A Closer Look at the Email Phishing Landscape

Hamza Biabani

University of Twente, The Netherlands

h.a.biabani@student.utwente.nl

ABSTRACT

While we may strive for the latest security measures and stringent protocols, the human element remains the weakest and most effective link of attack for malevolent individuals. Exploiting these human links for crucial information is what's referred to as social engineered attacks. To counter this, interventions are developed to protect uninformed or unaware users of such systems. In this paper we focus on the most common type of social engineering attack: email phishing. The goal is firstly, to review the efficacy of currently employed interventions, and secondly to identify the factors that can inform susceptibility. To do this, a meta-analysis and sub-group analysis were conducted on current literature. All studies from 2000 till present were screened for relevance and would be narrowed down further via eligibility criteria. In the end, 8 papers are used for the analysis, with 14 sub-category variables included. Current interventions have a medium effect size (SMD=0.527). The most effective interventions include elements of warning, a focus on URL identification, and use multiple formats of delivery.

Additional Key Words and Phrases: Effect size, intervention, meta-analysis, phishing, social engineering, susceptibility,

1 INTRODUCTION

The usage of deception and manipulation to gain access to sensitive information is called Social Engineering [4]. It is a method of attack that preys on human weakness and psychology. Humans are seen at the weakest link in cybersecurity systems [5]. Social engineers exploit psychological factors of human beings, such as our inherently trusting nature or our carelessness [6]. Socially engineered attacks come in all shapes and varieties. They can be done in close proximity, as a face-to-face deception, or, more commonly, through more distant means such as online malware or telephone scams.

As the years advance, technology is advancing with it. The extensive digitization of human communication means that email usage has become a necessity to thrive in today's world [1]. However, with this advancement comes an increase in digital attacks as well [2]. In this paper, the focus will be on a type of social engineering attack known as "e-mail phishing".

TScIT 37, July 8, 2022, Enschede, The Netherlands

While there are mechanisms such as email filtering systems in place to prevent email attacks, some will eventually make it to peoples' inboxes [18] [19], exposing the human element as a weakness that can be targeted. Many of these attacks make use of social engineering by including elements of deception such as clicking on malicious URLs or giving confidential information. In fact, 90% of malware is delivered via email means [20]. According to a report by the FBI, in 2020 [21], phishing attacks were the most common type of cybercrime carried out, occurring more than double the amount compared to other types of online crime for this year.

To combat email phishing, researchers and experts have devised many interventions to minimize the effects of such attacks. There are many such interventions that have been carried out in the past, each with differing experimental setups. For example, one uses interactive games [7], another uses cybersecurity training [8], and one simply warns users about the threat of phish [9]. Some papers saw a significant improvement in how people responded [7], while some had insignificant[10] or even negative effects[11].

The goal of this paper is to provide a bird's eye view of the current phishing intervention landscape. This will help understand not only how effective current interventions are, but also other factors that can help predict and inform attacks. To this end, the following research question is posed:

RQ: What is the current state of socially engineered phishing attacks?

To adequately answer the question in its entirety, the following sub-research questions have also been generated:

RQ1: What is the effectiveness of current phishing interventions?

RQ2: What factors predict and counter phishing?

2 METHODOLOGY

The methodology consists of 3 parts. First, scientific literature concerning the topic is sourced and classified. Then, a meta-analysis is performed to gauge the effectiveness of the

interventions of the studies (**RQ1**). Finally, a sub-group analysis is performed to gauge the effects of the factors on countering phishing (**RQ2**).

