Effectiveness of Social Engineering Intervention: A meta-analysis

Kristjan Haavel
University of Twente
PO Box 217, 7500 AE Enschede
The Netherlands
k.haavel@student.utwente.nl

ABSTRACT

Due to the ongoing digitisation of different processes, cybersecurity is an integral part of life. The weakest link in cybersecurity is people who can be manipulated through social engineering attacks. There are many different preventative measures against such attacks, and with this paper, we find out how effective the most recent measures are. To find out the effectiveness, a systematic literature review was conducted with data starting from 2018, and then a meta-analysis of the studies was conducted. The analysis consists of 8 studies with 45 effect sizes and a total of 28,277 subjects. We found that design-based interventions are the most effective type of preventative measure and that not all the conclusions are in line with the data found on the similar research into the topic with data pre-2018.

Keywords

Social Engineering, Phishing, Training and Awareness Programs, Meta-analysis & Systematic Review

1. INTRODUCTION

The COVID19 pandemic was, in modern times, an unprecedented event that caused society to become more technology-driven. Thanks to increased digitisation, people are now spending more time online, and there has been an increase in the rate of cyber-attacks. Criminals have become more creative in their attacks, and due to people losing their jobs and spending more time at home, people have likely turned to cyber-crime to support themselves [6]. With the global pandemic, social engineering attacks have become more prevalent [12], so having up-to-date data and awareness of the best preventative measures will be helpful for companies to know to keep their cybersecurity protocols in order.

Humans are believed to be the weakest link in cyber-security [9]. This link can be exploited through social engineering, a form of technical assault that relies heavily on human interaction and entails manipulating people into disobeying standard security protocols [4]. Modern-day social engineers create ways to exploit people's greed, ambition, or vanity with the help of technology the same way a con artist exploits people in person.

Bullee et al. [4] found in their research that the different ways that social engineers attack their victims are:

- Voice call
- Email
- Face-to-Face
- Text Messages / Social Media

With voice calls and face-to-face social engineering, the offender usually contacts their victim and introduces themselves as someone essential to gain their targets' trust. They then gather sensitive information which should not be shared with the attacker [17]. With text-based social engineering, the main form of attack is phishing, where the attacker disguises their text to appear as if it comes from a legitimate source. However, it usually contains a link to a malicious file that can take control of the victim's data without them even being aware of it [7].

Several high-profile web services have been compromised over the past decade through social engineering based attacks, resulting in millions of leaked passwords. These kinds of leaks violate people's privacy and give attackers access to different platforms that victims use throughout the internet [1].

With these attacks, a lot of personal data can be collected and exploited without the victims being aware of it, so preventative interventions have to be set so such attacks would not happen. There are many different intervention mechanisms against social engineering, and this paper will find out how effective these measures are and what specific elements affect the effectiveness the most.

The effectiveness of social engineering interventions has been previously studied by Jan-Willem Bullee and Marianne Junger [5], but their study used data up to the end of 2017. Their work will be further expanded with current research with data starting from 2018 up until now. With the rise of cybercrime, it can be expected that preventative measures have also been improved upon since the last research on this subject.

Hussain Aldawood and Geoffrey Skinner also carried out a thorough literature review in 2019 concerning the pitfalls and ongoing issues of social engineering training and awareness programs [1]. Furthermore, Workman found with his work in 2008 that even trained people fail to recognise such attacks [24]. Nevertheless, these works are not a systematic review and are more focused on the shortcomings of current interventions and not specifically looking into the preventative methods' effectiveness.

With this research, we found out that interventions are effective in preventing text-based social engineering attacks. This result was found by doing a systematic literature review and a random-effects meta-analysis. A random-effects model is used because the true effect size varies from study to study [3]. With this research, we will find answers to the following research questions:

RQ: How effective are recent interventions in preventing text-based social engineering attacks?

To support this research question following sub-questions are also answered:

RQ1.1: What training methods are used to educate people on such attacks?

RQ1.2: What elements are the most successful in preventing social engineering attacks against people?

The following ten subgroups were defined and analysed to understand the specific elements that contribute to the intervention effectiveness:

1. *Type of intervention* explains which of the four different kinds of intervention a specific observation belongs to.
2. *Pre-victimisation* is used to define if the subjects were victimised before they were presented with the intervention.
3. *Modality* explains how effective the different modes of how subjects were trained are.
4. *Priming* shows if the subjects were implicitly warned before participating in the study.
5. *Warnings* is used to understand if the participants were explicitly warned about the dangers.
6. *Focus* shows if the intervention was focused on either email attacks or webpage attacks.
7. *Intensity* explains how intense was the preventative measure.
8. *Retention* is used to understand how well the subjects retained their knowledge about such attacks.
9. *Awareness* shows how aware people were about participating in a study.
10. *The environment* is used to understand if the intervention was based on a homely environment or lab conditions.

