Dogana

ADVANCED SOCIAL ENGINEERING AND VULNERABILITY ASSESMENT FRAMEWORK

D 4.2 DOGANA-Model Version 1

Work Package:	4
Lead partner:	AIT
Author(s):	Michaela Reisinger, Marc Busch, Peter Wolkerstorfer
Submission date:	
Version number:	1.0 **Status:** Final

Grant Agreement N°:	653618
Project Acronym:	DOGANA
Project Title:	Advanced Social Engineering and Vulnerability Assessment Framework
Call identifier:	H2020-DS-06-2014-1
Instrument:	IA
Thematic Priority:	Trustworthy ICT
Start date of the project:	September 1st, 2015
Duration:	36 months

Dissemination Level	
PU: Public	
PP: Restricted to other programme participants (including the Commission)	
RE: Restricted to a group specified by the consortium (including the Commission)	
CO: Confidential, only for members of the consortium (including the Commission)	✓

Revision History

Horizon 2020 European Union Funding for Research & Innovation

Project co-funded by the European Commission under the Horizon 2020 Programme.

Revision	Date	Who	Description
0.1	05.11.2015	Marc Busch	First Draft
0.2	18.11.2015	Marc Busch	Implementation of comments by Yung Shin Van der Sype
0.3	20.04.2016	Michaela Reisinger	Update of Susceptibility Factors, Interrelation
0.4	28.04.2016	Michaela Reisinger	Input Preliminary Studies
0.5	02.05.2016	Michaela Reisinger	Review wording, consistency
0.6	10.05.2016	Michaela Reisinger	Amendments/improvements to several parts according to the input of Enrico Frumento
0.7	12.05.2016	Michaela Reisinger	Revisions according to review by Filipe Custódio
0.8	13.05.2016	Michaela Reisinger	Appending of the legal and ethical checklist
0.9	13.05.2016	Michaela Reisinger	Finalization, Input for ethical checklist by Ioana Cotoi

Quality Control

Role	Date	Who	Approved/Comment
Project Partner	12.05.2016	VIS / Filipe Custódio	Approved
Scientific/Technical Coordinator	10.05.2016	CEFRIEL / Enrico Frumento	Approved

Disclaimer:

This document has been produced in the context of the DOGANA Project. The DOGANA project is part of the European Community's Horizon 2020 Program for research and development and is as such funded by the European Commission. All information in this document is provided "as is" and no guarantee or warranty is given that the information is fit for any particular purpose. The user thereof uses the information at its sole risk and liability. For the avoidance of all doubts, the European Commission has no liability is respect of this document, which is merely representing the authors' view.

Table of Contents

List of Figures

List of Tables

DEFINITIONS AND ACRONYMS

Term	Meaning
CYBERCONNECTOR	An internal knowledge collaboration site and social network that is used to share all the information among partners.
CC	Cyberconnector
MST	Management and Support Team
SC	Scientific Coordinator
DOW	Description of Work
VCS	Victim Communication Stack

Executive Summary

Task 4.2. consolidates the knowledge gathered in Task 4.1 into a single psychological social engineering model. By screening existing psychological models, DOGANA will develop the DOGANA Social Engineering Model: a “Human centric framework for SE 2.0 attack defense, and prevention”. This model will deliver the basis DOGANA awareness methods and the DOGANA tools (design and implementation strategies in Tasks 4.3 and 4.4). It will be evaluated during the field trials (WP7) and iterated based on the field trials results.

This deliverable reports on the outcomes of Task 4.2 and consists of the first version of the underlying psychological model of DOGANA. It will be updated to a second version after evaluation.

1 EXISTING PSYCHOLOGICAL MODELS

1.1 Susceptibility Models

A number of previous models are concerned with different strategies and concepts employed in phishing to influence the psychological susceptibility of individuals. This includes Cialdini's six persuasive strategies (Authority, Social Proof, Liking, Commitment and Consistency, Reciprocity, Scarcity [1]), Gragg's seven psychological triggers (Authority, Diffusion of Responsibility, Deceptive Relationships, Integrity and Consistency, Reciprocation, Overloading, Strong Affect [2]), Stajano and Wilson's seven principles of scam (Social Compliance, Herd principle, Deception, Dishonesty, Need and Greed, Time, Distraction [3]) and Ferreira, Coventry and Lenzini's combined model of Principles of Persuasion in Social Engineering [4]. Their model includes five principles that combine Cialdini's. Gragg's and Stajano and Wilson's work (Figure 1): Authority, Social Proof, Liking, Similarity and Deception, Commitment, Reciprocation and Consistency, Distraction. For details see D.4.1.

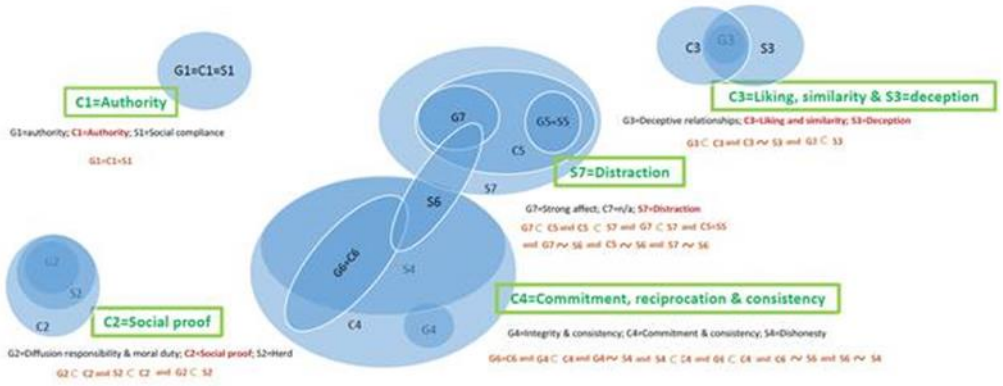


Figure 1 – A mapping of Cialdini's principles of influence [1], Gragg's psychological triggers [2] and Stajano and Wilson's principles of scam [3] by Ferreira, Coventry, & Lenzini [4].

Especially the effectiveness of Cialdini's persuasive strategies [1] have been linked to personality traits [5], gender and age [6], and masculinity and femininity [7], which is an interrelation important to phishing susceptibility.

Kim et al. [8] additionally describe using rational, emotional and motivational appeals, which they link to Cialdini's persuasion principles [1]. These will evoke different responses and influence psychological processes as well as coping mechanisms.

1.2 Information Processing Models

The Theory of Deception [9] focusses on the individual and subjective process of recognizing deception, which relies on prior knowledge and spans four stages: *activation*, *deception hypothesis generation*, *hypothesis evaluation*, and *global assessment*. Another model concerned with information processing is the Elaboration Likelihood Model (ELM, [10]), which distinguishes processing per central or peripheral route: Cuing the peripheral

route is of interest in social engineering, because the peripheral route – in contrast to central processing – does not consider individual aspects of a message, but focusses on specific, pre-definable cues. Decision-making processes have been also described via O-S-I-R models [11], [12]: *O* represents personality characteristics, motivational states and cultural factors, which influence an individual's interpretation of a message, *S* the signal, *I* the interpretation (including motivational and experiential factors), and *R* the response.