2.1 Effectiveness of current interventions

To begin with the selection, preliminary research on the types of articles associated with this topic were performed. After discussion with a supervisor, and working through several iterations to reduce as many results as possible, the following final query was entered into Scopus on 12.05.2021:

```
TITLE-ABS-KEY
( "social engineering" OR phishing OR whaling OR "spear phishing" ) AND TITLE-ABS-KEY ( intervention OR training OR countermeasure OR warning ) AND TITLE-ABS-KEY ( experiment OR survey OR analysis OR campaign OR *mail OR attack OR test OR simulation ) AND NOT TITLE-ABS-KEY ( "spam filter" OR "machine learning" OR "artificial intelligence" OR "neural network*" OR "natural language processing" OR algorithm )
```

As can be seen from the query above, the goal was to return articles related to social engineering subjects, some sort of involved training or intervention, as well as an experimental design for the study. There are also many exclusion terms, which are there to filter out results featuring non-human approaches to phishing, such as automated spam filters or machine learning. The query returned 656 results.

The abstracts and titles of the papers were screened for further eligibility. During this process, 557 articles were removed, and the remaining 99 papers would be checked against eligibility criteria.

The final screening looked at the following criteria:

1. The paper must include an intervention to reduce susceptibility to social engineering attacks
2. The paper must be written in English
3. The paper must cover email phishing specifically. Other social engineering attacks are not considered
4. Participants must be exposed to a new experience of phish. Papers that focus on participants' experiences with phishing, for example, are excluded
5. There must be a comparison between two groups or more, in order to gauge the effect of the intervention
6. The papers are published after the year 2000, to further remove non-email phish

8 papers made the final cut, with 91 being rejected. Common reasons for rejection include spotty or undetailed data, insufficient sample size of groups in the study (<20), and not having a proper experimental design/setup. The overview of this entire process can be found in Figure 1.

2.2 Meta-analysis

The dependent variable of the studies is the effect size of the intervention. To calculate this, we use the Standardised Mean Difference (SMD), specifically Cohen's d, along with the intervals of confidence for each effect. Some papers can have more than one intervention type used, which leads to multiple effect sizes. The SMD works well, as it allows us to easily compare effects between studies [12]. A SMD of 0.2 or less signifies a low effect size, with 0.5 signifying a medium effect, and anything above 0.8 is considered a high effect size.

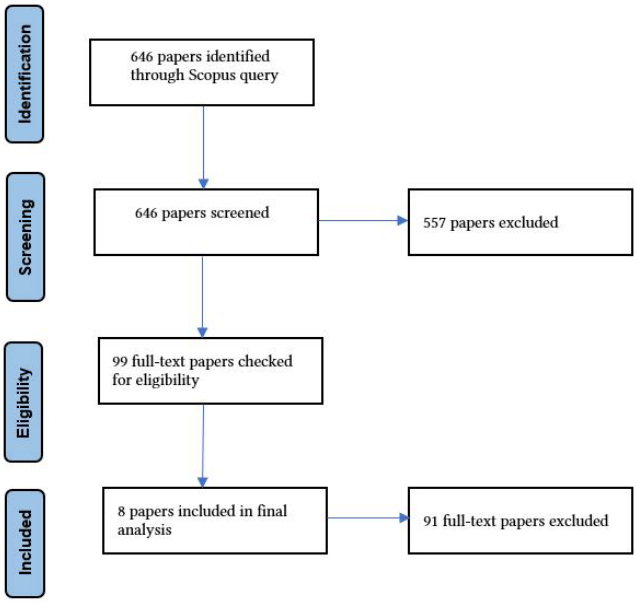


Figure 1: Data collection

After the SMD for each intervention is calculated, it is time to perform the meta-analysis. To do so, Comprehensive Meta Analysis (CMA) [3] is the software of choice. We are comparing the SMD scores for all interventions. The SMD is the difference between the two means divided by the pooled standard deviation of both groups. A random effects model is used as the studies used are not identical in design [12]. This will help answer RQ1.

2.3 Sub-group analysis

For the subgroup analysis, we need to first define the list of independent variables whose effects we are interested in. To do so, 16 variables are defined, split across 3 categories; Study Context, Attack Characteristics, and Intervention Characteristics. The categories and variables are inspired by those used by Bullée and Junger [13], who have performed a similar meta-analysis on social engineering interventions.

