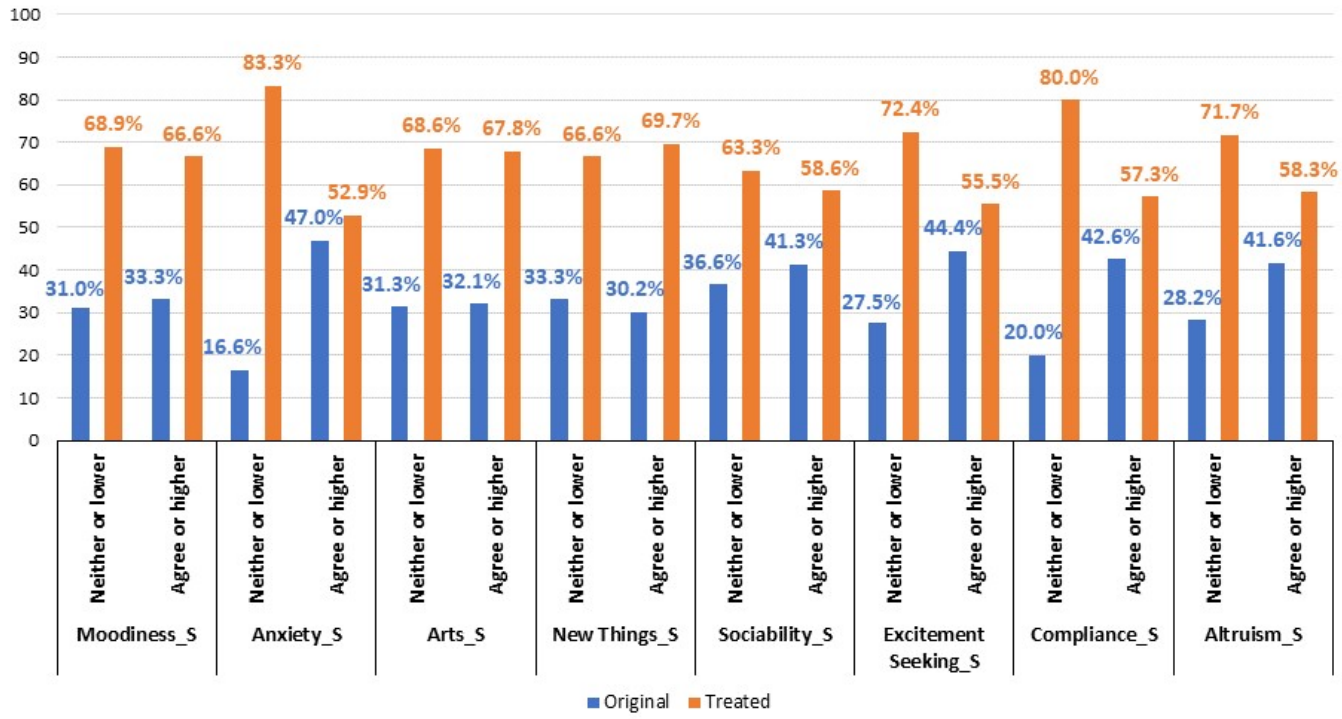

Fig. 11. Comparison between the preferred version of the text by dividing the participants' answers to two ranges of 'Neither or lower' and 'Agree or higher' for Type S questions to evaluate the effect of the introduced trait-specific treatments.

In the comment section of the preference question, we analyzed the different reasoning behind the participants' choices. For the original version, the reason that it was to the point, concise, and more objective was mentioned 16 times. Similarly, including all the workers and not only a specific group was specified 16 times. Usage of phrases belonged to the principles of influence for Agreeableness, Extraversion, and Openness in the treated version was mentioned by 9 participants for preferring the original version. The top reason for choosing the treated version was having more detail and targeting the participants with 101 mentions, showing the important role of personalization in general. Containing the number of members that have already joined, and having a higher chance to be qualified for the HIT were stated 15 times and 14 times, respectively. The contradictory outcomes of some elements, like targeting the user's specific age group or usage of phrases tailored to their personality, shows the need for further investigation.

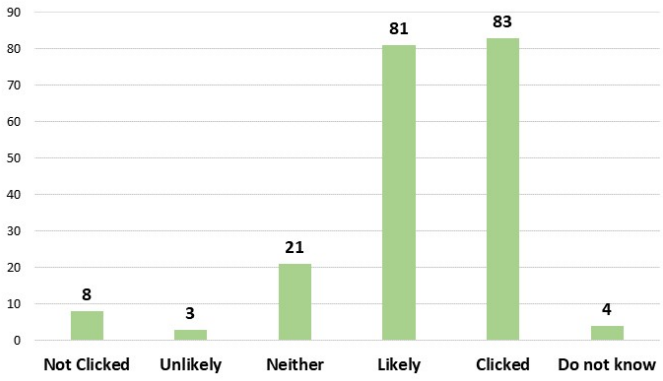

Fig. 12. Number of answers to the question related to the likelihood of click on the link per each option.

Concerning the participants' answers to the likelihood of click on the link (figure 12), 83 chose the option of Certainly Clicked, and 81 chose Likely while other options were chosen 36 times collectively. In the corresponding comment section, the dedicated reward had the highest number of mentions with 49 times as the reason for clicking on the link. Also, the willingness to share their opinion about the new platform and wanting their voice to be heard was mentioned 17 times. For 14 participants, the reason was the fact that the email looks legitimate, is sent from MTurk, or they trust MTurk. Other interesting comments regarding the reason of their choice were, having details, being targeted, and being informative (13 times); Because it is their routine to receive invitations or click on such emails daily (9 times); The text is inviting, interesting, honest, or with good wording (9 times); Others have joined, or they have only 24 hours to act (5 times). Additionally, although only 36 participants had not answered with Likely or Clicked, legitimacy of the email was mentioned 30 times, that they would check the sender, hover over the link, or they knew that the email was not legitimate.

Overall, the results indicate that trait-specific personaliza-

tion caused the treated version to be less appealing for the participants. Only `New Things_S` showed a slightly higher preference for the treated version after the treatment, and a 30.4% decrease for `Anxiety_S` was particularly noticeable. Nevertheless, customization of the message, targeted to the recipient, showed to lead to a more effective phishing attack in general. Meanwhile that a greater number of the participants chose the treated version, for most of them, including more details and being more targeted, were directly mentioned as the reason for preferring it. This clearly shows the effectiveness of personalization and highly emphasizes the role of having more information about the targets. In addition, Anxiety and Compliance showed the highest percentage of preference for the treated version in type S questions, and Sociability and Excitement seeking the lowest.

B. A preliminary qualitative evaluation of choice motivations

We now investigate the reasoning behind the participants' preferences qualitatively. We aim to achieve this objective by further scrutinizing the text preference question's comment section.

Among seven comments concerning the Agreeableness treatment's phrase (`to be its eyes and ears`), five of them were positive. We found that the participants who also had high rates in any of the Neuroticism facets, with at least one choice of Agree for any of them, tend to like it more. Whereas, the two participants with negative comments scored considerably lower in Neuroticism's facets. This may be due to the emotional weight of the phrase and the higher sensitivity of the Neurotic people to these kinds of stimuli. One of the participants mentioned

> *It seems more personalized like I am one of several and not one of many as in the second email. I feel like a person and not a number.*

This can show the possibility that personality traits are interconnected variables, need to be considered as a whole, that can affect each other and the decision-making procedure.

Furthermore, 18 comments were related to the phrase `586 members of your community have already joined`. Among the 15 positive comments for the treated version, three participants explicitly stated that it is the indicator that the text is "safe", "trustworthy", and gives "a sense of feeling of belonging", which shows the desired influence of the phrase on them. Interestingly, one participant clearly mentioned his fear as one of the best examples of the induced feeling of the corresponding phrase:

> *Because email number 2 has information about how many other members already joined. So i feel comfortable to join with that survey. I don't want to feel lonely.*

Additionally, another comment clearly shows the essence of the social proof and its direct effect:

> *it is more exclusive, well not exclusive exactly but now that i know people are already joining it makes me not want to miss out on it.*

However, although the mentioned comments show the effect of social proof on some participants, it had the opposite effect for three other participants. One of them had the lowest score (Disagree) for the Sociability facet among all the ten mentioned participants, and two had strikingly low scores for both Altruism and Compliance facets (with 3 Strongly Disagree and 1 Disagree), with similar comments. We speculate that these can result in less affection for other people, which one of the comments can be a good exemplar for it:

> *I don't necessarily need to know how many people aside from me have been chosen or invited or whatever. The number of other people is here nor there.*

While we could not find a specific connection between the usage of the `Amazon MT Research Team` for Neuroticism and the text preference, one comment showed that it had caused reciprocation on the participant:

> *The email makes it seem that amazon has more specifically target me for a response. I feel almost some responsibility that I should respond and provide what I can from my experience working with their platform.*

For 5 participants, time pressure, introduced by the `24 hours` phrase, was a reason to choose the original version since they could still be able to do the task if they forgot to subscribe. Even one participant who chose the treated version complained about the time constraint of it. Meanwhile, for 6 other participants, not only was this not a negative point but somewhat helpful to not "forget about the task" and therefore doing it right away. Also, two comments mentioned that it was an indication of the importance of the time and the task for the researchers.

Scarcity is effective on open people as they want to have new experiences and do not want to miss that chance. However, although a convincing pattern for the personality facets could not be found, we think the motivation behind the choices is similar. We presume due to our setting, that the participants had the chance to choose among the two texts, the freedom of having more time in the original version caused some of them to like it better. Additionally, most of the comments in favor of the treated version had the same idea of not missing out on the opportunity because of forgetting the task. Therefore, in a real-life scenario, when there is only one email to decide, this constraint can act as an incentive for the recipient to act quickly, hence, become more vulnerable.

Another notable finding is related to the context. Several comments were regarding the subjects that were used in our pretext explicitly, such as the new MTurk platform. Besides, many of the participants saw higher customization as a sign of having a higher chance to take part in the fictitious study:

> *Because it specifically targets workers with less than a year on MTurk, and I qualify. This will shrink the pool of applicants and I will have a better chance of being chosen.*

Nevertheless, some workers mentioned that they did not like the qualifications and rather prefer the text that includes all workers. Among many possible reasons, one can be the result of reading the two texts superficially and not paying attention to some cues, like the age range. Our further analysis showed that some of them were in the boundary ages of the mentioned age range in the text, such as 40 years old, that would fit in the age range of 36 to 40. Subsequently, they may have been worried that they would not qualify when the study is going to begin.