Phishing has previously been analysed in the context of information processing [13], yet few aspects have been individually studied so far: [13] included e-mail load as personal characteristic (O), the individual's level of involvement, domain specific knowledge, and technological efficacy as interpretation aspects (I), as well as leakage cures as Stimulus characteristic (S). They show that certain cues designed to short-circuit evaluation process (e.g. by invoking fear) are used, how motivation can increase the level of attention, how attention decreases under stress and in habitual use, and how knowledge influences the ability to effectively elaborate. Their information processing model of phishing susceptibility was able to explain 46% of the variance in individual phishing susceptibility. While this model is thus a very interesting one, it has not been combined with more factors of individual susceptibility. These factors, elaborated in D.4.1., are the basis of the DOGANA social engineering model.

1.3 Technology Acceptance Models

Technology Acceptance Models (TAM, TAM-2, TAM-3) and the Unified Theory of Technology Acceptance (UTAUT) will be considered in the further steps of awareness methods and tool generation. They are especially important to consider processing in combination with educative tools.

2 DOGANA Social Engineering Model

"Human centric framework for SE 2.0 attack defense, and prevention"

D.4.1. elaborated on several factors for susceptibility and social engineering design. Building on human attack vectors such as personality, susceptibility, personas and connected attributes (e.g. semantic and syntax), this document will present the *DOGANA Social Engineering Model* (DOSE-Model) of factors that contribute to the social engineering situation and outcome from the victim's perspective of view. The DOSE-Model thus takes additional factors into account, which are not attack- but victim-specific, to explain social engineering success and failure.

The Human Attack Vector (D.4.1.) shows the use of certain vulnerabilities in social engineering. These vulnerabilities are exploited *by design*, and form part of the victim's *state* during attack. Table 1 shows their connection to the factors of the DOSE-Model.

Social engineering and phishing specifically, uses several leverage points to succeed as an attack – psychological landmarks that can be triggered and exploited:

Individual factors include age [14]–[20], gender [14]–[19], [21]–[25], cultural background and identity [24], [26]–[29], and organizational/social position and network characteristics [30]–[34].

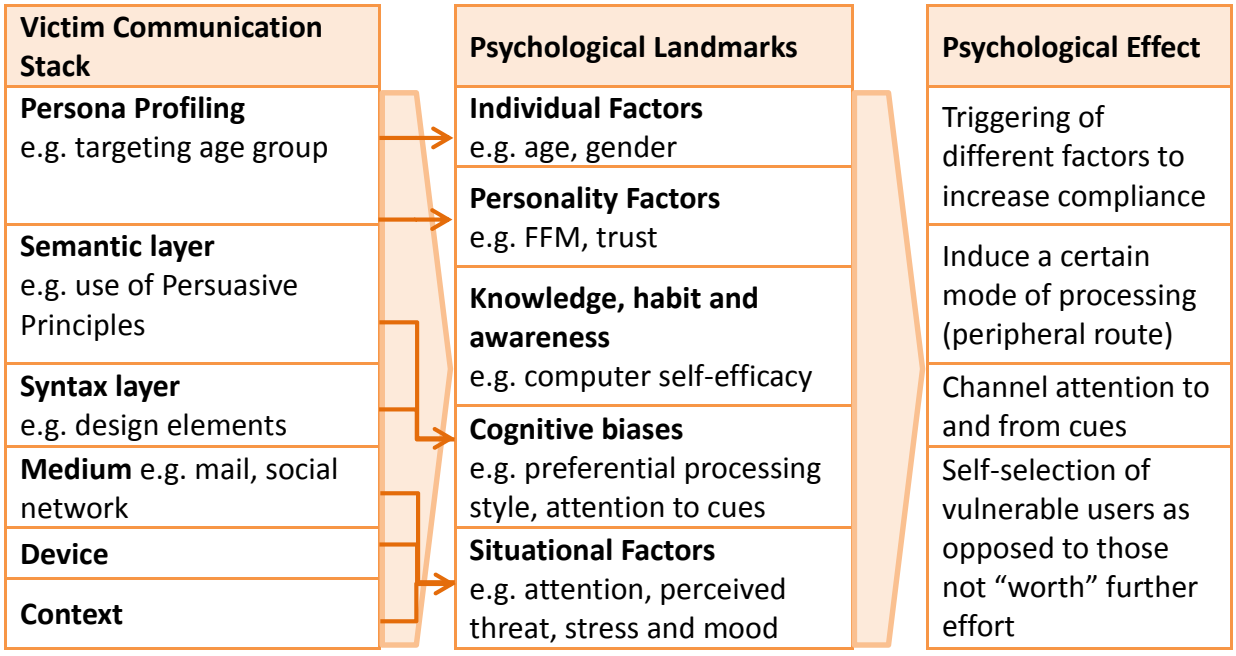
Personality factors include personality traits (FFM; [22], [23], [35]–[39]), trust [20]–[22], [26], [37], [38], [40]–[43], risk behavior [15], [21], [41], [44], helpfulness [21], impulsiveness [37], submissiveness [36], [38], curiosity [45], novelty seeking [45], and personality driven attention patterns [46].

Knowledge, habit and awareness factors describe awareness of, knowledge about and experience with social engineering and phishing [23], [34], [36], [46]–[50]. Technological knowledge [15], [47], [51], and especially domain-specific knowledge, computer self-efficacy [13] and experience [21], [37], internet [19] and e-mail experience [36], [38] and e-mail habits [35] have been shown to be important influencing factors for psychological phishing susceptibility.

Cognitive biases and processing styles are concerned with cognitive capacity and general individual cognitive preferences which can, for example, influence how much attention is paid to e-mail richness [36], [38] and certain triggers [13], and by that modulate elaboration and evolvement of a message.

Situational factors describe the situational modulation of the more general cognitive capacity and attention e.g. via e-mail load [13] and stress. They also include perceived threat/risk and perceived safeguard attributes: Perceived threat [19] is modified through perceived severity and perceived susceptibility, perceived safeguard attributes through perceived safeguard effectiveness and cost [19]. Perceived risk [42], [50] is an important situational factor – it includes perceived voluntariness, immediacy, control, chronic-catastrophic and severity [52]. We note current mood, emotional well-being and well-being at the workplace as important additional situational factors influencing the outcome.

Table 1: The Victim Communication Stack (D.4.1.) viewed side-by-side with the psychological landmarks it targets, and the psychological effects. The individual connections shown are primary connections i.e. the landmarks that are mainly, but not exclusively, targeted by the VCS.



Many of these factors are **interrelated**, and function as (possible) mediators of other factors' effects on susceptibility. The effect of gender on phishing susceptibility, for example, has been shown to be mediated by age, level of knowledge and technical training. Figure 2 shows the interrelation of all factors mentioned in the above literature with an explicit focus on gender and age as most researched factors.