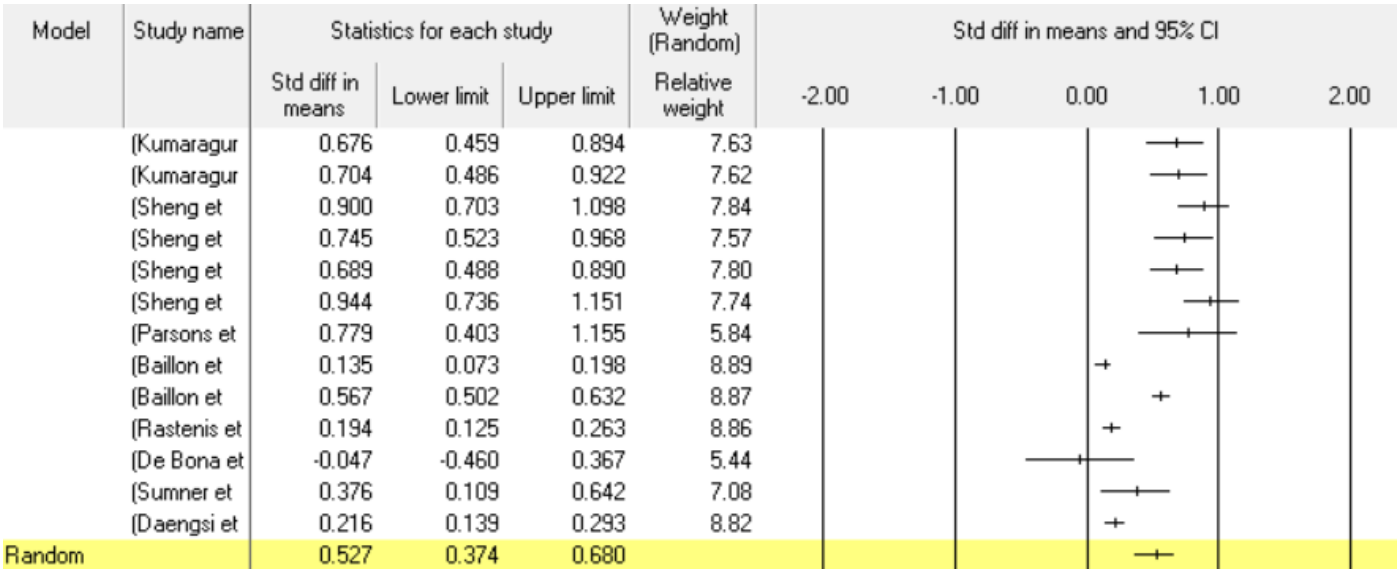
The Study Context category looks at the broad effects of how the study is designed. The following variables are used:

1) Randomisation: How the groups in the experiment are formed.
    a) Yes, participants are randomized

b) Quasi, not truly randomized
c) No, there are no randomization measures used at all

2) Control: Whether a control group is used in the experiment.
   a) Yes, a control group is used
   b) No, a control group is not used

3) Environment: Whether the experiment takes place in a controlled, laboratorial setting or a more naturalistic one.
   a) Real, naturalistic conditions for the experiment
   b) Lab, participants tested in laboratory conditions



Model	Study name	Std diff in means	Lower limit	Upper limit	Weight (Random) Relative weight
	(Kumaragur	0.676	0.459	0.894	7.63
	(Kumaragur	0.704	0.486	0.922	7.62
	(Sheng et	0.900	0.703	1.098	7.84
	(Sheng et	0.745	0.523	0.968	7.57
	(Sheng et	0.689	0.488	0.890	7.80
	(Sheng et	0.944	0.736	1.151	7.74
	(Parsons et	0.779	0.403	1.155	5.84
	(Baillon et	0.135	0.073	0.198	8.89
	(Baillon et	0.567	0.502	0.632	8.87
	(Rastenis et	0.194	0.125	0.263	8.86
	(De Bona et	-0.047	-0.460	0.367	5.44
	(Sumner et	0.376	0.109	0.642	7.08
	(Daengsi et	0.216	0.139	0.293	8.82
Random		0.527	0.374	0.680	