Our analysis showed the possibility that each personality facet, or in a broader sense, each personality trait may have complementary roles for each other in the phishing subject. Therefore, although each of them can be informative by themselves, considering the full picture of the target's personality can lead to a more effective attack. Moreover, context proved to have a chief role in affecting the participants' choices in our experiment, emphasizing its high importance.

VI. Discussion

A. Effects of the personalization

Overall, our study shows any variation of personalized text is more preferred compares to a non-targeted version. However, trait-specific treatments resulted in a lower preference for the treated version in all the personality facets, except for `New things` with a modest rise. The mentioned decrease in the preferability of the text can have different reasons. Generally, participants had the possibility of choosing between two versions of the text rather than receiving only one email to act upon. Coupled with that, our pretext and context for the experiment could have largely impacted the workers' choice. We hypothesize that the workers may find the brevity and conciseness of the invitation messages, as was the text of the baseline version, more desirable since it is equal to the opportunity of saving more time and having more income in the MTurk platform. Also, the fact that we only had one template and one phrase as a Representative of the corresponding principle of influence can also affected the result. For example, using 'authority' principle may not be a good fit with the context that is inviting the workers to another HIT with an included reward.

Moreover, although the difference is not large and further experiments are needed, our study indicates on which personality facets the personalization has the most effect. In total, `Anxiety`, `Compliance`, and `New things` showed the highest percentage of the preference for the treated version, among other traits. Interestingly, `Anxiety` and `Compliance` had the highest decline for the treated version's desirability after the addition of the trait-specific treatment. This shows that the proposed framework can be used to systematically evaluate the effectiveness of social engineering attacks, for example, by adding across "likelihood" and "pretext" variables.

B. Personalization is effective but complicated

Some customization elements had contradictory results, and some of them backfired in some scenarios. For example, targeting a specific group to participate in the fictitious study was why some participants did not like the treated message. In comparison, it gave the other participants confidence that they would have a higher chance of participating in the study and receiving the reward. Additionally, the usage of particular phrases or sentences to introduce principles of influence was interpreted differently by the participants. While the study has limited samples to conclude from, some personality facets show the possibility of affecting others in a different trait, which context of the message can have a highly influential role. Consequently, we speculate that personality facets should be considered in an interconnected way to result in a more effective attack. The introduced framework offers a wide variety of possibilities to conduct further, similar experiments with the presented variables to investigate more about their effects on personalization and the likelihood of success.

C. Availability of OSINT for targeted phishing

The analysis confirms that OSINT can be a valuable source of information about the target, even for evaluating personality, which is a tough subject to measure [96]. Considering Moodiness, which had the lowest percentage for the possibility of inference from social media, there is still a high chance (40.5%) of concluding the existence of that facet in the target. Results for our experiment show the highest chance of inference for the level of Extraversion, and the lowest for Neuroticism. This can be due to the nature of these two traits, that the facets of Extraversion have more visible manifestations in the target's social media than those for Neuroticism that mostly concern inner feelings. It can also give the attacker the possibility to focus more on the traits that are easier to measure and use the corresponding useful principles of influence. Besides, our experiment consisted of only two facets per each personality trait. Including other facets of the personality can result in a more accurate estimation of the target's personality. Meanwhile that our framework suggests effective variables in phishing success; it affords a focused view on OSINT. The provided proxy variables and the sample measurement strategy for the framework's variables are based on OSINT to make the best use of these available resources.

D. High likelihood of click

The study also demonstrates a high clicking possibility on the link if the participants had received the email in reality, that the majority of them chose either option of Clicked or Likely for the corresponding question. While the legitimacy of the link, sender, and the email, in general, was mentioned by a few numbers of participants in the comment section, only A small minority chose Not Clicked. We speculate that one of the reasons can be the fit between different elements of the phishing. Income and the MTurk platform itself are for sure of importance for each worker that is currently active. While the reward was attractive to the participants, it was also in a reasonable form and range, close to what a good HIT would pay in reality. Also, the reason for the reward was well justified in the context, in a way that a few participants only wanted to take part in making The MTurk platform, as the basis of the pretext, better.

All the mentioned reasons, combined with the other elements, like the communication norm and visual cues, could have been influential in not to raise the participants' doubt. Therefore, not only this indicates the need for training the users to raise their awareness, but it also shows the framework's true potential in helping with this goal by conducting more targeted attacks in training. Using the framework helps

the organizations to measure their phishing attack surface and therefore reduce it more efficiently.

E. Mood and stress relationship

Our experiment showed a close connection between mood and stress. A general pattern can be found for the number of chosen corresponding answers between the available options, such as the Positive mood with the Relaxed state. This pattern could be seen in the dedicated type S questions, preferred version of the texts, and the possible action if they have had received the email in the real-life. Even though this experiment's purpose is not to confirm the relationship between Mood and Stress, the result can imply that they can be suitable representatives of each other when one can not be measured from OSINT. Especially, with the chief role of stress in phishing vulnerability and different available ways that can be used to measure mood from social networks, such as sentiment analysis, or more efficiently, through mood and feelings fields in different social networks. Some of these ways are mentioned in the sample measurement strategy.

F. The framework

The presented framework provides structured critical points in phishing attacks that can be areas of focus to evaluate the phishing vulnerability and weak points in the organizations. Also, the framework's variables can highlight various items and details that can be used in inventing a believable, customized message with a fitting pretext, and the factors that need to be considered for a successful attempt. Therefore, the framework offers a wide range of opportunities to conduct different related experiments and studies. By using the framework, our experiment presented an interactive way of evaluating the phishing susceptibility for the researchers that do not have access to the organizations or needed infrastructure to conduct related experiments.

Additionally, it provides organizations and security specialists with use-cases to implement tailored phishing vulnerability tests and training for their employees to make them familiar with more real-life scenarios in a more effective way.

G. Limitations and future work

This research, however, is subject to several limitations. The most important limitation is that our experiment did not use real emails in its setting. Receiving the email by the participants in their mailboxes while they are not expecting it can be more indicative of their real reaction. In our setting, participants were asked to compare two versions of the text, which can decrease the effectiveness of some elements of the message, like the time pressure for scarcity.

Also, we measured the participants' personality traits by asking them to rate their traits. This makes the measurement less accurate compare to measuring them objectively from observations. Additionally, the number of questions for inferring each personality trait in a participant was restricted to two facets for each trait. Relevant to this, the treatments were used when any of the trait's facets scored at least Agree that can lower the precision of the treatments' effects in contrast with when there is a comprehensive way for introducing the treatment to the text.

Our sample size was limited, and our dataset for the treatments was restricted to one phrase for each principle of influence. Further experiments are needed to find more patterns regarding the personality and the effects of different principles of influence.

VII. Conclusion

By providing a structured map of the human attack surface in phishing, the presented framework seeks to assist the researchers and specialists in their fight against phishing. Our experiment confirmed that our framework's usage resulted in a more preferred message and, therefore, a more effective phishing campaign. However, the usage of trait-specific treatments caused lower desirability for the treated version in most cases. Furthermore, the true potential of OSINT in evaluating the target vulnerability was shown. Also, the high likelihood of clicking on the link, even in the experiment setting, showed an alarming phishing susceptibility result.

References

[1] 2020 Global Encryption Trends Study. Technical report, Ponemon Institute LLC, 2020.

[2] 2020 State of the Phish. Technical report, Proofpoint, 2020.

[3] Steven Albert and Duffy. Differences in risk aversion between young and older adults. *Neuroscience and Neuroeconomics*, page 3, 2 2012.

[4] Samar Albladi and George R.S. Weir. Vulnerability to social engineering in social networks: A proposed user-centric framework. *2016 IEEE International Conference on Cybercrime and Computer Forensic, ICCCF 2016*, 2016.

[5] Samar Muslah Albladi and R. S. George. Personality traits and cyber-attack victimisation: Multiple mediation analysis. *Joint 13th CTTE and 10th CMI Conference on Internet of Things - Business Models, Users, and Networks*, 2018-Janua:1–6, 2017.

[6] Abdullah Algarni. What message characteristics make social engineering successful on Facebook: The role of central route, peripheral route, and perceived risk. *Information (Switzerland)*, 10(6), 2019.

[7] Abdullah Algarni, Yue Xu, and Taizan Chan. An empirical study on the susceptibility to social engineering in social networking sites: The case of Facebook. *European Journal of Information Systems*, 26(6):661–687, 2017.

[8] Luca Allodi, Tzouliano Chotza, Ekaterina Panina, and Nicola Zannone. On The Need for New Antiphishing Measures Against Spear-Phishing Attacks. *IEEE Security & Privacy*, 18(2):23–34, 3 2020.

[9] Ibrahim Alseadoon, Taizan Chan, Ernest Foo, and Juan Gonzalez Nieto. Who is more susceptible to phishing emails?: A Saudi Arabian study. *ACIS 2012 : Proceedings of the 23rd Australasian Conference on Information Systems*, (Trusteer 2009):1–11, 2012.

[10] Ibrahim Alseadoon, M. F. I. Othman, and Taizan Chan. What Is the Influence of Users' Characteristics on Their Ability to Detect Phishing Emails? pages 949–962. 2015.

[11] Mary Jean Amon, Rakibul Hasan, Kurt Hugenberg, Bennett I Bertenthal, and Apu Kapadia. Influencing Photo Sharing Decisions on Social Media: A Case of Paradoxical Findings. *the Proceedings of the IEEE Symposium on Security & Privacy (SP '20), To appear*, pages 1–17, 2020.

[12] Belal Amro. Phishing Techniques in Mobile Devices. *Journal of Computer and Communications*, 06(02):27–35, 2018.

[13] Mohd Anwar, Wu He, Ivan Ash, Xiaohong Yuan, Ling Li, and Li Xu. Gender difference and employees' cybersecurity behaviors. *Computers in Human Behavior*, 69:437–443, 2017.

[14] Zinaida Benenson, Freya Gassmann, and Robert Landwirth. Unpacking spear phishing susceptibility. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 10323 LNCS:610–627, 2017.

[15] Cristina Bicchieri, Jan W. Lindemans, and Ting Jiang. A structured approach to a diagnostic of collective practices. *Frontiers in Psychology*, 5(DEC):1–13, 2014.

[16] Jan Willem Bullee, Lorena Montoya, Marianne Junger, and Pieter Hartel. Spear phishing in organisations explained. *Information and Computer Security*, 25(5):593–613, 2017.