All factors combined form the **state** in which the victim is attacked (comparable with the O-Factor in [13]), and thus all influence the mode of processing and the attention to (specific) cues (I-Factor) while encountering a phishing attack. This affects coping or non-coping, employing or not employing avoidance behavior, and falling or not falling for phishing. For example, leveraging personality factors can tap into a range of emotions, such as guilt or diffusion of responsibility, to increase compliance, which increases the likelihood of falling for the deception. Some of such psychological effects can be seen in Table 1.

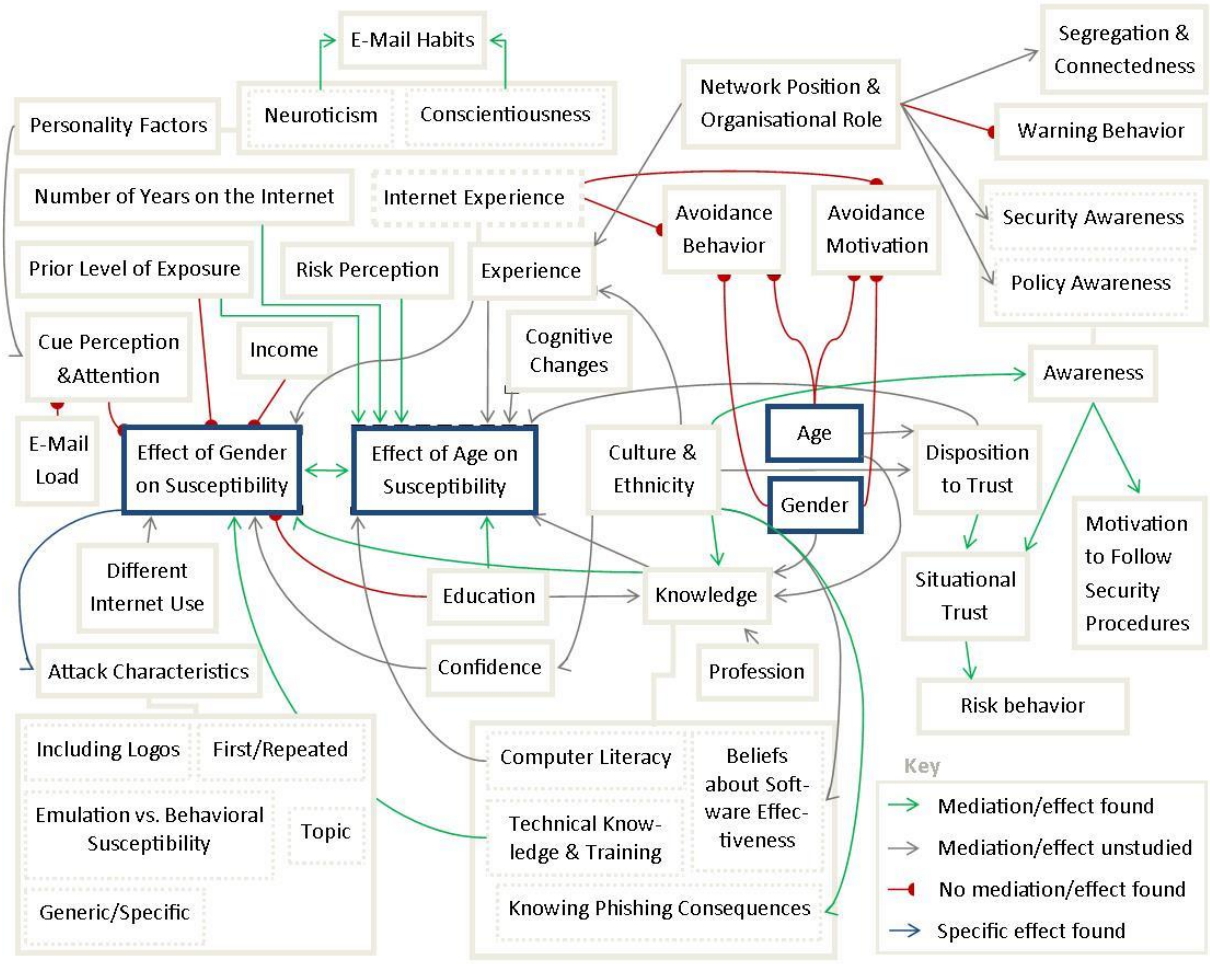


Figure 2 – Interrelation of Factors and Mediations for the Effect of Age and Gender on Phishing susceptibility (based on the literature reviewed in 2). Note circular dependencies e.g. age possibly influencing the disposition to trust, which may mediate the effect of age on susceptibility.

3 PRELIMINARY STUDIES

3.1 A real phishing incident via Skype

In February 2016, 170 people were targeted by a real phishing event on Skype (Figure 3). The individual whose account was hacked created a group to inform people who had been targeted not to open the link. After receiving permission from the hacked individual, AIT was able to post a questionnaire to this group. 11 People responded, five of whom had fallen for the phish.

Figure 3 – Link received by a targeted individual during the Skype phishing incident. Note the personalization of the link, using the target's name.

This real incident provided us with a unique opportunity to collect ecologically valid data - an actual phishing study is insofar superior to phishing emulations (also called *phishing IQs*) and survey data, because it allows an actual look at individual, natural behaviour – including a number of situational parameters, which are usually not present in lab settings. Since this has been a real incident, it also has different ethical implications than a phishing study – participants were not deceived by researchers but only questioned following an incident in their daily life.

3.1.1 Survey

The survey included questions about the phish, about the situation and about their person (Table 2).

Table 2: Questions from the survey sent out after the real phishing incident.

Question Type	Question
Phishing Incident Details	Did you see the link that was posted on [the senders] Skype Account (i.e. the phishing link)?
	Did you see the warning about the phishing link?
	What did you see first? (actual phishing link or warning)[a]
	What were your first thoughts when you saw the phishing link?
	Have you clicked on this link?
	Why did you click on it? / Why did you refrain from clicking on the link? [a]
	Did you at any point realize that this is a phishing link?
	What made you realize that this is a phishing link? [a]
	At what point did you realize that this is a phishing link? [a]
	What did you do after realizing this is a phishing link? [a]
	What could have helped you to realize that this is a phishing link? [a]

Situational Factors and Habits	Overall, how was your mood when you saw the phishing link? (Very unpleasant - Very pleasant) What did you do before you saw the Skype message from [the sender]? How busy were you before you saw the Skype message from [the sender]? (Not at all busy - Extremely busy) Where were you when you were opening the Skype message from [the sender]? (Home/Work/Public Transport/ Other) Approximately, how many messages (e-mail, WhatsApp, Facebook, Skype) did you receive the hour before you were opening the Skype message from [the sender]? How do you usually handle links from Skype contacts?
Individual Factors	Age, Gender, native language

[a] Presence of the question depending on answers on questions above

3.1.2 Results and Discussion

The responses indicated personal and situational factors for susceptibility – three of five of the individuals who fell indicated this being due to their trust in the sender:

"[The sender] is a trusted professor."