Figure 2: Meta-analysis results

4) Awareness: Degree of realization of participants that they are taking part in an experiment.
   a) Full, the participants know not only that they are part of an experiment, but also the purpose of said experiment
   b) Partial, participants know that they are part of an experiment, but do not know its true purpose
   c) None, participants are not aware at all

The Attack Characteristics Category looks at variables related to the type of exposure participants in the study have to phishing. The following variables are used:

5) Method: The way in which participants experience the attack.
   a) Sorting, participants are asked to roleplay and sort an inbox of emails
   b) Normal, participants receive the email attacks in a way that is indistinguishable from normal emails

6) Type of phishing: The subtype of email phishing used in the experiment.
   a) Spear phish, the emails are targeted to specific individual groups and/or organisations
   b) Normal, a normal phishing email, which utilizes more broader messaging than spear phishes

7) Frequency: How often the victims are targeted.
   a) Single, the victims receive the attack once
   b) Multiple the victims receive the attack more than once

8) Pre-victimisation: Whether participants are made a victim before receiving the intervention.
   a) Yes, only victimized participants receive the intervention
   b) No, all participants receive the intervention

The Intervention Characteristics category looks at variables related to the methods of treatment used to reduce participants' vulnerability. The following variables are used:

9) Modality: The method of delivery for the intervention.
   a) Static, such as documents and infographics
   b) Oral, such as lectures and video
   c) Dynamic, such as games and interactive learning

10) Format: The format of the provided intervention.
   a) Text, the intervention consists of textual information only
   b) Comic, a comic strip is used to present the information
   c) Game, an interactive game or learning platform
   d) Comic + Game, a combination of comic as well as game was used
   e) Other, an unspecified combination of previous formats was used, such as Comic + Text for example

11) Focus: The aspect of victimization that the intervention is focused on.

a) URL, how to recognize malicious URL links
b) Email, how to recognize phishing emails
c) Both, how to recognize both malicious URLs as well as phishing emails

12) Intensity: How intense the measure is.
a) Low, such as a warning message or leaflet
b) High, such as a lecture or interactive game/quiz

13) Warning: Whether the participants are warned about phishing.
a) No, there was no warning provided
b) Yes, participants were warned about being targeted by phish
c) Yes + Train, participants were not only warned, but also provided training

14) Priming: Whether participants exposed to certain stimuli that can change their behaviour to phish.
a) Yes, participants are primed
b) No, participants are not primed

The SMD is again calculated for all variables. These values are used for the sub-group analysis, in order to calculate the overall strength of each type. Again, this was performed using CMA. This will provide a meaningful look at the categories' individual effects, helping answer RQ2.

3 RESULTS

The results of the meta-analysis can be found in Figure 2. In it, we see a forest plot with all the individual studies. We can see the individual as well as pooled effect size and limits for the 95% confidence interval. Each study's relative weight towards the pooled effect size is also displayed. Table 1 shows the results of the subgroup analysis.

3.1 Effectiveness of current phishing interventions

From Figure 2, we can see that the overall SMD for all the interventions is 0.527, suggesting a medium effect, with a 95% confidence interval of [0.374–0.680]. The effect is positive, signifying that the interventions do reduce susceptibility.

3.2 Factors that predict and counter phishing

3.2.1 Randomisation

Interventions that employed randomization had moderate effect sizes (SMD=0.748). When quasi randomness (SMD=0.342) or no randomness at all (SMD=0.210) were used, there was only a small effect size. The effect of randomness was statistically significant (p=.00).

3.2.2 Control Group

Experiments with a control group (SMD=0.436) had double the effect of those with no control (SMD=0.210), although both had a small effect. The effect was statistically significant (p=.00).