[17] M Butavicius, K Parsons, M Pattinson, A Mccormac, D Calic, and M Lillie. Understanding Susceptibility to Phishing Emails: Assessing the Impact of Individual Differences and Culture. *Proceedings of the Eleventh International Symposium on Human Aspects of Information Security & Assurance (HAISA 2017)*, 2016(Haisa):12–22, 2017.

[18] Blake Butler, Brad Wardman, and Nate Pratt. REAPER: An automated, scalable solution for mass credential harvesting and OSINT. *eCrime Researchers Summit, eCrime*, 2016-June:71–80, 2016.

[19] Robert B Cialdini. *Influence: The new psychology of modern persuasion*. Morrow, 1984.

[20] Dan Conway, Ronnie Taib, Mitch Harris, Shlomo Berkovsky, K. Yu, and Fang Chen. A qualitative investigation of bank employee experiences of information security and phishing. *Proceedings of the 13th Symposium on Usable Privacy and Security, SOUPS 2017*, (Soups):115–129, 2019.

[21] Paul T. Costa and Robert R. McCrae. Normal personality assessment in clinical practice: The NEO Personality Inventory. *Psychological Assessment*, 4(1):5–13, 1992.

[22] Giulio Costantini, Juliette Richetin, Emanuele Preti, Erica Casini, Sacha Epskamp, and Marco Perugini. Stability and variability of personality networks. A tutorial on recent developments in network psychometrics. *Personality and Individual Differences*, 136:68–78, 2019.

[23] Ali Darwish, Ahmed El Zarka, and Fadi Aloul. Towards understanding phishing victims' profile. In *2012 International Conference on Computer Systems and Industrial Informatics*, pages 1–5. IEEE, 12 2012.

[24] Marieke de Vries, Rob Holland, and Cilia Witteman. Fitting decisions: Mood and intuitive versus deliberative decision strategies. *Cognition and Emotion*, 22(5):931–943, 2008.

[25] Rachna Dhamija, J. D. Tygar, and Marti Hearst. Why phishing works. *Conference on Human Factors in Computing Systems - Proceedings*, 1(November 2005):581–590, 2006.

[26] J. Digman. Personality Structure: Emergence Of The 5-Factor Model. *Annual Review of Psychology*, 41(1):417–440, 1990.

[27] Thomas Dohmen, Armin Falk, Bart Golsteyn, David Huffman, and Uwe Sunde. Identifying the effect of age on willingness to take risks. 2018.

[28] Julie S. Downs, Mandy Holbrook, and Lorrie Faith Cranor. Behavioral response to phishing risk. *ACM International Conference Proceeding Series*, 269:37–44, 2007.

[29] Chaiken S. Eagly AH. The Psychology of Attitudes. Fort Worth, TX: Harcourt Brace. *The Psychology of Attitudes. Fort*, 1993.

[30] Matthew Edwards, Robert Larson, Benjamin Green, Awais Rashid, and Alistair Baron. Panning for gold: Automatically analysing online social engineering attack surfaces. *Computers and Security*, 69:18–34, 2017.

[31] Frank Enos, Stefan Benus, Robin L. Cautin, Martin Graciarena, Julia Hirschberg, and Elizabeth Shriberg. Personality factors in human deception detection: Comparing human to machine performance. *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, 2:813–816, 2006.

[32] Waldo Rocha Flores and Mathias Ekstedt. A Model for Investigating Organizational Impact on Information Security Behavior. *WISP 2012 Proceedings*, pages 12–15, 2012.

[33] Waldo Rocha Flores, Hannes Holm, Marcus Nohlberg, and Mathias Ekstedt. Investigating personal determinants of phishing and the effect of national culture. *Information and Computer Security*, 23(2):178–199, 2015.

[34] Joseph P. Forgas and Rebekah East. On being happy and gullible: Mood effects on skepticism and the detection of deception. *Journal of Experimental Social Psychology*, 44(5):1362–1367, 2008.

[35] Edwin D. Frauenstein and Stephen V. Flowerday. Social network phishing: Becoming habituated to clicks and ignorant to threats? *2016 Information Security for South Africa - Proceedings of the 2016 ISSA Conference*, pages 98–105, 2016.

[36] Edwin Donald Frauenstein and Stephen Flowerday. Susceptibility to phishing on social network sites: A personality information processing model. *Computers and Security*, 94:101862, 2020.

[37] Diane Gan and Lily Jenkins. Social Networking Privacy—Who's Stalking You? *Future Internet*, 7(4):67–93, 2015.

[38] Roy Godson and James J Wirtz. Strategic denial and deception. *Trends in Organized Crime*, 6(1):5–16, 9 2000.

[39] Diksha Goel and Ankit Kumar Jain. Mobile phishing attacks and defence mechanisms: State of art and open research challenges. *Computers and Security*, 73:519–544, 2018.

[40] Roderick Graham and Ruth Triplett. Capable Guardians in the Digital Environment: The Role of Digital Literacy in Reducing Phishing Victimization. *Deviant Behavior*, 38(12):1371–1382, 12 2017.

[41] Kristen Greene, Michelle Steves, Mary Theofanos, and Jennifer Kostick. User Context: An Explanatory Variable in Phishing Susceptibility. In *Proceedings of the Network and Distributed Systems Security (NDSS) Symposium*, number February, pages 1–14, 2018.

[42] Frank L. Greitzer, Justin Purl, Yung Mei Leong, and D. E.Sunny Becker. SOFIT: Sociotechnical and organizational factors for insider threat. *Proceedings - 2018 IEEE Symposium on Security and Privacy Workshops, SPW 2018*, pages 197–206, 2018.

[43] Zhen Guo, Jin-Hee Cho, Ing-Ray Chen, Srijan Sengupta, Michin Hong, and Tanushree Mitra. Online Social Deception and Its Countermeasures for Trustworthy Cyberspace: A Survey. 1(1), 2020.

[44] Christopher Hadnagy. *Social engineering: The art of human hacking*. John Wiley & Sons, 2010.

[45] Tzipora Halevi, Jim Lewis, and Nasir Memon. Phishing, Personality Traits and Facebook. 2013.

[46] Tzipora Halevi, Nasir Memon, and Oded Nov. Spear-Phishing in the Wild: A Real-World Study of Personality, Phishing Self-Efficacy and Vulnerability to Spear-Phishing Attacks. *SSRN Electronic Journal*, 2015.

[47] Eszter Hargittai, Anne Marie Piper, and Meredith Ringel Morris. From internet access to internet skills: digital inequality among older adults. *Universal Access in the Information Society*, 18(4):881–890, 2019.

[48] Farkhondeh Hassandoust, Harminder Singh, and Jocelyn Williams. How Contextualisation Affects the Vulnerability of Individuals to Phishing Attempts. *AIS Electronic Library (AISeL)*, 2019.

[49] Diane Henshel, Char Sample, Mariana Cains, and Blaine Hoffman. Integrating cultural factors into human factors framework and ontology for cyber attackers. In *Advances in Intelligent Systems and Computing*, volume 501, pages 123–136. Springer Verlag, 2016.

[50] Jacob B. Hirsh and Michael Inzlicht. The devil you know: Neuroticism predicts neural response to uncertainty. *Psychological Science*, 19(10):962–967, 2008.

[51] Jacob B. Hirsh, Sonia K. Kang, and Galen V. Bodenhausen. Personalized Persuasion: Tailoring Persuasive Appeals to Recipients' Personality Traits. *Psychological Science*, 23(6):578–581, 2012.

[52] Geert Hofstede. *Culture's consequences: Comparing values, behaviors, institutions and organizations across nations*. Sage publications, 2001.

[53] Geert Hofstede and Robert R. McCrae. Personality and Culture Revisited: Linking Traits and Dimensions of Culture. *Cross-Cultural Research*, 38(1):52–88, 2004.

[54] Geert Hoftede, Gert Jan Hofstede, and Michael Minkov. *Cultures and organizations: software of the mind: intercultural cooperation and its importance for survival*. McGraw-Hill, 2010.

[55] Yu-feng Huang and Feng-yang Kuo. Positive Moods Can Encourage Inertial Decision Making: Evidence from Eye-Tracking Data. pages 229–238. 2020.

[56] Cristian Iuga, Jason R.C. Nurse, and Arnau Erola. Baiting the hook: factors impacting susceptibility to phishing attacks. *Human-centric Computing and Information Sciences*, 6(1), 2016.

[57] Tom N. Jagatic, Nathaniel A. Johnson, Markus Jakobsson, and Filippo Menczer. Social phishing. *Communications of the ACM*, 50(10):94–100, 2007.

[58] Markus Jakobsson. The Human Factor in Phishing. *Privacy Security of Consumer Information*, 7:1–19, 2007.

[59] Mohammad S. Jalali, Maike Bruckes, Daniel Westmattelmann, and Gerhard Schewe. Why Employees (Still) Click on Phishing Links: Investigation in Hospitals. *Journal of medical Internet research*, 22(1):e16775, 2020.

[60] Jongkil Jeong, Joanne Mihelcic, Gillian Oliver, and Carsten Rudolph. Towards an improved understanding of human factors in cybersecurity. *Proceedings - 2019 IEEE 5th International Conference on Collaboration and Internet Computing, CIC 2019*, pages 338–345, 2019.

[61] W. Jiang. The relationship between culture and language. *ELT Journal*, 54(4):328–334, oct 2000.

[62] Oliver P John, Sanjay Srivastava, and others. The Big Five trait taxonomy: History, measurement, and theoretical perspectives. *Handbook of personality: Theory and research*, 2(1999):102–138, 1999.

[63] W. D. Kearney and H. A. Kruger. Considering the influence of human trust in practical social engineering exercises. *2014 Information Security for South Africa - Proceedings of the ISSA 2014 Conference*, 2014.

[64] Gail Kinman. Work stressors, health and sense of coherence in UK academic employees. *Educational Psychology*, 28(7):823–835, 2008.

[65] Richard Klimoski. Critical Success Factors for Cybersecurity Leaders. *People & Strategy*, 39(1):14–18, 2016.

[66] Kostadin Kushlev and Elizabeth W. Dunn. Checking email less frequently reduces stress. *Computers in Human Behavior*, 43(February):220–228, 2015.

[67] Elmer E.H. Lastdrager. Achieving a consensual definition of phishing based on a systematic review of the literature. *Crime Science*, 3(1):1–10, 2014.