"Because I trust anything [the sender] sends me!!!"

"It is impossible someone hacked [the sender]"

Two others noted their habits and attention as reasons for falling for the phish:

"[It fitted a circumstance -] I didn't reflect on it, I clicked it habitually." [Translated from German by the author]

"I did not read carefully what is the address in the link. [sic!]"

The survey also showed that the participants had varying views on how much messages/e-mails are "normal" or "many" for them – such varying perceptions of message load and stress could also differentially contribute to susceptibility.

Of the five people who clicked on the link, only one did not realize it was a phishing attack. This individual noted that they would have realized the attack had there been a message attached which would use different expressions/language than the hacked person normally would. Those who did realize the deception after clicking on the link noted the content of the website, and in one instance, the changing of the URL as clues. Those who did not click on the phishing link additionally mentioned the absence of an accompanying message, the link being unexpected and the awareness of a similar situation. They for the most part reported to have been suspicious but not sure that it was a phish, and confirmed via the warning posted in the group or via asking the sender. Though four individuals reported they usually ask the sender what a link is about when the receive links on Skype, only one did.

Coping strategies for individuals who fell for the phish included closing the website, clearing the history and launching their antivirus. One individual who didn't fall for the phish still cleared the chat in which they received the link; another warned their colleagues of the attack.

Though the sample size is limited, this phishing attack provided AIT with the opportunity of establishing trends that could be important to ascertain in following studies. Trust was a major factor in this small study – while this is an interpersonal event, trust also plays a role in more standard e-mail phishing – mostly in the form of credibility.
Apart from trust, the results also indicate situational factors, which will especially looked at in the following study.

3.2 EmoPhish – Emotional status as susceptibility factor

Situational susceptibility factors are rarely studied so far: While some progress has been achieved in including e-mail load [13] and e-mail habits [35] in the picture of phishing susceptibility, the emotional component provides a stark research gap. This is a major shortcoming to phishing research, since it is very likely the mixture of targeting (randomly via quantity or specific via spear-phishing), context and situational factors which succeed as opposed to only one element. We thus wish to address two possible emotional pathways: 1) To elucidate situational emotional predisposition as deciding factor if people fall for phish and 2) to report on the emotional impact of phishing via legitimate e-mail, which contributes to phishing susceptibility.

For the study, AIT will create an assessment centre-like setting. This will on one hand mask the phishing detection element as e-mail sorting in order to "assess organizational skills", and on the other hand allow us to assess individual characteristics e.g. cognitive style, personality measures via separate "assessment steps". It will also aid to put participants in a situation where they feel involved and put in effort, and will overall provide better immersion than usual phishing emulations. Controlled message load and engagement will additionally elucidate habitual e-mail behaviour.

3.2.1 Methods

Participants will be briefed in the beginning of the study, and debriefed at the end. The briefing will include their first consent to participate in the study, which will be explained to them as a study testing assessment-centre proceedings. They will be informed about the procedure and about any recordings made, specifically about the use of eye-tracking in one of the assignments. The phishing focus of the study will not be disclosed in this briefing.

Instead, they will be introduced to the situation of an assessment-centre and asked to imagine they were really taking part in it, and wanting to succeed to a job interview. This will be done to give some sort of framing to the inbox sorting task and the whole test battery. Giving participants framing is important so they do not wonder about the purpose of the study and focus their attention on recognizing the study purpose. The "assessment centre" will consist of three parts:

In the first part, we will assess four characteristics – need for cognition (18-item, [53]), executive functioning (via the Tower of London Task), field dependence (via an Embedded Figures Test) and processing style (e.g. via OSIVQ, [54]).

In the second part, participants will perform an e-mail sorting task, which also includes phishing e-mails. We will measure time for processing each e-mail, correctness of the sorting and overall processing time. Though these do not constitute our primary measurements, they will be used for comparison with other studies and are necessary for a thorough investigation. We will measure general emotional state and emotional effect via the Noldus® FaceReader (Figure 4) and/or physiological stress measures. An eye-tracker will be used to understand the elements individuals triggering emotional measures. Subsequent to the task they will assess how stressful they experienced the task, how attentive they experienced themselves to be and how the incoming e-mail load felt to them.

Figure 4 – The Noldus® FaceReader in Use (Picture © AIT – Aris Venetikidis)

As a third part, participants will receive a questionnaire to assess personality factors (e.g. FFM, [55]). They will then answer some concluding questions (technical and domain-specific knowledge, self-efficacy, demographics) and be debriefed. In the debriefing they will be informed about the real intention of the study (to capture the emotional component of phishing and to measure cognitive processing styles with the management of phishing e-mails). They will be offered to withdraw their data from the study. More information on deception in DOGANA can be found in section 4 (1. Pre-Study Informed Consent).

4 PLANNING OF USER STUDIES TO DEVELOP PSYCHOLOGICAL MODEL

In WP4, a psychological model underlying social engineering will be developed and validated. In order to do that, AIT will perform several user studies in which companies will be attacked. The goal of the attacks is to determine which factors contribute to the susceptibility to fall for social engineering attacks. Participants in the studies will be the DOGANA partners as well as external companies recruited by AIT (see below the proposed recruiting letter). There will be no specific selection criteria for participants as we aim for a heterogeneous sample to draw meaningful conclusions. We aim at a gender balanced sample and a broad range of age.

All user studies will be performed according to the following scheme (Figure 5):

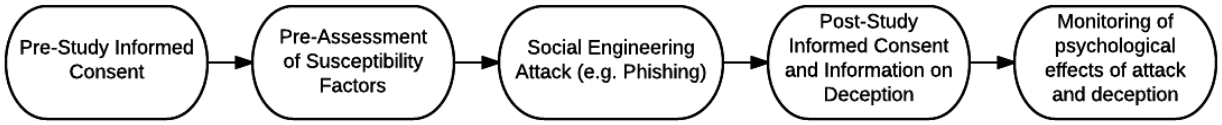


Figure 5 – Succession of user studies within the project.

1. **Pre-Study Informed Consent:**

Participants receive a pre-study informed consent (see below), in which they give consent that they participate in studies around "ICT at the workplace (covering diverse aspects such as health, wellbeing, safety, security, culture". The pre-study informed consent will not disclose information about the real intention of the study (to attack people with social engineering methods and to assess their susceptibility to these attacks and the factors that determine this susceptibility). Disclosing the study focus and intent to participants would compromise the vulnerability assessment taken by the study: Participants would be biased through their knowledge and heightened concern, which in turn would influence their reactions. To avoid such invalid results, we opt to use deception in the pre-study informed consent and to employ a second informed consent with full disclosure at the end of the study.