3.2.3 Environment

Experiments that took place in a lab setting had a moderate effect (SMD=0.765), while the effect of the real world environment was small (SMD=0.298). The effect was statistically significant (p=.00).

3.2.4 Awareness

All effects of awareness had a small effect size (SMD= 0.194, 0.439, 0.207) for the conditions of being none, partially, or fully aware respectively. Interestingly, being partially aware had the largest effect (SMD=0.439). The effect was statistically significant (p=.00).

3.2.5 Method

Having participants sort emails had a moderate effect (SMD=0.765), while more traditional email attacks had a small effect (SMD=0.298). The effect was statistically significant (p=.00).

3.2.6 Type of phish

The type of phish had a big impact. Spear phishing had a small effect (SMD=0.299), while traditional phishing had a large effect (SMD=0.819). The effect was statistically significant (p=.00).

3.2.7 Frequency of attack

The frequency of attack had little difference, as being attacked once yields a small effect (SMD=0.374), while multiple attacks had a slightly smaller effect (SMD=0.323). The effect was statistically significant (p=.00).

3.2.8 Pre-victimisation

Being pre-victimised actually had a small, but negative effect (SMD=-0.047), meaning it increased susceptibility. Not being pre-victimised also had a small effect, albeit positive (SMD=0.352). The effect was statistically significant (p=.00).

3.2.9 Modality

The different modes of delivery had similar, smaller effect sizes, with oral delivery having the highest effect (SMD=0.391), followed
by dynamic content (SMD=0.348), with static content (SMD=0.301) having the least effect. The effect was statistically significant (p=.00).

3.2.10 Format

The combination of comic+game had the most effect (SMD=0.944), which is also a large effect. Comics on their own (SMD=0.690), as well as formats that were made up of other combinations (SMD=0.567). had a moderate effect. The game (SMD=0.279) and text (0.205) formats had small effects. The effect was statistically significant (p=.00).

3.2.11 Focus

There are no studies that focused solely on email identification. The ones with the URL focus had a medium effect (SMD=0.743),

while those that focused on both email as well as URL had a small effect (SMD=0.283). The effect was statistically significant (p=.00).

3.2.12 Intensity

The effects of intensity were small, and similar. Low intensity had a SMD of 0.370, while those with a higher intensity had a SMD of 0.327. The effect was statistically significant (p=.00).

3.2.13 Warning

All the studies had at least a warning of some kind. Warnings by themselves had a small, negative effect (SMD=-0.047), while warnings along with training had a small, positive effect (SMD=0.352). The effect was statistically significant (p=.00).

3.2.14 Priming

When priming was done, there was a small effect (SMD=0.335), while a lack of priming yielded a moderate effect (SMD=0.690). The effect was statistically significant (p=.00).