[68] Patrick Lawson, Carl J. Pearson, Aaron Crowson, and Christopher B. Mayhorn. Email phishing and signal detection: How persuasion principles and personality influence response patterns and accuracy. *Applied Ergonomics*, 86(December 2018):103084, 2020.

[69] Patrick Lawson, Olga Zielinska, Carl Pearson, and Christopher B. Mayhorn. Interaction of personality and persuasion tactics in email phishing attacks. *Proceedings of the Human Factors and Ergonomics Society*, 2017-Octob:1331–1333, 2017.

[70] Mario Luca, Jonathan Nagler, and Joshua A Tucker. You Won't Believe Our Results ! pages 1–30, 2020.

[71] Gloria Mark, Shamsi T. Iqbal, Mary Czerwinski, Paul Johns, Akane Sano, and Yuliya Lutchyn. Email Duration, Batching and Self-interruption. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pages 1717–1728, New York, NY, USA, 5 2016. ACM.

[72] Gloria J. Mark, Stephen Voida, and Armand V. Cardello. "A pace not dictated by electrons": An empirical study of work without email. *Conference on Human Factors in Computing Systems - Proceedings*, pages 555–564, 2012.

[73] Leonard L Martin, Teresa Abend, Constantine Sedikides, and Jeffrey D Green. How would it feel if...? Mood as input to a role fulfillment evaluation process. *Journal of Personality and Social Psychology*, 73(2):242, 1997.

[74] Christopher B. Mayhorn and Patrick G. Nyeste. Training users to counteract phishing. *Work*, 41(1):3549–3552, 2012.

[75] Frank T. McAndrew. When Do Personality Traits Predict Behavior?, 2018.

[76] Agata McCormac, Tara Zwaans, Kathryn Parsons, Dragana Calic, Marcus Butavicius, and Malcolm Pattinson. Individual differences and Information Security Awareness. *Computers in Human Behavior*, 69:151–156, 2017.

[77] Robert R. McCrae. Human nature and culture: A trait perspective. *Journal of Research in Personality*, 38(1):3–14, 2004.

[78] Robert R. McCrae and René Mõttus. What Personality Scales Measure: A New Psychometrics and Its Implications for Theory and Assessment. *Current Directions in Psychological Science*, 28(4):415–420, 2019.

[79] María M. Moreno-Fernández, Fernando Blanco, Pablo Garaizar, and Helena Matute. Fishing for phishers. Improving Internet users' sensitivity to visual deception cues to prevent electronic fraud. *Computers in Human Behavior*, 69:421–436, 2017.

[80] David Mundie. Unintentional Insider Threats: Social Engineering. (January):82, 2014.

[81] Kevin Munger. All the News That's Fit to Click: The Economics of Clickbait Media. *Political Communication*, 4609, 2019.

[82] Ajaya Neupane, Nitesh Saxena, Jose Omar Maximo, and Rajesh Kana. Neural Markers of Cybersecurity: An fMRI Study of Phishing and Malware Warnings. *IEEE Transactions on Information Forensics and Security*, 11(9):1970–1983, 2016.

[83] Yuan Niu, Francis Hsu, and Hao Chen. iPhish : Phishing Vulnerabilities on Consumer Electronics. *UPSEC'08 Proceedings of the 1st Conference on Usability, Psychology, and Security*, pages 10:1–10:8, 2008.

[84] Gareth Norris and Alexandra Brookes. Personality, emotion and individual differences in response to online fraud. *Personality and Individual Differences*, (January):109847, 2020.

[85] Gizem Ögütçü, Özlem Müge Testik, and Oumout Chouseinoglou. Analysis of personal information security behavior and awareness. *Computers and Security*, 56:83–93, 2016.

[86] Daniela Oliveira, Harold Rocha, Huizi Yang, Donovan Ellis, Sandeep Dommaraju, Melis Muradoglu, Devon Weir, Adam Soliman, Tian Lin, and Natalie Ebner. Dissecting spear phishing emails for older vs young adults: On the interplay of weapons of influence and life domains in predicting susceptibility to phishing. *Conference on Human Factors in Computing Systems - Proceedings*, 2017-May:6412–6424, 2017.

[87] Daniela Seabra Oliveira, Tian Lin, Harold Rocha, Donovan Ellis, Sandeep Dommaraju, Huizi Yang, Devon Weir, Sebastian Marin, and Natalie C. Ebner. Empirical analysis of weapons of influence, life domains, and demographic-targeting in modern spam: an age-comparative perspective. *Crime Science*, 8(1):1–14, 2019.

[88] J.L. Parrish Jr., J.L. Bailey, and J.F. Courtney. A personality based model for determining susceptibility to phishing attacks. *Little Rock: University of Arkansas*, pages 285–296, 2009.

[89] Kathryn Parsons, Marcus Butavicius, Paul Delfabbro, and Meredith Lillie. Predicting susceptibility to social influence in phishing emails. *International Journal of Human Computer Studies*, 128(July 2018):17–26, 2019.

[90] Javier Pastor-Galindo, Pantaleone Nespoli, Felix Gomez Marmol, and Gregorio Martinez Perez. The not yet exploited goldmine of OSINT: Opportunities, open challenges and future trends. *IEEE Access*, 8:10282–10304, 2020.

[91] Malcolm Pattinson, Marcus Butavicius, Kathryn Parsons, Agata McCormac, and Dragana Calic. Factors that Influence Information Security Behavior: An Australian Web-Based Study. pages 231–241. 2015.

[92] John Pescatore. SANS Top New Attacks and Threat Report. Technical report, SANS Institute, 2020.

[93] Richard E. Petty and Pablo Briñol. Emotion and persuasion: Cognitive and meta-cognitive processes impact attitudes. *Cognition and Emotion*, 29(1):1–26, 2015.

[94] Swapan Purkait, Sadhan De Kumar, and Damodar Suar. An empirical investigation of the factors that influence Internet user's ability to correctly identify a phishing website. *Information Management and Computer Security*, 22(3):194–234, 2014.

[95] Elissa M. Redmiles, Neha Chachra, and Brian Waismeyer. Examining the demand for spam: Who clicks? *Conference on Human Factors in Computing Systems - Proceedings*, 2018-April:1–10, 2018.

[96] William Revelle. Hans Eysenck: Personality theorist. *Personality and Individual Differences*, 103:32–39, 2016.

[97] S Rothmann and E P Coetzer. The big five personality dimensions and job performance. *SA Journal of Industrial Psychology*, 29(1):68–74, 10 2003.

[98] Fraser Sampson. Intelligent evidence. *The Police Journal: Theory, Practice and Principles*, 90(1):55–69, 2017.

[99] Hossain Shahriar, Tulin Klintic, and Victor Clincy. Mobile Phishing Attacks and Mitigation Techniques. *Journal of Information Security*, 06(03):206–212, 2015.

[100] Steve Sheng, Mandy Holbrook, Ponnurangam Kumaraguru, Lorrie Faith Cranor, and Julie Downs. Who falls for phish? A demographic analysis of phishing susceptibility and effectiveness of interventions. In *Conference on Human Factors in Computing Systems - Proceedings*, volume 1, pages 373–382. Association for Computing Machinery, 2010.

[101] Mario Silic and Andrea Back. The dark side of social networking sites:Understanding phishing risks. *Computers in Human Behavior*, 60:35–43, 2016.

[102] Mark A. Staal. Stress, Cognition, and Human Performance: A Literature Review and Conceptual Framework. In *Aeronautics and Space Administration*, 2004.

[103] Michelle P. Steves, Kristen K. Greene, and Mary F. Theofanos. A Phish Scale: Rating Human Phishing Message Detection Difficulty. (February):1–14, 2019.

[104] Guillermo Suarez-Tangil, Matthew Edwards, Claudia Peersman, Gianluca Stringhini, Awais Rashid, and Monica Whitty. Automatically Dismantling Online Dating Fraud. *IEEE Transactions on Information Forensics and Security*, 15:1128–1137, 2020.

[105] Jerry Chih Yuan Sun, Shih Jou Yu, Sunny S.J. Lin, and Shian Shyong Tseng. The mediating effect of anti-phishing self-efficacy between college students' internet self-efficacy and anti-phishing behavior and gender difference. *Computers in Human Behavior*, 59:249–257, 2016.

[106] Joe Tidy. Google blocking 18m coronavirus scam emails every day, 2020.
[107] Twitter Inc. An update on our security incident, 2020.
[108] Sven Uebelacker and Susanne Quiel. The social engineering personality framework. *Proceedings - 4th Workshop on Socio-Technical Aspects in Security and Trust, STAST 2014 - Co-located with 27th IEEE Computer Security Foundations Symposium, CSF 2014 in the Vienna Summer of Logic 2014*, pages 24–30, 2014.
[109] Kota Uehara, Kohei Mukaiyama, Masahiro Fujita, Hiroki Nishikawa, Takumi Yamamoto, Kiyoto Kawauchi, and Masakatsu Nishigaki. Basic Study on Targeted E-mail Attack Method Using OSINT. In Leonard Barolli, Makoto Takizawa, Fatos Xhafa, and Tomoya Enokido, editors, *International Conference on Advanced Information Networking and Applications*, volume 926 of *Advances in Intelligent Systems and Computing*, pages 1329–1341. Springer International Publishing, Cham, 2020.
[110] Rohit Valecha, Adam Gonzalez, Jeffrey Mock, Edward J. Golob, and H. Raghav Rao. Investigating phishing susceptibility—an analysis of neural measures. *Lecture Notes in Information Systems and Organisation*, 32(November 2019):111–119, 2020.
[111] Amber Van Der Heijden and Luca Allodi. Cognitive triaging of phishing attacks. *Proceedings of the 28th USENIX Security Symposium*, (2019):1309–1326, 2019.
[112] Verizon. 2019 Data Breach Investigations. Technical report, 2019.
[113] Arun Vishwanath. Mobile device affordance: Explicating how smartphones influence the outcome of phishing attacks. *Computers in Human Behavior*, 63:198–207, 2016.
[114] Arun Vishwanath, Brynne Harrison, and Yu Jie Ng. Suspicion, Cognition, and Automaticity Model of Phishing Susceptibility. *Communication Research*, 45(8):1146–1166, 2018.
[115] Arun Vishwanath, Tejaswini Herath, Rui Chen, Jingguo Wang, and H. Raghav Rao. Why do people get phished? Testing individual differences in phishing vulnerability within an integrated, information processing model. *Decision Support Systems*, 51(3):576–586, 2011.
[116] Rick Wash and Molly M. Cooper. Who provides phishing training? Facts, stories, and people like me. *Conference on Human Factors in Computing Systems - Proceedings*, 2018-April:1–12, 2018.
[117] Paul A. Watters. Why do users trust the wrong messages? A behavioural model of phishing. *2009 eCrime Researchers Summit, eCRIME '09*, 2009.
[118] Dirk Weirich and Martina Angela Sasse. Pretty good persuasion. page 137, 2001.
[119] Mark Wiggins and David O'Hare. Expertise in aeronautical weather-related decision making: A cross-sectional analysis of general aviation pilots. *Journal of Experimental Psychology: Applied*, 1(4):305–320, 1995.
[120] Emma J. Williams, Joanne Hinds, and Adam N. Joinson. Exploring susceptibility to phishing in the workplace. *International Journal of Human Computer Studies*, 120(June):1–13, 2018.
[121] Emma J. Williams and Danielle Polage. How persuasive is phishing email? The role of authentic design, influence and current events in email judgements. *Behaviour and Information Technology*, 38(2):184–197, 2019.
[122] Michael Workman. Wisecrackers: A theory-grounded investigation of phishing and pretext social engineering threats to information security. *Journal of the American Society for Information Science and Technology*, 59(4):662–674, 2 2008.