2. METHODOLOGY AND APPROACH

2.1 Data Collection

First, we conducted a systematic literature review of social engineering preventative methods. To perform this review, we followed a research article on the guidance of doing a systematic review [25]. This review led to a better understanding of the current interventions, what elements contribute to them, how effective they are, and the improvements that have been made since this topic was last researched. Literature for the systematic review was identified through the Scopus database by querying results starting from 2018, due to the latest research on this having data up until the end of 2017.

The Scopus database was queried on 02.05.2021 to obtain studies for the systematic analysis. The database query was inspired by the previous research into this topic [5]. However, additional criteria, highlighted in bold, were added to fit the eligibility criteria better, and the query was as follows:

TITLE-ABS-KEY ("social engineering" OR "phishing" AND "online" OR "cybercrime" OR "internet" OR **"prevention"** AND "experiment" OR "training" OR "survey" OR "warning" OR "intervention") **AND PUBYEAR > 2017** AND (EXCLUDE (SUBJAREA, "MEDI")) **AND (LIMIT-TO (PUBSTAGE, "final")) AND (LIMIT-TO (DOCTYPE , "cp") OR LIMIT-TO (DOCTYPE , "ar") OR LIMIT-TO (DOCTYPE, "ch")) AND (LIMIT-TO (LANGUAGE, "English"))**

The reasoning behind the added parts is that "prevention" was added as we are looking at preventative methods. PUBYEAR was added so we can look at studies published after 2017. Limit to final was added, so it is a published paper. Doctype limits were added to fit the eligibility criteria, and language limit was added to understand the content of papers.

This query identified 226 results. With a thorough literature review, these results were screened through to reduce the amount to only relevant studies which fit the following eligibility criteria inspired from previous research on this topic [5]:

1. To be published as a scientific paper or a PhD thesis.
2. The manuscript must be written in English.
3. The study involves human subjects and no technical solutions against attacks as this study focuses on human behaviour in the context of social engineering.
4. The studies are published after 2017
5. The experiment and intervention should aim to reduce victimisation by social engineering. Deception or malicious attack must be involved.
6. There should be a comparison of at least two groups and 20 observations per group to back the legibility of data. The comparison of groups is required to state the effectiveness of new interventions.
7. An experimental design should be used. Questionnaires or surveys that only measure attitude or intention are excluded.

The search query automatically covered by criteria 2 and 4. Criteria 1 was partly covered by results having scientific papers, and afterwards accepted papers were manually confirmed if they were indeed published scientific papers or a PhD thesis. However, for checking the search results, criteria 3, 5, 6, and 7 were looked for by scanning the papers. Out of the 226 papers, we found 14 papers that looked like they should fit the criteria and additional 15 papers that could fit the criteria. These 29 papers were read through to see if they can be used for the meta-analysis.

Besides the papers that could and might fit the criteria, we also found additionally 58 papers that were separated into different colour codes. Separation was conducted as good for content writing, suitable for better topic understanding, and possible papers of interest to look through that do not seem to fit the criteria directly but might still be helpful for the meta-analysis.

The main reasons for excluding most of the papers were:

- No preventative methods but for example, were about general awareness-raising or victim action analysation
- Not enough participants
- Had no control group / no two groups
- Bad data
- Technical preventative methods
    - Machine learning
    - Neural networks
    - Algorithmic predictions
    - Automatic blacklisting

At the end of the collected data, we found only eight papers with k= 45 observations that fit the meta-analysis [2, 8, 10, 13, 16, 17, 21, 26]. In addition, for one of the studies, the authors were contacted to gather more exact information about their research population sizes [16].

2.2 Data / Statistical Analysis

Once the data was collected, a random-effects meta-analysis and subgroup analysis was conducted. The final data consisted of 8 studies with k= 45 effect sizes and a total of n= 28,277 subjects.