Characteristic	Types	SMD	95% CI	n	I^2	p
STUDY CONTEXT						0.000
Randomisation	Yes	0.748	[0.667 , 0.830]	8	66.41	
	Quasi	0.342	[0.297 , 0.387]	2	98.87	
	No	0.210	[0.160 , 0.260]	3	0.00	
Control Group						0.000
	Yes	0.436	[0.397 , 0.475]	10	95.05	
	No	0.210	[0.160 , 0.260]	3	0.00	
Environment						0.000
	Lab	0.765	[0.672 , 0.858]	6	62.75	
	Real	0.298	[0.265 , 0.331]	7	95.47	
Awareness						0.000
	Full	0.207	[0.132 , 0.283]	2	33.31	
	Partial	0.439	[0.400 , 0.478]	10	94.91	
	None	0.194	[0.125 , 0.263]	1	0.00	
ATTACK CHARACTERISTICS						
Method						0.000
	Normal	0.298	[0.265 , 0.331]	7	95.47	
	Sorting	0.765	[0.672 , 0.858]	6	62.75	
Type of Phish						0.000
	Normal	0.819	[0.720 , 0.919]	5	2.56	
	Spear	0.299	[0.267 , 0.332]	8	94.73	
Frequency						0.000
	Multiple	0.323	[0.278 , 0.368]	8	94.07	
	Single	0.374	[0.332 , 0.417]	5	96.39	
Pre-victimisation						0.000
	Yes	-0.047	[-0.460 , 0.367]	1	0.00	
	No	0.352	[0.321 , 0.383]	12	95.17	
INTERVENTION CHARACTERISTICS						
Modality						0.000
	Dynamic	0.348	[0.281 , 0.414]	4	94.58514	
	Oral	0.391	[0.343 , 0.438]	2	98.32093	
	Static	0.301	[0.249 , 0.354]	7	94.56446	
Format						0.000
	Comic + Game	0.944	[0.736 , 1.151]	1	0.00	
	Comic	0.690	[0.567 , 0.812]	3	0.00	
	Game	0.279	[0.209 , 0.349]	3	89.98	
	Text	0.205	[0.161 , 0.250]	5	93.64	
	Other	0.567	[0.502 , 0.632]	1	0.00	
Focus						0.000
	URL + Email	0.283	[0.250 , 0.317]	6	95.58	
	URL	0.743	[0.661 , 0.825]	7	57.40	
Intensity						0.000
	High	0.327	[0.282 , 0.373]	7	94.78	
	Low	0.370	[0.327 , 0.413]	6	95.64	
Warning						0.000
	Warning + Training	0.352	[0.321 , 0.383]	12	95.17	
	Warning	-0.047	[-0.460 , 0.367]	1	0.00	
Priming						0.000
	No	0.690	[0.536 , 0.844]	2	0.00	
	Yes	0.335	[0.304 , 0.367]	11	95.28	

Table 1: Sub-group analysis

4 DISCUSSION

The results from the study indicate that that interventions used to counter phishing have a medium effect size. When it comes to the effect of variables of categories relating to the overall study, we see that most of them also have a positive effect, and that some have a larger effect than others.

4.1 Meta-analysis

As discussed in the previous section, the interventions had a medium effect, with SMD of 0.527. What can be understood from this is that interventions have a positive effect in reducing victimization, but this can always be improved as the effect is medium and not large. This answers RQ1.

4.2 Sub-group analysis

Firstly, let's take a look at the study context variables. The studies which employed randomization had a medium to large effect on participants' effectiveness. Comparing this to the small effects of the studies which had no randomization, and we can see that randomization changes the study outlook quite drastically. These findings are in line with those of Bullée and Junger [13]. Similar trends can be seen with the control variable. The interventions which used a control had double the effect size compared to those that did not. This is perhaps explained to the fact that using randomization and control groups helps isolate the effects of the intervention better, and so the effectiveness of the interventions is more apparent.

Lab studies are also much more effective than real life studies, with the former having a medium-large effect compared to the small effect of the latter. This is likely due to participants being a lot more aware and alert in the lab environment. In fact, the awareness variable points to this, as we can observe a higher effect size for studies in which participants are at least partially or fully aware compared to those in which they are not. These effects are in check with those of [13].

For the attack characteristics, it can be seen that the experiments in which participants had to sort emails yielded a medium-high effect, compared to the low effect of traditional email attacks. This is likely a byproduct of priming, as participants who are told to sort emails are likely more alert and vigilant for fraudulent activity.

Interventions that were tested by normal phishing yielded a much higher, large effect size than spear phishing. This makes sense, as spear phishing tends to be a lot more targeted and specialized towards members of organisations, so experiments using this method of attacking would lead to more victimization. This is also supported by previous findings in which spear phishing was the more effective attack [15]. There was little difference in effect size for the frequency of attacks. Those who were pre-victimised performed worse. This result is in line with [13], but contradicts [16], which suggested that being attacked, while also being made aware of it, helps with learning. For the pre-victimised criterion, the effect of the intervention is only observed for those who fell victim prior. Hence, it possible that these populations, by virtue of being selected for falling victim when others did not, were less likely to respond to effects of interventions than their non victimised peers, which could explain why they performed worse.