Deception in DOGANA: In some psychological studies researchers deceive their participants in order to get unfiltered and unbiased results. According to the American Psychological Association[1], deception in research in justified if:

1. The study's significant prospective scientific, educational, or applied value requires deception and effective non-deceptive alternative procedures are not feasible.
2. The study is not expected to cause physical pain or severe emotional distress
3. Psychologists explain any deception as early as is feasible and permit participants to withdraw their data

In DOGANA, we see a high prospective scientific (development and validation of psychological social engineering model) and applied value (design input for awareness methods) (1). We don't expect the study to cause physical pain or emotional distress (2) and we will debrief participants with a post-study informed consent (see below) as soon as possible (directly after the attack) and permit participants to withdraw their data (3).

Some studies show that participants in deception studies vs. participants in non-deception studies "enjoyed the experience more, received more educational benefit from it, and did not mind being deceived or having their privacy invaded" [56]. Deception is a commonly used research method, since the by the 1970s, the use of deception of social psychology studies has reached 50% [57].

[1] http://www.apa.org/monitor/2009/04/ethics.aspx

2. **Pre-Assessment of Susceptibility Factors**: Participants receive a link to an online questionnaire in which we will assess factors that could potentially contribute to the susceptibility to fall for social engineering attacks. Not all factors will be assessed in all studies; their use depends on the concrete goal of the study and will be based on previous studies of social engineering susceptibility. While a number of factors have been previously researched, we will also include several new factors. It is especially important to note that some previously investigated factors have not been studied in this context or in combination with specific attacks – their renewed inclusion is therefore on one hand due to fill this gap and on the other hand necessary to understand the interrelation of susceptibility factors. Some of these factors are sensitive information, thus more strict data protection provision will be applicable.
3. **Social Engineering Attack (e.g. Phishing):** Attacks can be administered via several modes (e.g. e-mail, WhatsApp, face to face, via USB sticks, etc.). This project will focus on e-mail messages as a major distribution vector. The design of the attacks will all be based on collections of persuasive strategies (e.g., [1]), existing phishing examples and analyses of real attacks.
4. **Post-Study Informed Consent:** Participants will receive a written explanation of the actual goal and procedure of the study and will have the right to withdraw their data from the study.
5. **Monitoring of psychological effects**: Participants will be interviewed and assessed with questionnaires sometime after the attack to make sure they are not emotionally stressed. If so, we will suggest aftercare measures. However, the chances that they will be emotionally stressed are estimated to be really low, also empirical studies on deception back this estimation [56].

The data collected in these studies will be hosted on the web space from AIT Technology Experience Department (www.tech-experience.at), which is hosted at all-inkl.com (https://all-inkl.com/). All servers are located in Germany, hence German data protection law applies, which is one of the toughest in Europe. Current state-of-the-art security is guaranteed by our web space provider all-inkl.com. Access is only granted to dedicated AIT personnel involved in the data analysis activities in DOGANA (Marc Busch, Peter Wolkerstorfer, Michaela Reisinger, Peter Fröhlich, Manfred Tscheligi). All collected raw data will be deleted 1 year after the DOGANA project has ended.

5 STUDY MATERIALS AND INFORMED CONSENTS

5.1 Pre-Study Informed Consent

Research project on new information and communication technologies at the workplace

Dear Sir or Madam,

Your employer provided your e-mail address to us – we are the AIT Austrian Institute of Technology (http://www.ait.ac.at/?L=1), a non-profit independent research organization located in Vienna, Austria. Together with our partners we would like to invite you to participate in our research project:

Upcoming information and communication technologies (ICTs) will affect the way we work. New devices, cloud computing and new ways of working (e.g. increase remote work) can either make our lives easier or more difficult. To shape the way how we will interact with ICTs in the future, we perform a series of online and field studies to understand how people interact with ICTs at the workplace and how they are affected by ICTs in aspects related to health, wellbeing, safety, security and culture.

All data that you provide during the studies will be securely hosted at AIT. We guarantee that the data will be analyzed and viewed in an aggregated way only – we will not analyze data of single participants. Only employees of AIT will have access to the data and can analyze the data. The data that we collect will be kept until 2019. For this study, Austrian privacy law and requirements are applicable.

If you'd like to participate click on the "Continue" button below and fill out our short registration form (It will take less than a minute). We will contact you in the near future to invite you to online questionnaires or telephone interviews. You can withdraw from this project at any time without providing any reason. If you like to withdraw your data, or if you have other questions or concerns please confidently contact Marc Busch from AIT (Marc.Busch@ait.ac.at; 0043 664 8894935).

"Continue"

5.2 Post-Study Informed Consent

You have just been phished. Phishing is an attempt to acquire sensitive information of people or organizations (e.g. passwords). This phishing attempt was part of the research project on information and communication technologies at the workplace for which you registered some time ago. In this registration, you were told that we would investigate how information and communication technologies change future workplaces. However, the real purpose of the study was to find out which factors (for example age and knowledge) influence the tendency of people to provide sensitive information.

You have not been told that we would attempt to "phish" you, and we apologize for that. We are sorry we could not tell you the aim of the study beforehand – if we had, you would have been prepared and that would have changed the outcome of the study. We hope for your understanding. With the knowledge that we gather in this study we are able to develop tailored and innovative awareness and education programs (for example digital mini-games) to prepare people for "phishes" and to ultimately work on an improved personal and organizational information security. We now know for example if we should develop tailored programs for specific age groups.

We guarantee that your data will be kept safe and will not be shared with your employers or colleagues. Collected data (e.g. answers to our questionnaire) will be stored in an anonymized way using non-identifiable codes. Identifiable personal data (e.g. the e-mail address we used to contact you) will always be stored separately from the collected data. Only employees from AIT will have access to the data. Your data will be only analyzed in an aggregated way; we will not analyze data of single persons. The data will only be analyzed for the research purpose of developing tailored programs to prepare people against phishing attacks. We will only collect, store and process data strictly necessary for this research purpose.

However, we fully understand if you would like to withdraw your data from the study. If you would like to do so, please indicate here:

( ) Yes, I would like to have my data withdrawn from the study

( ) No, you can use my data for research purposes as stated above.

Either way, if you feel uncomfortable having participated in this study or experience any unpleasant after-effects (e.g. stress), we can put you in contact with psychologists from AIT Austrian Institute of Technology, who will talk to you about your experience and suggest means to deal with it.

( ) Yes, please contact me. This is my preferred way of being contacted: _____

( ) No, I don't like to be contacted at this point.

If you experience unpleasant effects any time after you saw this message, or have any further questions, please note the contact details from Marc Busch, psychologist at AIT (Marc.Busch@ait.ac.at; 0043 664 8894935). He will help you with any questions, problems or concerns related to this study.

If you don't want to have your data withdrawn from the study, we might also contact you at a later point to ask you about your experience after participating in the study.

One last thing: Please don't talk to your fellow employees about this study, it might be that they also participate(d) in this study and that they have been, or will also be "phished". The success of the project relies on keeping the real intention of the study secret for as long as possible – we rely on you to keep it so.