APPENDIX A
FURTHER DISCUSSION OF THE EFFECTIVE VARIABLES

A. Effective variables in likelihood of phishing success

1) Gender: Most of the research has found that females are more susceptible to phishing than men [100], [57], [70], [86], [101]. Sun et al. [105] argue that female students have greater Internet anxiety and less positive attitudes towards the Internet, which causes them to be less comfortable with Internet tools and respond less effectively to anti-phishing instructions. In another study [13] female participants had a significantly lower security self-efficacy. Furthermore, Jagatic et al. [57] state that females in their study had less technical knowledge and training than the male participants that they believe could have affected their vulnerability to the phishing to a certain extent. As a side note related to the gender, they state that receiving the phishing message from a person of the opposite gender causes the attack to be more successful, which has a higher effect on male recipients that can be helpful for pretext [57].

Meanwhile, some researchers think of gender as a proxy for other effective variables. Redmiles et al. [95] in a related study categorizes the spam topics in three broad categories of sales-oriented (e.g., clothes for sale), media (e.g., photos and videos), and interactive (e.g., games). They found that sales-oriented spam messages were mostly viewed by females and media spam by men. Authors believe the fact that most of the sales-oriented spam featured female products or explicitly targeted females and the majority of the media spam was porn or violent content, which is more desirable to men, influenced the result. Therefore, they conclude that gender is a facilitator considering the spam message's content rather than being the true reason for clicking on it by itself.

2) Age: Studies regarding this subject show contradictory results that some of them have found younger adults the most vulnerable group to the phishing among other groups [57], [100], [85] while others have stated older adults are more susceptible [86], [70], [95], with older women as the most vulnerable group.

Oliveira et al. [86] state that one possible reason for this contradiction is that studies supporting the younger adults' higher susceptibility to the older participants have considered the middle-aged group and not older adults. They argue that adults aged around 55 years benefit from high processing speed and working memory, united with a high level of experience. On the other hand, general cognitive processing capacities, and sensitivity to deception and untrustworthy information decline with age, whereas perceived trust increases, resulting in a higher probability of falling for phishing attacks in older adults. It is also mentioned that older adults have been especially unaware of their susceptibility to the phishing attacks where there was a huge difference between their self-reported susceptibility awareness and their behavioral susceptibility to the phishing emails. However, Sheng et al. [100] argue that the vulnerability of the younger adults is because of the lower level of education, fewer years on the Internet, less exposure to training materials, and lower risk aversion. These factors can cause a younger adult to become vulnerable to phishing messages and fall for them, making them a susceptible group as well.

Further, [86] finds that younger users fall more for phishing emails that use the "scarcity" principle of influence [19]. In comparison, older users are more likely to succumb to the ones that use reciprocity and liking. Furthermore, it has been found that teenagers have a higher tendency to include and share their personal information online, such as their photo, location, or phone number [37].

3) Level of Internet usage: The user activity level on the Internet, and its effects on phishing susceptibility, is two-folded in the literature. Redmiles et al. [95] found that more

active Facebook users were less likely to click on spams and also in different studies [7], [6], more elapsed time since the user's Facebook membership was related to the lower phishing susceptibility. In one study, more frequent online purchasers were more successful at identifying phishing websites [94], and in another one, participants with higher Internet usage were more aware of its risks [45]. Ögütçü et al. [85] discuss, higher usage results in more exposure of the user to crime or negative experience, and therefore higher threat perception. Also, in other studies, participants who had higher computer usage in general, significantly performed better in phishing susceptibility [89], [56].

On the other hand, higher usage of each platform, or the Internet in general, results in habit creation that users perform automatically and unconsciously with minimal mental effort [35]. Habitual usage triggers heuristic processing and decision making that is defined by Eagly and Chaiken [29] as "a limited mode of information processing that requires less cognitive effort and fewer cognitive resources" and is also referred to as "rule of thumb". In this mode, users fail to pay attention to the deception cues and act based on their intuition [114], [117]. In one study, habitual email use was highly associated with the higher likelihood of responding to phishing emails that the authors mention automatic response processing as the cause of that rather than simply having more trust in those emails [121]. Another study [11] found that users who had a high frequency of photo-sharing were more likely to act automatically and reflexively without thinking properly before deciding to share photos and were more likely to share more photos in the future, regardless of jeopardizing other people's privacy.

Therefore, it can be inferred that higher internet usage, in general, heightens users' awareness. However, habitual usage can be of great importance when other parts of the phishing message are well-crafted enough that do not trigger the target's systematic thinking. In these scenarios, habitual usage of the platform, that the phishing message is delivered on, like a specific social networking site or the company's mailbox, can act in favor of the attacker when the message fits the user's expectation and the essence of that platform.

4) Education: Education has been proved to have a positive relationship with Internet skills [47], information security awareness [85], lower preference for clickbait [70], and being one of the causes of younger adults' susceptibility to phishing [100]. Moreover, knowledge of the target's educational background can help the attacker access a rich source for a more personalized pretext.

5) Computer security literacy: Digital literacy, in general, has been proved to play an important role in the users' online behavior. The less digitally literate users have a higher tendency to fall for the sources that are not credible [70], are more inclined to judge the accuracy of the source based on their intuition, or are not able to do so [81], and have a higher chance of responding to a phishing attempt [40].

Regarding the focus of this study, computer security literacy has received special attention. Phishing awareness strikingly influences the user's phishing detection ability [94] and their perceived protective practices and reactions [48]. Also, in different studies, less knowledgeable users about phishing were more vulnerable to it [80], [7], [6], and users that were able to define phishing correctly, were less susceptible to the phishing attacks [28].

6) Previous victimization: The previous victimization consists of the user's experience of being phished and encountering phishing attempts where a strong positive relationship has been found between these experiences and the capability of phishing websites identification [94]. Past experiences make users better execute cybersecurity measures by positively affecting their overall awareness and risk perception [60]. In a study [120], users with exposure to more spear-phishing had a higher awareness of its underlying risks and better knew how to handle and report those emails. In another study, some users mentioned such experiences, like catching a virus after clicking on a link, as a cause for their cautious behavior [14].

7) Information Security Awareness: Information security awareness concerns "individuals' knowledge of what policies and procedures they should follow, their understanding of why they should adhere to them (their attitude) and what they actually do (their behavior)" [76].

Higher information security awareness causes users to judge based on the systematic, deliberate processing rather than heuristic mode [48], have a higher detection rate of spam emails [17], and less inclination towards risky decisions [76]. Information security awareness' education and training for users have been emphasized as effective measures against phishing attacks in different studies [101], [30].

8) Training: The role of training has been emphasized by different studies, as one of the most effective measures and the key mitigation strategy for the phishing attacks for each organization. These effects are regarding increasing individuals' information security awareness, improving users' protective behaviors, and preventing users from falling for the attacks [80], [74], [85], [101], [30], [48].

Different kinds of training, such as web-based, contextual, and interactive games, can heighten the user's phishing detection ability [100]. However, although training has proved its usefulness for increasing users' awareness, it has been found that its effect starts decreasing after one month [116]. Therefore, the need to have regular refresher training for users to remember their training should not be neglected [80].

9) The Big Five: It is believed that different personality traits affect various aspects of a person, like beliefs, decisions, and behavior, that can be good predictors of related outcomes, such as happiness and life choices [78], and hence their vulnerability to social engineering attacks [80]. Due to its consistency across time, culture, and age groups, the common Big Five personality model [26], also known as the five-factor model (FFM), has been the most widely used model among the researchers to distinguish different personality traits [36].

The Big Five model consists of five main personality traits, namely `neuroticism`, `extraversion`, `openness`, `agreeableness`, and `conscientiousness`. Possession

of these traits and their rate in the person's personality can be a good indicator of that person's future behavior. However, measurement is a main problem in personality [96]. It needs observation of the person's behavior in different times and contexts [75] that is a hard task to achieve in the context of open-source intelligence gathering. Each of these traits consists of different facets and correlated trait adjectives. Different measurements have been proposed to use these facets, in order to measure the main Big Five traits, such as the Revised NEO Personality Inventory (NEO PI-R) [21] and the Big Five Inventory (BFI) [62].

a) Neuroticism: Neuroticism, also known as emotional instability, is the tendency to experience negative emotions, such as anger or sadness, easily, which causes the person with a high amount of neuroticism not to be able to handle stress appropriately, think clearly, and make decisions [97].

In a study [31], a high level of neuroticism found to have a relationship with the lack of ability to detect lies which the authors believe is because neurotic people will be more upset when facing sad situations and tend to believe that others are truthful. Neuroticism is also related to the probability of answering phishing emails in another study [45]. Nevertheless, some researchers believe in the opposite, that the higher neuroticism causes lower trust in others [5] and fear of being held responsible for the outcomes of the security incident [118], and therefore causes lower vulnerability to the phishing attempts.