Finally, there are the variables for the intervention itself. Oral content is slightly more effective than dynamic and static content, but not by a large degree. The combination of comic and game is the most effective, evident by large effect size. This is perhaps due to the fact that more than one format is used here and is hence much more effective. It could also explain why the "other" category also did very well, as this covers alternate combinations that are not common in other studies. Interventions which solely had the URL as the focus performed much better. The chosen intensity of the intervention did not have a big difference on effect. Interventions that just warned participants had a slightly negative effect, while those that provided training had an overall small but positive effect. What this could indicate is that training is a necessary component to have an effective intervention, and a sole warning may not be sufficient. A new observation from the research is that not priming participants was more effective. This is different from previous research such as [14] and [13], which suggested that primed subjects performed better than their non primed counterparts.

The previously discussed findings can help make clear what factors should be focused on more to more effectively counter phishing. While it is nice to see some more straightforward results, such as having combinations of modalities be more effective, there are also several unexpected ones, such as priming interventions performing less effectively compared to their counterparts.

4.3 Practical Implications

Based on the findings of the research, we can suggest more secure interventional measures for practitioners in the field. What we would suggest is using a multimedia approach, such as a combination of textual information, such as training manuals, and interactive learning formats, such as game-based training. The focus should be on recognizing one topic at a time, and the training should include warning elements, while also trying to avoid the effects of priming.

4.4 Future Research

For future research in this field, we would suggest looking more closely at priming and its effects. In this study, our findings suggest that priming is detrimental, but cannot completely explain why this is the case. We also suggest studying multiple combinations of formats, instead of solely individual formats, and see if they corroborate with the findings of this paper.

4.5 Limitations of study

This study has a few limitations:

- First, the research in this study was performed solely by one person, so the results are not double checked by any peer researchers.

- Second, there are a limited number of studies used in the paper. For some of the categories specified, there is only one study, for example. This can drastically skew results.
- Third, due to publication bias, it is possible that studies on this topic that have negative or insignificant effects are not published. Therefore, the real effectiveness of interventions may be lower than proposed in this paper.
- Finally, most of the participants of the studies are from WEIRD (Western, Educated, Industrialised, Rich, and Democratic) countries which make up a minority of the world's population [17]. Therefore, the results of the study cannot be applied to countries that are not WEIRD.

REFERENCES

[1] OECD 2019.*How's Life in the Digital Age?* . DOI:https://doi.org/10.1787/9789264311800-en

[2] Gupta, B., Arachchilage, N. and Psannis, K., 2017. Defending against phishing attacks: taxonomy of methods, current issues and future directions. *Telecommunication Systems*, 67(2), pp.247-267

[3] Why use Comprehensive Meta-Analysis Software: 2021. https://www.metaanalysis.com/pages/why_use.php?cart=BXMJ5572595. Accessed: 2022- 06- 20.

[4] Christopher. Hadnagy. 2011. Social engineering : the art of human hacking. (2011), 382.

[5] Taimur Bakhshi. 2018. Social engineering: Revisiting end-user awareness and susceptibility to classic attack vectors. Proceedings - 2017 13th International Conference on Emerging Technologies, ICET2017 2018-January, (February 2018), 1–6. DOI:https://doi.org/10.1109/ICET.2017.8281653

[6] Francois Mouton, Louise Leenen, Mercia M. Malan, and H. S. Venter. 2014. Towards an Ontological Model Defining the Social Engineering Domain. IFIP Advances in Information and Communication Technology 431, (2014), 266–279. DOI:https://doi.org/10.1007/978-3-662-44208-1_22

[7] Sheng, S. et al. 2010. Who falls for phish? A demographic analysis of phishing susceptibility and effectiveness of interventions. Conference on Human Factors in Computing Systems - Proceedings. 1, (2010), 373–382. DOI: https://doi.org/10.1145/1753326.1753383