This research is partly funded by the European Union.

5.3 Recruiting letter for external companies participating in DOGANA studies

Dear Sir or Madam,

we are the AIT Austrian Institute of Technology (http://www.ait.ac.at/?L=1), a non-profit independent research organization located in Vienna, Austria. We invite your organization to participate in a research project about organizational information security. Information security is an important asset for companies; however a recent PricewaterhouseCoopers survey on information security (http://www.pwc.com/gx/en/issues/cyber-security/information-security-survey.html) shows that employees are still the greatest risk for information security. A big issue is employees falling for phishing attacks, which are deceptive e-mails and messages asking for sensitive organizational data.

In a research project partly funded by the European Union we are investigating the factors that are responsible for employees falling for such phish. This helps us to design and to investigate effective countermeasures, such as trainings, awareness programs and mini games aiming at educating employees about the risks of phishing. For this study, Austrian privacy law and requirements are applicable.

How could your company participate in this research project?

If you decide to participate, you would provide company e-mail addresses of employees who you suggest to participate in studies. Before each of the following measures, we would provide you with e-mails and documents we planned to send out to the participants, and access to our panel, for review and approval.

We would send an e-mail to the employees you provided us with, telling them that they could voluntarily participate in research studies about the "use of information and communication technologies at the workplace" and that you provided their e-mail address. If they chose to participate, they would register on a participant panel and would be contacted at a later date to fill out an online questionnaire. Some time after that, they would receive a phishing attack sent by us. Immediately after clicking on the "fake" link, they will be debriefed in writing, explaining the real intention and goal of the study with full right to withdraw their data from the study. If they agree, we will contact them after a while to make sure they were not distressed from participating in the study.

We will not tell participants beforehand that the study is about phishing, as this would "warn" them for an upcoming "phishing attack" and would bias the results of the study. We would thus ask you not to disclose more information than necessary to your employees before the study is finished.

What happens with the data that is collected?

We guarantee that your data will be kept safe from third parties. Only employees from AIT will have access to the data. Data about the companies (including name of the company) or employees will not be made public. Data will be only analyzed in an aggregated way; we will not analyze data of single persons. The data will only be analyzed for the research purpose of developing tailored programs to prepare people against phishing attacks. The data will be used for reports and publications only in an aggregated and anonymous way that does not allow any conclusions to single companies or employees.

What is your benefit in participating in the study?

After the study is finished, you will get an aggregated report on the outcomes – tailored to your organization. This will show you weaknesses in your organizations' information security. Based on the results, we will make suggestions on how you could improve your organizations information security through training and awareness programs. Additionally, your company can participate in exclusive (and free) tests of the newly developed training and awareness programs.

If you have any questions, please don't hesitate to contact us!

Best

Marc Busch

(Signature)

6 LEGAL AND ETHICAL CHECKLIST

Table 3: Self-assessment of legal and ethical issues as described in D 1.3. This deliverable follows the checklist for user studies and trials (D 1.3. Section 5.2.) as per Section 6 of D 1.3.

	Risk (as described in D1.3 Section 3)	**Requirement**	**Argumentation**
Stage 1. Preliminary measures	The research results may have a severe negative impact on the human rights of individuals or groups (e.g. privacy, discrimination, stigmatization)	Risk mitigation, such as 1 a human rights impact assessment 2 the involvement of human rights experts in the research 3 training of personnel and/or technological safeguards Risk-assessment 4 details on how the research could affect human rights 5 details on the measures which were taken to prevent abuse	The research could potentially be a danger for the privacy of the employees, as data (e.g. personality or gender in the light of susceptibility to phishing [see factors in section 2]) about them will be collected. However, identifiable personal information will be always stored separately from these data. Furthermore, the data will be stored in a safe and secure way and only a restricted number of people from AIT (Michaela Reisinger, Marc Busch, Peter Wolkerstorfer, Manfred Tscheligi) have access to the raw data.Data presented in publications or to employers will use aggregated data only, in which individuals cannot be identified. If certain groups are identified as more vulnerable to phishing, care will be taken to include and suggest their underlying and mediating factors and use our focus on intervention to prevent discrimination against this group.
	The research has the potential to be abused or misused	Risk-assessment 6 details on the measures taken to prevent abuse 7 if applicable, copies of personnel security clearances Details on the storage and destiny of the research data	Research data will be stored on secure servers and stored only there. Raw data will only be accessible to a limited number of AIT employees (see above), the prevent misuse.

	Incompliance with data protection law	Notification of the processing of personal data to the DPA	The data will be handled according to the Austrian data protection law. AIT has a registered number for processing of research data.
	Incompliance with data protection law > human volunteers	Assessment of the ethical implications of the chosen methodology Description of the recruitment procedures Consent forms Information sheets – describe the goal of the research, the categories of data you are going to collect, the reasoning behind the collection of the personal data, the way how this information will be used and how it will be stored and what will happen with the data after the research	All ethical implications have be considered, ethical procedures for the conduction of studies have been developed, a two-stage consent process is implemented and participants will be informed about research goal, purpose, procedures and their rights.
Stage 2. Research considerations	The research may have a negative impact on human rights	Research methods for correct interpretation of the research results should be provided	Standardized psychological methods (e.g. for assessment of personality) will be used to ensure correct interpretation of results.
	Incompliance with data protection law > human volunteers	Post-consent forms and specified procedures must be provided	A two-stage consent form process is implemented with pre- and post-study informed consent. The informed consent includes information on the types of data (examples given in the deliverable will be expanded according to the study at hand). The data will only be analyzed for the research purpose of developing tailored programs to prepare people against phishing attacks. We will only collect, store and process data strictly necessary for this research purpose.
	Confidential Dogana internal information could be disclosed through the	Caution when publishing or otherwise disseminating those results Compliance with non-disclosure agreements and other	In all publications only aggregated data will be presented. We will comply with non-disclosure agreements with third parties (e.g. New York

	research	contracts in relation to the research data Compliance with the technical partner couples relationships Internal check of the Trials deliverables by the End-Users, according to the general criteria described in D2.5	Universities) and all study and trial deliverables will be internally checked.
	Data loss during the execution of tests	Pro-active protection of the SVA platform against potential intrusion during the execution of tests and before destruction of the data	Research data will be stored on secure severs to prevent data loss.
Stage 3. Post measures	Loss of personal data	Details on the personal data storage assessment, including access control Details on the access control etc. safety measures	Access control will be according to AITs quality management standards (AIT is certified according to ISO standards).
	The research may have a negative impact on human rights	Caution when publishing or otherwise disseminating research results Statement that no data other than the results of the project (software and documentation) will be exported to non-EU Member States.	In all publications, only aggregated data will be presented. No other data other than the results of the project will be exported to non-EU Member States.
	Abundant personal data are stored for an unreasonable time	Erasure of data or anonymisation of the data	All research data will be anonymized. We will create non-identifiable codes for research participants

7 REFERENCES

[1] R. B. Cialdini, "Harnessing the science of persuasion," *Harv. Bus. Rev.*, vol. 79, no. 9, pp. 72–81, 2001.