However, the high amount of sensitivity to the threats [50] and inability to properly manage stressful situations in neurotic people can make them vulnerable to well-crafted phishing attempts when the suspicious person cannot find clear evidence of the ongoing attack.

b) Extraversion: A higher level of extraversion is related to more inclination towards being in the other people's companionship and is associated with characteristics, like sociability or activity [97], [80]. Different studies have shown relationships between higher extraversion level and higher phishing vulnerability [69], [68], [23], [9].

c) Openness: Openness is the tendency to try new things and experiences without anxiety accompanied by intellectual curiosity [80]. Studies regarding openness have shown that high level of openness is related to higher phishing vulnerability [9] and having less strict privacy settings while posting more on Facebook [45].

d) Agreeableness: Agreeableness is the tendency towards altruism, sympathy, and willingness to help rather than being competitive and egocentric [97]. Although it has been found that higher agreeableness is related to a higher information security awareness [76], various research has shown higher agreeableness results in more phishing vulnerability and being at a high rate of security risk for the possessor[88], [23], [80].

e) Conscientiousness: Conscientiousness consists of traits such as self-control, organizing, and determination [97], which causes the conscious person more likely to obey the security guidelines and training [80], [23]. Higher agreeableness is related to higher information security awareness and inclination to less risky behavior [76], [91]

10) Mood: Researchers argue that the positive mood gives a sense of security to the person that the environment is safe and triggers a low level of cognitive effort, hence giving the possessor confidence. In contrast, the negative mood indicates that the environment is unsafe and needs a higher level of cognitive resources that makes the possessor doubtful. Therefore, positive mood activates heuristic processing, and negative mood sets off careful, systematic processing, which causes the target to be more aware of the phishing cues [24], [102]. The higher impulsivity and inertia of the positive mood possessors are emphasized in the study by Huang and Kuo [55]. Also, in another study [34], participants with a happy mood were more gullible, while sad mood had caused the possessors to be more skeptical and more accurate in the deception detection.

Some theories have been stated and analyzed regarding mood maintenance and mood repair. These theories state that people try to maintain their positive mood or try to repair their negative mood and change it to positive; however, further research is still needed in this domain, especially considering phishing susceptibility. [84]

Martin et al. [73] argued that although positive mood, in general, produces a more favorable result, mood congruency is more important. In their studies, subjects with positive mood had done better in the evaluation of the targets with positive mood role fulfillment (like a laughable comedy), and subjects with negative mood had done better with negative mood role fulfillment (like a sad tragedy). When the mood of the target is known, tailoring the emotional tone of the message based on that can increase the success likelihood as it 'feels right' for the recipient, especially when other factors leading to thinking is kept at minimum [93].

11) Work experience: Work experience includes attributes associated with the person's present and previous jobs, such as expertise in the job, years of functioning in each role, gathered skills, and job description.

In an interesting study on pilots [119], a higher level of experience was related to the more effective cognitive strategy in the decision making of the pilots, and while expertise was helpful in general, task-specific expertise found to be the most effective factor in decision making. This emphasizes two aspects; First, the importance of the previous victimization and computer security literacy as the specific phishing relevant expertise that help users to use their cognitive resources more effectively in making decisions when facing phishing attempts in general. Second, as have been seen before, many phishing emails were particularly pertinent to the person's job, tried to deceive the user into executing a job-related demand, like asking an accountant to transfer money to an account, or use it for pretexting, such as impersonating an authority. In these scenarios, the user's work experience can be a major aid for better reacting to the phishing attempt. Experience becomes a source of information over time [60], helping the possessor make the right decision, which successful experience was mentioned as a chief factor of a CISO credibility [65].

Moreover, information regarding the person's work experience can be a rich source to create various personalized pretexts [109].

12) Years in the current company: Despite the overall working experience discussed, years of service in the current company can be of importance regarding phishing susceptibility as well. Although limited numbers of previous studies have considered this variable, two studies have shown that the employees who were hired for a longer period by a company were less likely to fall for the phishing attempts [16], [63]. The mentioned result can be due to the higher familiarity of the employees with more years of service with the organization's processes, customs, procedures, and rules, which make the employee more aware of the related organizational phishing attacks compared to those employees with fewer years of service and experience in that organization [16].

Another study [120] mentions 'unfamiliarity' with the sender or the topic of the received phishing email, as an aspect that has caused suspicion in their study respondents where new employees that are not familiar with the people that they or their company cooperate with are considered more vulnerable to the work-related phishing attacks.

13) Stress: Stress has been proved to be related to lower performance, attention or memory deficits, higher task error rate, errors in judgment, narrowing visual attention, and reduced cognitive resources that all make users vulnerable to phishing attacks [80], [103]. Furthermore, stress causes the tunneling effect that results in focusing on the main task and decreasing the attention on peripheral information [102]. This is especially important in phishing that peripheral information plays a key role in phishing detection by the user.

Different kinds of daily stressors exist, such as heavy workload, time pressure, email load, resource constraints, poor management, noise, and fatigue [80], [64], [102], [20], [42]. Mentioned stressors make the person switch to the faster heuristic mode due to the diminished available cognitive resources when processing emails and hence more vulnerable to phishing [103].

In different studies regarding the relationship between emails and stress, participants who were asked to turn off their emails experienced a lower stress [72]. In another study, even participants who had limited their email checking frequency were observed to have lower stress [66], and higher daily time spent on emails found to have a relationship with higher stress and lower productivity [71]. Furthermore, in a study associated with workload [59], a higher workload was related to a higher probability of non-compliance behavior.

14) Role: Despite being invaluable information to be used in the pretext, studies related to the target's role in the organization proved that some roles are more vulnerable to phishing. In a study [101], employees from the call center, management, and HR/legal function groups were among the most susceptible groups to the phishing attacks with the rate of 41.8%, 18%, and 9.8% respectively.

Effects of stress on phishing vulnerability has been discussed previously; therefore, roles that are more stressful in their nature and roles that have a periodical high peak of workload and stress, like accountants near the end of the fiscal year, can be good indicators of susceptibility as well. As an example, in a study [120], one of the call center employees has mentioned receiving 200-300 emails per day, making it hard for them to distinguish valid emails as opening the attachments is an essence of their job. Some effective factors in the relationship between organizational roles and the phishing vulnerability are heavy workload, long working hours, time pressure, resource constraints, role overload, poor management, lack of support, and poor communication [64].

Another study regarding the email and cognitive overload [71] shows that the social pressure from peers to respond quickly is more when a person is higher in the organizational hierarchy and hence causes cognitive overload for the person.

15) Risk aversion: Risk aversion causes the person to be knowingly inclined towards choices and decisions that contain less risk, which has been shown that the higher level of risk aversion is related to lower susceptibility to phishing attacks [100]. Additionally, risk perception has been proved to be effective in the users' behavior that in a study [114] users' awareness of the underlying risk of their actions caused them to process the email systematically. At the same time, heuristic processing was triggered when the action was perceived as safe, with no risk.

Risk aversion increases by age that younger adults reported having the lowest risk aversion [3], [27], and teenagers were more willing to share their personal information, like the photo, phone number, or exact location in social medias [37].

16) Culture: In the importance of the culture, Godson and Wirtz [38] mention that understanding the target's culture strengthens the deception, and in order to be successful in the deception, "the deceiver must recognize the target's perceptual context to know what (false) pictures of the world will appear plausible". Furthermore, studies have shown that there is an association between an individual personality and cultural factors [53].

Cultural differences can cause various reactions towards the phishing emails. Individualism culture, which is higher in countries like the US, means people give higher priority to their concerns than their groups [77]. Butavicius et al. showed in their study [17] that respondents from countries with a lower level of individualism were less likely to identify phishing emails. The authors argue this can be because of their higher motivation for keeping the group harmony and thus responding to others' requests.

Also, feminine cultures proved to promote risk-averse attitude [49], which can be a good indicator of lower phishing vulnerability. In a related cultural experiment [33] Swedish employees' behavior was significantly correlated with their general information security awareness and intention to resist social engineering, compared to US and Indian employees. Authors suggest that it was the outcome of their less individualistic, feminine culture that promotes moderation.

Another effective cultural factor is Power Distance (PDI), defined as "the extent to which the less powerful members

of the organizations and institutions accept and expect that power is distributed unequally" [54] where individuals expect the higher individuals to tell them what to do [54]. In a study [16], individuals from cultures with higher PDI proved to be more susceptible to phishing attacks where the experiment's email contained a manager signature and was asked them to follow a guideline.

In a more general view regarding Hofstede's dimensions of culture [52] and their association with the Big Five traits, it has been shown that Power Distance was related to low Extraversion and Openness, and high Conscientiousness; Uncertainty Avoidance was associated with high Neuroticism and low Agreeableness; Individualism with Extraversion and Openness; Masculinity with Neuroticism and Openness. Furthermore, McCrae has created a map based on which Asians and Africans are generally considered as introverts and Europeans as extraverts, with Europeans having a higher rate of openness and a lower rate of agreeableness and conscientiousness. [77]

Also, in a study related to the relationship between culture and language, the author states that "language and culture cannot exist without each other" [61] and are inseparable of each other. This can indicate the possibility of using the language, which can be relatively easy to measure from OSINT, as an indicator of the target's culture.

17) Devices: In the present days, technological devices are inseparable parts of each person's life. Whether personally or professionally, every aspect of our lives appears to be intertwined with our devices in a way that makes them look like a part of ourselves.

Different studies have considered the importance of the devices that a person uses in increasing the likelihood of success for phishing attempts. Examples of such effective factors considering mobile devices are hiding or truncating the complete URL in their browsers or their small screens [83], [12], that make it hard for the user to investigate the signs of the illegitimacy of a phishing website, simple user interface for entering the credentials in mobile apps, that allows the attacker to develop a similar believable one easier [39], [99], enhanced habituation caused by the device affordances [113], or the owners' feelings of trust in their mobile devices [12].

The knowledge regarding the target's used devices in different environments and the time that those devices are used can surely affect different aspects of the attack. Besides, each device has its characteristics, like the operating system, known issues, and built-in applications that can lead to the usage of a more suitable payload for that special device. Additionally, this information can be a good source for the pretext and a better-crafted message.

B. Effective variables in personalization

1) Web platform: Each website and social media has a different nature and essence. Users expect more to see content related to the image of those media they have in mind. In the study by Redmiles et al. [95] on the Facebook platform, the number of clicks on sales spam was twice as the number of media spam. They hypothesize that sales spam fitted the Facebook platform more than media spam that were mostly related to porn or violent content, which users are not expecting to see in such a platform. Conversely, sales spam can be analogous to Facebook's advertisements. Therefore, knowing the platforms that the user has activity on can be truly helpful in creating a more suitable and more believable phishing message.