[8] Il kwon Lim, Young Gil Park, and Jae Kwang Lee. 2016. Design of Security Training System for Individual Users. Wireless Personal Communications 90, 3 (October 2016), 1105–1120. DOI:https://doi.org/10.1007/S11277-016-3380-Z

[9] Jing Chen, Scott Mishler, Bin Hu, Ninghui Li, and Robert W. Proctor. 2018. The description-experience gap in the effect of warning reliability on user trust and performance in a phishing-detection context. International Journal of Human-Computer Studies 119, (November 2018), 35–47. DOI:https://doi.org/10.1016/J.IJHCS.2018.05.010

[10] Marco de Bona and Federica Paci. A Real World Study on Employees' Susceptibility to Phishing Attacks. Proceedings of the 15th International Conference on Availability, Reliability and Security. DOI:https://doi.org/10.1145/3407023

[11] Alex Sumner, Xiaohong Yuan, Mohd Anwar, and Maranda McBride. 2021. Examining Factors Impacting the Effectiveness of Anti-Phishing Trainings. https://doi.org/10.1080/08874417.2021.1955638 (2021). DOI:https://doi.org/10.1080/08874417.2021.1955638

[12] Michael Borenstein, Larry v. Hedges, Julian P.T. Higgins, and Hannah R. Rothstein. 2010. A basic introduction to fixed-effect and random-effects models for meta-analysis. Res Synth Methods 1, 2 (April 2010), 97–111. DOI:https://doi.org/10.1002/JRSM.12

[13] Jan Willem Bullee and Marianne Junger. 2020. How effective are social engineering interventions? A meta-analysis. Information and Computer Security 28, 5 (November 2020), 801–830. DOI:https://doi.org/10.1108/ICS-07-2019-0078

[14] Kathryn Parsons, Agata McCormac, Malcolm Pattinson, Marcus Butavicius, and Cate Jerram. 2015. The design of phishing studies: Challenges for researchers. Computers and Security 52, (July 2015), 194–206. DOI:https://doi.org/10.1016/J.COSE.2015.02.008

[15] Bartlomiej Hanus, Yu Andy Wu, and James Parrish. 2022. Phish Me, Phish Me Not. Journal of Computer Information Systems 62, 3 (2022), 516–526. DOI:https://doi.org/10.1080/08874417.2020.1858730

[16] Richard A. Schmidt and Robert A. Bjork. 1992. New Conceptualizations of Practice: Common Principles in Three Paradigms Suggest New Concepts for Training. Psychological Science 3, 4 (1992), 207–217. DOI:https://doi.org/10.1111/J.1467-9280.1992.TB00029.X

[17] Joseph Henrich, Steven J Heine, and Ara Norenzayan. The weirdest people in the world? DOI:https://doi.org/10.1017/S0140525X0999152X

[18] Christopher N. Gutierrez, Taegyu Kim, Raffaele della Corte, Jeffrey Avery, Dan Goldwasser, Marcello Cinque, and Saurabh Bagchi. 2018. Learning from the ones that got away: Detecting new forms of phishing attacks. IEEE Transactions on Dependable and Secure Computing 15, 6 (November 2018), 988–1001. DOI:https://doi.org/10.1109/TDSC.2018.2864993

[19] Rakesh Verma, Narasimha Shashidhar, and Nabil Hossain. 2012. Detecting phishing emails the natural language way. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 7459 LNCS, (2012), 824–841. DOI:https://doi.org/10.1007/978-3-642-33167-1_47/COVER/

[20] Purplesec 2022 Cyber Security Statistics: The Ultimate List Of Stats, Data & Trends Retrieved July 3, 2022 from https://purplesec.us/resources/cyber-security-statistics/

[21] Internet Crime Report 2020. *FBI Internet Crime Complaint Centre.*