[2] D. Gragg, "A Multi-Level Defense Against Social Engineering," 2003.

[3] F. Stajano and P. Wilson, "Understanding scam victims," *Commun. ACM*, vol. 54, no. 3, p. 70, 2011.

[4] A. Ferreira, L. Coventry, and G. Lenzini, "Principles of Persuasion in Social Engineering and Their Use in Phishing," in *Human Aspects of Information Security, Privacy, and Trust*, vol. 9190, T. Tryfonas and I. Askoxylakis, Eds. Springer International Publishing, 2015, pp. 36–47.

[5] N. Alkış and T. T. Temizel, "The impact of individual differences on influence strategies," *Pers. Individ. Dif.*, vol. 87, pp. 147–152, Dec. 2015.

[6] R. Orji, R. L. Mandryk, and J. Vassileva, "Gender, Age, and Responsiveness to Cialdini's Persuasion Strategies," in *Persuasive Technology*, 2015, vol. 9072, pp. 147–159.

[7] M. Busch, E. Mattheiss, M. Reisinger, R. Orji, P. Fröhlich, and M. Tscheligi, "More than Sex: The Role of Femininity and Masculinity in the Design of Personalized Persuasive Games," in *11th International Conference on Persuasive Technology*, 2016, vol. 9638, no. October, pp. 219–229.

[8] D. Kim, J. Hyun Kim, and J. H. Kim, "Understanding persuasive elements in phishing e-mails," *Online Inf. Rev.*, vol. 37, no. 6, pp. 835–850, 2013.

[9] P. E. Johnson, S. Grazioli, K. Jamal, and I. A. Zualkernan, "Success and failure in expert reasoning," *Organ. Behav. Hum. Decis. Process.*, vol. 53, no. 2, pp. 173–203, Nov. 1992.

[10] R. E. Petty and J. T. Cacioppo, "The Elaboration Likelihood Model of Persuasion," in *Advances in experimental social psychology*, vol. 19, 1986, pp. 123–205.

[11] H. Markus and R. Zajonc, "The cognitive perspective in social psychology," in *Handbook of social psychology*, 3rd ed., vol. 1, G. Lindzey and E. Aronson, Eds. 1985, pp. 137–230.

[12] J. M. McLeod, G. M. Kosicki, and D. M. McLeod, "The expanding boundaries of political communication effects," in *Media effects: Advances in theory and research*, J. Bryant and D. Zillmann, Eds. Hillsdale, NJ: LEA, 1994, pp. 43–67.

[13] A. Vishwanath, T. Herath, R. Chen, J. Wang, and H. R. Rao, "Why do people get phished? Testing individual differences in phishing vulnerability within an integrated, information processing model," *Decis. Support Syst.*, vol. 51, no. 3, pp. 576–586, Jun. 2011.

[14] P. Kumaraguru, J. Cranshaw, A. Acquisti, L. Cranor, J. Hong, M. A. Blair, and T. Pham, "School of phish," in *Proceedings of the 5th Symposium on Usable Privacy and Security - SOUPS '09*, 2009, p. 1.

[15] S. Sheng, M. Holbrook, P. Kumaraguru, L. F. Cranor, and J. Downs, "Who falls for phish? A Demographic Analysis of Phishing Susceptibility and Effectiveness of Interventions," in *Proceedings of the 28th international conference on Human factors*

in computing systems - CHI '10, 2010, pp. 373 – 382.

[16] J. Lee, L. Bauer, and M. Mazurek, "Studying the Effectiveness of Security Images in Internet Banking," *IEEE Internet Comput.*, pp. 1–1, 2014.

[17] J. G. Mohebzada, A. El Zarka, A. H. Bhojani, and A. Darwish, "Phishing in a University Community," in *2012 International Conference on Innovations in Information Technology (IIT)*, 2012, pp. 249–254.

[18] M. Aston, S. McCombie, B. Reardon, and P. Watters, "A preliminary profiling of internet money mules: An australian perspective," *UIC-ATC 2009 - Symp. Work. Ubiquitous, Auton. Trust. Comput. Conjunction with UIC'09 ATC'09 Conf.*, pp. 482–487, 2009.

[19] N. A. G. Arachchilage and S. Love, "A game design framework for avoiding phishing attacks," *Comput. Human Behav.*, vol. 29, no. 3, pp. 706–714, 2013.

[20] R. Chakraborty, R. Rao, V. Sankaranarayanan, and S. Upadhyaya, "Mediated internet experience for senior citizens," *14th Am. Conf. Inf. Syst. AMCIS 2008*, vol. 3, pp. 1885–1894, 2008.

[21] W. Rocha Flores, H. Holm, G. Svensson, and G. Ericsson, "Using Phishing Experiments and Scenario-based Surveys to Understand Security Behaviours in Practice," in *European Information Security Multi-Conference (EISMC 2013); Lisbon, Portugal, May 8-10, 2013*, 2013, vol. 22, no. 4, pp. 79–90.

[22] K. W. Hong, C. M. Kelley, R. Tembe, E. Murphy-Hill, and C. B. Mayhorn, "Keeping Up With The Joneses: Assessing Phishing Susceptibility in an Email Task," *Proc. Hum. Factors Ergon. Soc. Annu. Meet.*, vol. 57, no. 1, pp. 1012–1016, 2013.

[23] T. Halevi, J. Lewis, and N. Memon, "A pilot study of cyber security and privacy related behavior and personality traits," *WWW 2013 Companion - Proc. 22nd Int. Conf. World Wide Web*, pp. 737–744, 2013.

[24] D. M. Downs, I. Ademaj, and A. M. Schuck, "Internet security: Who is leaving the 'virtual door' open and why?," *First Monday*, vol. 14, no. 1, pp. 1–7, Dec. 2008.

[25] M. Blythe, H. Petrie, and J. a Clark, "F for Fake : Four Studies on How We Fall for Phish," *Chi 2011*, pp. 3469–3478, 2011.

[26] M. Al-Hamar, R. Dawson, and L. Guan, "A culture of trust threatens security and privacy in Qatar," *Proc. - 10th IEEE Int. Conf. Comput. Inf. Technol. CIT-2010, 7th IEEE Int. Conf. Embed. Softw. Syst. ICESS-2010, ScalCom-2010*, no. Cit, pp. 991–995, 2010.

[27] R. Tembe, "Phishing in International Waters Exploring Cross-National Differences in Phishing Conceptualizations between Chinese, Indian and American Samples," pp. 0–6, 2014.

[28] M. Al-Hamar, R. Dawson, and J. Al-Hamar, "The need for education on phishing: a survey comparison of the UK and Qatar," *Campus-Wide Inf. Syst.*, vol. 28, no. 5, pp. 308–319, Nov. 2011.