2) Contacts network: Regarding the importance of knowing about the target's network and social contacts, studies suggest that the chance of the victims succumb to a targeted attack is four times higher if the sender is a known acquaintance. This results in a higher chance of ignoring the critical clues by the recipient of the message and becoming notably more vulnerable [57]. Additionally, the authors noted that some of the participants were not aware of the possibility of harvesting their personal information by somebody else rather than their friends. This lack of awareness can result in a higher success rate for the phishing messages from the target's contacts.

Another interesting study regarding this subject on the Facebook platform [95] showed that spam coming from friends are less likely to succeed as their content may not fit the recipient's expectation of a message from known friends, which emphasizes the key role of the context and the fit between the crafted message and the recipient's expectations. Nevertheless, recipients had a weaker assumption regarding the content coming from friends of friends and expected pages to offer promotional content. Therefore, they had a higher probability of falling for the spam coming from friends of friends or pages; however, they found spam re-shared by the friends of the recipients to have a higher chance of succeeding than when they are re-shared by an unknown source. It is also suggested that being friends of friends will raise the chance of being accepted by the user due to the networking and connecting nature of the social medias [101].

3) Community membership: Membership in a community, whether virtual or physical, can give a great amount of information about the target, such as their interests, skills, beliefs, personalities, or time of presence in a location. These pieces of information can help the phisher to craft a suitable pretext for targeting the person. Furthermore, members of a community generally have similar traits or interests. They are more open to giving out personal information, especially in communities for social causes, that members want to help others [36]. One of the notable aspects of most of these communities is their easy access for the attacker that in a study [101], authors could access a private discussion forum of a company consisted of 1200 employees without any verification.

In addition, Redmiles et al. [95] found that being present in a community with higher spam level causes users to be less inclined to fall for them since this gives them a higher ability in distinguishing trustworthy content from spams in time, which is aligned with the findings in the previous victimization part. They also found that clicking norms in the users' communities affect their individual behavior. Hence, targeting users in countries with high spam click-through rate can lead to a more effective result. One of the possible use cases for this is the

study that authors found a huge difference in the number of compromised accounts with the various country-code top-level domains, with ones from Russia (.ru), 35 times higher than the next country code [18].

4) Residence: Information regarding the user's residence and surrounding, such as the country of living, home address, or organized events in the neighborhood, are considered as high value and can lead to more personalized attacks [98], [30], [90], [37] that some researchers used geographic contexts to improve the believability of their emails [86]. Additionally, as discussed in the community membership part, the user's country of living as a large scale community is important regarding the spam prevalence and clicking norm, where users living in the countries with higher spam prevalence were 59% less likely to fall for them [95].

5) Work place: Just as the residency, knowledge about the user's workplace is valuable to the attacker to use in the pretext, like the company name, location, working hours, or colleagues. [37], [80], [30]

6) Life events: Life events consist of all the happenings in the person's life, such as a newborn baby, attending a seminar, having an accident, and so on. In a study, compliance of the phishing email, regarding an unpaid invoice, with the real-life context of the recipients, that had had a real unpaid invoice among their tasks, was resulted in a high click rate [41]. Such events can temporarily affect contextual relevancy [103] and give the attacker a valuable source for the pretext, hence increase the likelihood of the attack success. Also, email relevancy causes the recipients to ignore some important cues for the phishing detection, like email's source, and become more vulnerable to it [115].

As a recent example, in April 2020, with the Coronavirus pandemic outbreak, Google company stated that it was blocking 18 million scam emails related to the Coronavirus everyday [106]. Correspondingly, a sudden growth of 667% was reported by Barracuda Networks security firm for the relevant phishing attacks by the end of February [36] which is a good representation of using the ongoing events for pretexting.

7) Likes and interests: Knowledge of the targets' likes and interests, such as their hobbies, are important sources for pretexting [104], [90], [109], [30], especially when customized based on the life domains of the target [87]. This can also help enhance the liking of the message as one of the Cialdini's "Principles of Influence" [19], making the target perceives the source as "like me", which holding the same cultural beliefs is an example of [15]. Additionally, it is proved that individuals tend to spend more time with the person that they think is more similar to them [22].

8) Communication norms: Hadnagy in "The Art of Human Hacking" [44] indicates communication style as one of the key factors to successful elicitation. The communication norm is one of the principal parts of the attack implementation. It consists of relevant information regarding the communication aspects of the impersonated entity, that can enhance the message's credibility, which one of the recent attacks on foreign immigrants by impersonating the Dutch immigration service can be a good representative [8]. Some examples of the communication norm that the attacker should consider are how does the impersonated company contact the customers? Do they communicate through email, text messaging, or social media? What are the tone and language of their messages? At what time do they normally send messages? And so on. In a study [120], receiving emails contradictory to the company's communication norm was emphasized by the respondents as a sign for the email illegitimacy. Some examples from the mentioned study are receiving an email on an inappropriate day and time and receiving an external email where the employee normally only receives internal emails from other colleagues.

9) Visual cues: Visual cues, such as logos, images, copyright statements, slogans, fonts, or margins, are another staple factor in implementation for increasing the believability of the message [8]. The correct implementation of the visual cues can even fool the most sophisticated users [25]. These cues are linked to the perceived user trust from the brand, and perfectly imitated ones in a phishing email can stimulate that trust [79]. Moreover, users perceive appropriate visual cues as a sign of legitimacy and emails containing them as more trustworthy and persuasive [121], that users were more likely to fall for phishing emails containing logos [14].

APPENDIX B
USED SURVEY IN THE EXPERIMENT

Figures 13 to 16 show the used survey of the first part of the experiment. The survey is segmented into four parts; however, it was shown to the participants as a single page.

APPENDIX C
OVERVIEW OF THE EXPERIMENT SETUP

Figure 17 shows the overview of the experiment and its procedure.

APPENDIX D
SAMPLE MEASUREMENT STRATEGY FOR THE FRAMEWORK VARIABLES USING OSINT.

Table VIII, provides sample measurement strategy for the variables introduced in the framework by focusing on the possibility of measurement from OSINT.

Welcome to the **Eindhoven University of Technology** survey on **Personality and Email Preference**.
We **do not** collect or store your personally identifiable information and your answers will be used for scholarly purposes only.

More about the HIT

The expected completion time is 7-10 minutes. This assignment has two phases.
The first phase is to answer some (generic) demographic and personality questions about yourself with a focus on your social media.
In the second phase, you will be presented with two emails and need to answer 2 questions based on your preference.
Please note that your provided answers to the open questions are equally important in the evaluation phase of the assignment.

Instructions

This survey is about Personality and Email Preference.
As there are no right answers to the questions, the important factor is thinking carefully about your answers and your honesty.
We strongly advise you to use a laptop or a desktop to fill this survey.
Please note that this survey **does not** work with the **Safari** browser. Therefore, please use other browsers, such as Firefox or Chrome.

Demographic

This section will take **1 minute** to complete.

---

2. Please provide your **age:**

example: 24

---

3. What is the **highest degree or level of school** you have completed?

○ Primary school or lower ○ Secondary school ○ University degree

---

4. **How long** is it that you are working as a worker for Amazon Mechanical Turk?

Note: Please only consider the time that you were actively working.

○ Less than 12 months ○ Between 12 months and 24 months ○ More than 24 months

Fig. 13. The used survey in the experiment (1st segment)

Personality

These questions aim at characterizing how somebody looking at your social media feeds may perceive you in terms of your personality traits.
This will take from **2 to 3 minutes** to complete, and is completely anonymous.

---

1. A person **looking at my social media feed(s)** can infer changes in **my mood**.

○ Strongly Disagree ○ Disagree ○ Neither agree nor disagree ○ Agree ○ Strongly Agree

---

2. I consider myself as an **anxious** person.

○ Strongly Disagree ○ Disagree ○ Neither agree nor disagree ○ Agree ○ Strongly Agree

2.1. A person **looking at my social media feed(s)** can infer that I am an **anxious** person.

○ Strongly Disagree ○ Disagree ○ Neither agree nor disagree ○ Agree ○ Strongly Agree

---

3. I consider myself as a person who is interested in **Arts**.

○ Strongly Disagree ○ Disagree ○ Neither agree nor disagree ○ Agree ○ Strongly Agree

3.1. A person **looking at my social media feed(s)** can infer that I am a person who is interested in **Arts**.

○ Strongly Disagree ○ Disagree ○ Neither agree nor disagree ○ Agree ○ Strongly Agree

---

4. I consider myself as a person who is interested in **trying new things** (eg a new restaurant, traveling abroad, etc).

○ Strongly Disagree ○ Disagree ○ Neither agree nor disagree ○ Agree ○ Strongly Agree

4.1. A person **looking at my social media feed(s)** can infer that I am a person that likes **trying new things** (eg a new restaurant, traveling abroad, etc).

○ Strongly Disagree ○ Disagree ○ Neither agree nor disagree ○ Agree ○ Strongly Agree

---

5. I consider myself as a sociable person that likes **meeting new people and/or being among a group of people**.

Fig. 14. The used survey in the experiment (2nd segment)

○ Strongly Disagree ○ Disagree ○ Neither agree nor disagree ○ Agree ○ Strongly Agree

5.1. A person **looking at my social media feed(s)** can infer that I am a sociable person that likes **meeting new people and/or being among a group of people**.

○ Strongly Disagree ○ Disagree ○ Neither agree nor disagree ○ Agree ○ Strongly Agree

---

6. I consider myself as a person who oftentimes seeks **exciting activities** (e.g adrenaline pumping activities).

○ Strongly Disagree ○ Disagree ○ Neither agree nor disagree ○ Agree ○ Strongly Agree

6.1. A person **looking at my social media feed(s)** can infer that I am a person who oftentimes seeks **exciting activities** (e.g adrenaline pumping activities).

○ Strongly Disagree ○ Disagree ○ Neither agree nor disagree ○ Agree ○ Strongly Agree

---

7. I consider myself as a **compliant and cooperative** person rather than a **competitive and stubborn** one.

○ Strongly Disagree ○ Disagree ○ Neither agree nor disagree ○ Agree ○ Strongly Agree

7.1. A person **looking at my social media feed(s)** can infer that I am a **compliant and cooperative** person rather than a **competitive and stubborn** one.

○ Strongly Disagree ○ Disagree ○ Neither agree nor disagree ○ Agree ○ Strongly Agree

---

8. I consider myself as an **altruistic** person.

○ Strongly Disagree ○ Disagree ○ Neither agree nor disagree ○ Agree ○ Strongly Agree

8.1. A person **looking at my social media feed(s)** can infer that I am an **altruistic** person.

○ Strongly Disagree ○ Disagree ○ Neither agree nor disagree ○ Agree ○ Strongly Agree

---

9. A person **looking at my social media feed(s)** can infer that, in the **current period**, I am **mainly** in a

○ Negative mood ○ Neutral mood ○ Positive mood ○ They would not find relevant information to infer this

---

Fig. 15. The used survey in the experiment (3rd segment)

10. A person **looking at my social media feed(s)** can infer that in the **current period** I am

○ Stressed ○ Neither stressed nor relaxed ○ Relaxed ○ They would not find relevant information to infer this

Next

Fig. 16. The used survey in the experiment (4th segment)

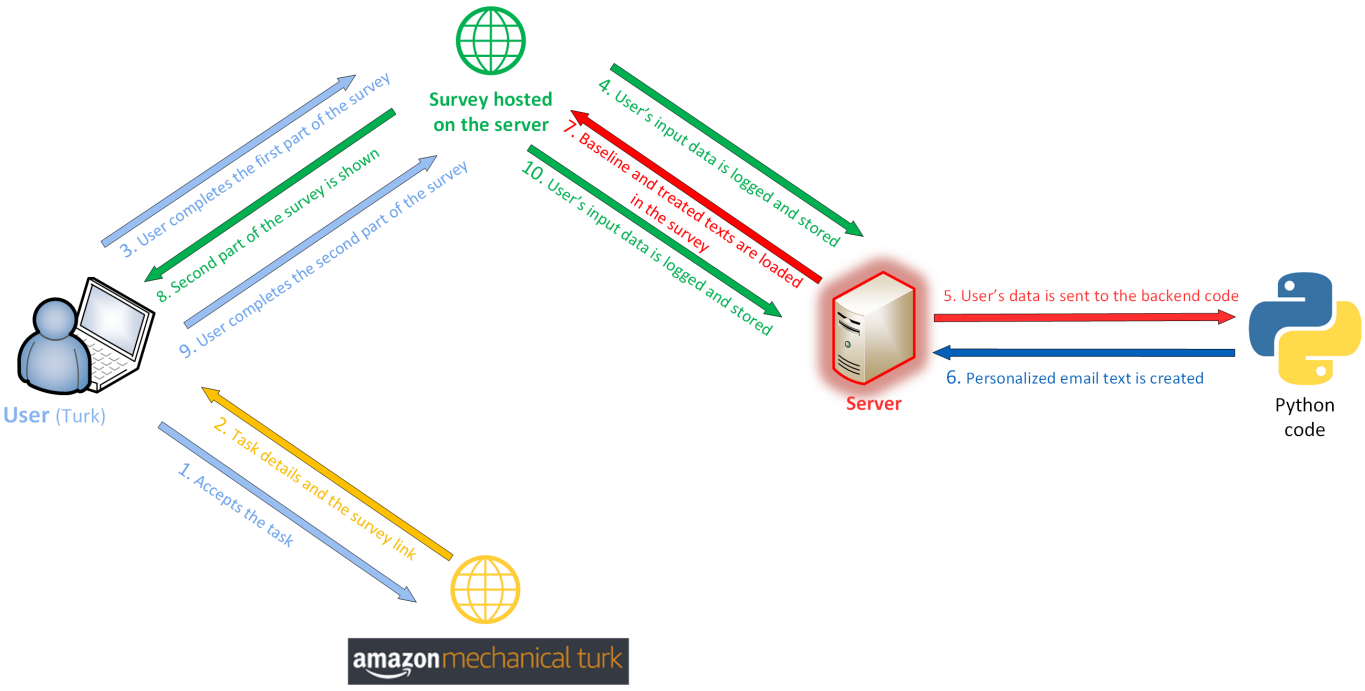

Fig. 17. Experiment setup overview.

TABLE VIII: **Sample measurement strategy for the framework variables using OSINT.**

Variable	Example measurement scenarios
Gender	Social networks, such as Facebook, LinkedIn, etc.
Age	Social networks, such as Facebook, LinkedIn, etc.
Level of Internet usage	Amount of posts, engaging activities, like commenting and liking other posts, etc. in different social networks.
	Online indicator in social networks. An example of such scenario is a tool developed for finding sleep pattern of the friends based on their online indicator (https://github.com/sqren/fb-sleep-stats).
	Increase in the number of posts and followers in a time frame, for example in a week, that can be monitored by different tools, like Visualping or Tinfoleak
	Person's professional role.
Education	Social networks, especially LinkedIn and Facebook.
	Educational institute website and alumni lists.
	Scientific and academic websites, like Researchgate, etc.
	Special events and conferences' websites.
Computer security literacy	LinkedIn's fields of Skills and Licenses and Certificates.
	Person's educational background.
	Previous work experience and present role.
	Privacy settings and security-related behavior in different social media.
Previous victimization	Dedicated tools and websites for leaked credentials, such as 'Have I been pwned?' (https://haveibeenpwned.com/).
	Using Google Dorks.
	Presence in the communities with higher spam level
	Shared content in social networks that contain being hacked or similar incidents.
Information security awareness	Previous work experience and present role.
	Social networks' fields of Education, Skills and Licenses and Certificates for referencing to security and awareness related certificate, training, educational background, etc.
	Knowledge of internal procedures of the person's working company, such as security-related training for new employees, periodic training, awareness-raising campaigns, etc.
	Using internet safe practices and behavior, like social networks' privacy settings.
	Using mentioned proxy variables in the proxy relationship figure, such as previous victimization.
Training	Social networks' fields of Education, Skills and Licenses and Certificates for referencing to security-related training.
	Knowledge of internal procedures of the person's working company, such as security-related training for new employees, periodic training, awareness-raising campaigns, etc.
	Company or its partners' websites, social networks, forums, etc.
The Big Five	Social networks' fields such as 'bio' or 'about' where people write about their personality, opinions, etc.
	Social networks' fields such as 'likes', 'interests', 'volunteer experiences', etc.
	Some social networks have a feature, like Facebook's 'Feeling/Activity' that people use to express their moods and feelings, which are among the best sources to examine the personality of the person, especially by considering posts in a wider period of time. Notably, most of the Big Five facets are present in the list.
	Analyzing the targets' shared contents, like texts, images, or videos, in various social networks in different periods of time with tools such as sentiment analysis, image processing, and manual examination to find the desired characteristics in their personalities.
	Dating websites are among the richest sources since they are the place where people are open to talk about themselves and their personalities.
	Analyzing target's membership in different groups, websites, forums, special social networks, etc.
Mood	Analyzing the social networks posts' mood and feelings fields.
	Analyzing the content and tone of the shared posts manually or through sentiment analysis.
	Life events that can affect mood.

Table VIII Continued:

Work experience	Social networks, especially LinkedIn and Facebook.
	Search engines.
	Company or its partners' websites, social networks, forums, etc.
Years in current company	Social networks, especially LinkedIn and Facebook.
	Search engines.
	Company or its partners' websites, social networks, forums, etc.
Stress	Person's professional role where some roles are stressful by their nature, contain high amount of workload or email load, and time pressure.
	Person's life events when stressful event has occurred, such as losing someone, experiencing a breakup, having an exam, etc.
	Usage of indicators of stress in social networks' posts for expressing moods and feelings. It should be noted that 'stressed' is one of the presented options on the list.
	Content of the shared posts and searching for referencing to stressful events, using stressed tone, etc.
	Some professional roles have periods with higher stress level, such as accounting in the end of each fiscal year.
Role	Social networks, especially LinkedIn and Facebook.
	Search engines.
	Company or its partners' websites, roster page, social networks, forums, etc.
Risk aversion	Social networks' fields such as 'likes', 'interests', etc. to find interests or hobbies that contain risk, like Parachuting, Scuba diving, etc.
	Analyzing the targets' shared contents, like texts, images, or videos, in various social networks in different periods to find risky hobbies or behaviors, such as drug usage, alcohol overuse, driving with high speed, etc.
	Past job experiences or present role can be a good indicator of the person's risk-taking behavior.
Culture	Social networks' fields such as 'languages', 'place of birth', 'living place', 'bio', 'about', etc.
	Language of the posts.
Devices	Shared photos' metadata.
	Metadata of the available documents on company's website, groups, forums, etc.
	Publicly available tools that identify the devices used for posting a content on social networks.
	Social networks' shared contents, such as photo showing the laptop, sharing excitement for the new phone, selfie photos in the mirror, etc.
	Predefined email signatures of some devices.
Web platform	Presence of the target in the social networks and adequate level of activity.
	Company's norm of the target that enforce usage of specific platforms.
Communities membership	Social networks' profiles that show membership in the groups.
	Social networks' 'likes', 'interests', 'following', and similar fields.
	Analyzing the company's website, forums, groups, etc. to find relevant organizational groups.
	Search engines.
	Social networks' shared contents, such as photo of a club, or mentioning the name of a community, etc.
Contacts network	Social networks' connection list, followers, following, or similar fields.
	Fields like 'teammates' and 'reports to' in work-related social networks.
	Company's website, roster page, forums, groups, etc. to find other related employees.
	Analyzing the shared contents in the social networks to find mentions of names, media contains another person, etc.
Communication norm	Emails, chats, shared contents in social networks, groups, website, and other similar options are reach resources to analyze the communication norm in different occasions.
	Company's newsletter, promotional messages, or other kinds of engaging messages.
Residence	Social networks' 'living place', 'current city', or similar fields.
	Using geotags of shared posts.
	Shared photos' metadata.
	Likes and interests that contain specific places to a neighborhood.

Table VIII Continued:

	Membership in the communities of a specific physical location, like the gym in the neighborhood or residents of a building.
Work place	Company's website.
	Search engines.
	Shared photos' metadata.
	Companies directories websites.
Life events	Shared contents in the social networks of the target, family members, or close friends.
	Events that are happening in the target's neighborhood, city, or even globally.
	Company and partner's website, social networks, etc. for professional events.
Likes and interests	Social networks' fields such as 'likes', 'interests', 'volunteer experiences', etc.
	Social networks' shared contents, such as showing the person in a sport, talking about the hobbies, anticipation for a new movie, etc.
	Shared reviews and ratings in social networks.
	Liking the posts related to a page, product, etc. or content and tone of the comments on those posts.
	Membership in the communities.
	Tools that tracks the behavior of the person based on the IP address, such as downloaded torrents.
	Using geotags to track the presence of the person in special places.
Visual cues	Emails, chats, shared contents in social networks, groups, website, and other similar options are reach resources to analyze the used visual cues in different occasions.
	Company's newsletter, promotional messages, or other kinds of engaging messages.