[29] R. Tembe, K. W. Hong, E. Murphy-Hill, C. B. Mayhorn, and C. M. Kelley, "American and Indian Conceptualizations of Phishing," *2013 Third Work. Socio-Technical Asp. Secur. Trust*, pp. 37–45, 2013.

[30] K. Coronges, R. Dodge, C. Mukina, Z. Radwick, J. Shevchik, and E. Rovira, "The

influences of social networks on phishing vulnerability," *Proc. Annu. Hawaii Int. Conf. Syst. Sci.*, pp. 2366–2373, 2011.

[31] B. M. Bowen, R. Devarajan, and S. Stolfo, "Measuring the human factor of cyber security," *2011 IEEE Int. Conf. Technol. Homel. Secur.*, vol. 2, no. May, pp. 230–235, 2011.

[32] M. Gupta and R. Sharman, "Social Network Theoretic Framework for Organizational Social Engineering Susceptibility Index," *Proc. 2006 Am. Conf. Inf. Syst.*, pp. 3383–3393, 2006.

[33] A. Vishwanath, "Habitual Facebook Use and its Impact on Getting Deceived on Social Media," *J. Comput. Commun.*, vol. 20, no. 1, pp. 83–98, 2015.

[34] T. Moore and R. Clayton, "Evaluating the wisdom of crowds in assessing phishing Websites," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 5143 LNCS, pp. 16–30, 2008.

[35] A. Vishwanath, "Examining the Distinct Antecedents of E-Mail Habits and its Influence on the Outcomes of a Phishing Attack," *J. Comput. Commun.*, vol. 20, no. 5, pp. 570–584, 2015.

[36] I. Alseadoon, T. Chan, E. Foo, and J. G. Nieto, "Who is more susceptible to phishing emails ?: A Saudi Arabian study," *23rd Australas. Conf. Inf. Syst.*, no. Trusteer 2009, pp. 1–11, 2012.

[37] M. Pattinson, C. Jerram, K. Parsons, A. McCormac, and M. Butavicius, "Why do some people manage phishing e-mails better than others?," *Inf. Manag. Comput. Secur.*, vol. 20, no. 1, pp. 18–28, Mar. 2012.

[38] I. Alseadoon, M. F. I. Othman, and T. Chan, "What Is the Influence of Users' Characteristics on Their Ability to Detect Phishing Emails?," in *Advanced Computer and Communication Engineering Technology*, vol. 315, H. A. Sulaiman, M. A. Othman, M. F. I. Othman, Y. A. Rahim, and N. C. Pee, Eds. Cham: Springer International Publishing, 2015, pp. 949–962.

[39] K. Ding, N. Pantie, Y. Lu, S. Manna, and M. I. Husain, "Towards Building a Word Similarity Dictionary for Personality Bias Classification of Phishing Email Contents," pp. 252–259, 2015.

[40] S. Chaudhary, E. Berki, L. Li, and J. Valtanen, "Time Up for Phishing with Effective Anti-Phishing Research Strategies," *Int. J. Hum. Cap. Inf. Technol. Prof.*, vol. 6, no. 2, pp. 49–64, Apr. 2015.

[41] W. D. Kearney and H. A. Kruger, "Considering the influence of human trust in practical social engineering exercises," *Inf. Secur. South Africa (ISSA), 2014*, pp. 1–6, 2014.

[42] R. T. Wright and K. Marett, "The Influence of Experiential and Dispositional Factors in Phishing: An Empirical Investigation of the Deceived," *J. Manag. Inf. Syst.*, vol. 27, no. 1, pp. 273–303, 2010.

[43] X. Luo, R. Brody, A. Seazzu, and S. Burd, "Social Engineering: The Neglected Human Factor for Information Security Management," *Inf. Resour. Manag. J.*, vol. 24, no. 3, pp. 1–8, 2011.

[44] I. Alseadoon, M. F. I. Othman, E. Foo, and T. Chan, "Typology of phishing email victims

based on their behavioural response," in *19th Americas Conference on Information Systems, AMCIS 2013 - Hyperconnected World: Anything, Anywhere, Anytime*, 2013, pp. 1–9.

[45] T. Vidas, E. Owusu, S. Wang, C. Zeng, L. F. Cranor, and N. Christin, "QRishing: The susceptibility of smartphone users to QR code phishing attacks," in *Lecture Notes in Computer Science*, 2013, vol. 7862 LNCS, pp. 52–69.

[46] J. Wang, T. Herath, R. Chen, A. Vishwanath, H. R. Rao, and R. H. Raghav, "Phishing Susceptibility: An Investigation Into the Processing of a Targeted Spear Phishing Email," *IEEE Trans. Prof. Commun.*, vol. 55, no. 4, pp. 345–362, Dec. 2012.

[47] A. Alnajim and M. Munro, "Effects of Technical Abilities and Phishing Knowledge on Phishing Websites Detection," in *Software Engineering*, 2009.

[48] J. S. Downs, M. B. Holbrook, and L. F. Cranor, "Decision strategies and susceptibility to phishing," in *SOUPS '06: Proceedings of the second symposium on Usable privacy and security*, 2006, vol. 15213, pp. 79–90.

[49] V. Anandpara, A. Dingman, M. Jakobsson, D. Liu, and H. Roinestad, "Phishing IQ Tests Measure Fear, Not Ability," in *Financial Cryptography and Data Security*, 2007, pp. 362–366.

[50] E. Metalidou, C. Marinagi, P. Trivellas, N. Eberhagen, G. Giannakopoulos, and C. Skourlas, "Human factor and information security in higher education," *J. Syst. Inf. Technol.*, vol. 16, no. 3, pp. 210–221, Aug. 2014.

[51] J. M. Drew and C. Cross, "Fraud and its PREY: Conceptualising social engineering tactics and its impact on financial literacy outcomes.," *J. Financ. Serv. Mark.*, vol. 18, no. 3, pp. 188–198, 2013.

[52] V. Garg, L. Huber, L. J. Camp, and K. Connelly, "Full paper: Risk communication design for older adults," *Gerontechnology*, vol. 11, no. 2, pp. 1–10, Jun. 2012.

[53] J. T. Cacioppo, R. E. Petty, and C. F. Kao, "The efficient assessment of need for cognition.," *Journal of personality assessment*, vol. 48, no. 3. pp. 306–307, 1984.

[54] O. Blazhenkova and M. Kozhevnikov, "The new object-spatial-verbal cognitive style model: Theory and measurement," *Appl. Cogn. Psychol.*, vol. 23, no. 5, pp. 638–663, Jul. 2009.

[55] L. Goldberg, "Language and individual differences: The search for universals in personality lexicons," *Review of Personality and Social Psychology*, vol. 2. pp. 141–165, 1981.

[56] L. Christensen, "Deception in psychological research: When is its use justified?," *Personal. Soc. Psychol. Bull.*, vol. 14, no. 4, pp. 664–675, Dec. 1988.

[57] J. H. Korn, *Illusions of reality: A history of deception in social psychology*. 1997.