3.2.1 Meta-analysis

To conduct the meta-analysis between the studies, we used a program called comprehensive meta-analysis [23]. Standardised Mean Difference (SMD) is used as the studies use different scales to assess their outcome [3]. SMD of 0-0.2 is seen as a small effect, 0.2-0.8 as a moderate effect and anything bigger than 0.8 as a large effect [14]. Both SMD and lower and upper limits, also

known as bounds of 95% confidence interval, were used to measure the effectiveness of the intervention. The analysis was conducted using a random-effects model as the results are drawn from different effect sizes of different studies. The program data was inserted either by events or event rates and sample sizes of intervention and control group. The meta-analysis consisted of 8 different studies with several effect sizes

Table 1: Frequency distribution of study effect sizes

Effect size	Frequency
1	2
2	2
3	1
6	2
24	1

2.2.2 Subgroups

We defined ten subgroups of interest to study within the preventative methods. Additionally to the results displayed in the meta-analysis, heterogeneity (I^2) and p-Value are also brought out. P-Value is used for the test of the null, and any value that falls under 0.05 is considered statistically significant. Heterogeneity is shown to see the amount of variance on a relative scale. Values on the order of 25%, 50%, and 75% are considered low, moderate and high, respectively [3].

2.2.2.1 Type of preventative measures

We differentiated four types of preventive measures against social engineering to see which one is the most effective. In the paper by Jansen et al. [11], it was written that the different types of precautionary measures could be categorised as security education, training, awareness-raising, and design. Training involves developing skills in information security, while security education is about spreading the knowledge and awareness of mitigating threats and understanding the threats online [11]. Awareness-raising is about warning users and focusing attention on specific threats and countermeasures [19]. Finally, design is changes in the environment to give protection to the user pushing them in the right direction [11].

2.2.2.2 Pre-victimisation

It has been found previously by Schmidt et al. [20] that people are more effective at learning right after being attacked. Suppose someone is taught without them feeling victimised. In that case, they might feel as if the subject is boring and not that important, so with this subgroup, we want to see if pre-victimisation caused any differences in the intervention methods.

2.2.2.3 Modality of intervention

The analysed studies used different ways of presenting intervention, so with this sub-group, it can be found out how much the different ways vary between each other. For example, some studies had just reading material while others had games or spoken lectures.

2.2.2.4 Subject priming

Parsons et al. [15] found that participants are better at distinguishing phishing attacks from non-phishing attacks when they are primed. So we are interested in finding out if priming the subject had any significant effect on the intervention effectiveness.

2.2.2.5 Warnings

Bullee et al. [5] found in their meta-analysis and literature search that warnings had no considerable effect on people. However, we are interested in finding out with this research if the reaction and awareness to warnings have increased with time.

2.2.2.6 Focus of the intervention

Another interesting point found with the literature review was that some of the interventions were just focused on email phishing detection. However, some were also focused on just the URL/website and some on both. So we are interested if different focuses had a significant effect on the knowledge understandability of subjects.

2.2.2.7 Intensity of intervention

The intensity of the intervention type varied between the researched papers. For example, the paper by Xiong et al. [26] just had warnings that people had to read, so that would be considered as low-intensity interventions. On the other hand, other papers had either different lectures or games about internet crime to be considered medium or high intensity. Also, Bullee et al. found in their paper [5] that higher intensity interventions had a much more significant effect than medium and low intensity, so it would be interesting to find out if this is still the case.

2.2.2.8 Retention of knowledge

Several studies in the meta-analysis separately tested retention of knowledge. We assume that with time knowledge disappears from memory. So with this subgroup, we can find out if preventative measures should be applied periodically. Also, Bullee et al. [5] found that time between intervention and victimisation is negatively associated, so we will also find out if similar results can be obtained.

2.2.2.9 Awareness of study

Awareness of participating in a study seems to matter for the outcome, as found in the meta-analysis by Bullee et al. [5]. The papers included currently had several different awareness factors. Therefore, with this subcategory, we can find out if the awareness resulted in any bias with the final result.

2.2.2.10 Environment

People act differently in their home environment from when being tested in lab conditions [12]. So with this sub-criteria, we will determine if the difference between environments caused any biased result between studies.

3. RESULTS

With the literature review, it was found that there are several different training methods against text-based social engineering attacks. They can be categorised as:

- Training games
- Warnings
- Lectures
- Fake phishing with debriefing

The results of the meta-analysis can be found in Figure 1. The table shows the forest plot of both individual and pooled effect sizes, the lower and upper limit, standardised mean difference (SMD), the sample sizes of the intervention and control groups, and each effect size's weights. The values of the subgroup analysis can be seen in Table 2. From there, we can see each characteristic SMD, 95% confidence interval, the number of observations, the heterogeneity, and the p-Value separately.

3.1 Individual studies analysis

The SMD of all the studies was 0.382, which shows a moderate effect of interventions, and the lower and upper limits are [0.301,

0.463]. From this, we can conclude that an average new intervention increases the effectiveness of not falling for a social engineering attack by 0.382 of the standard deviation of the specific outcome measure.

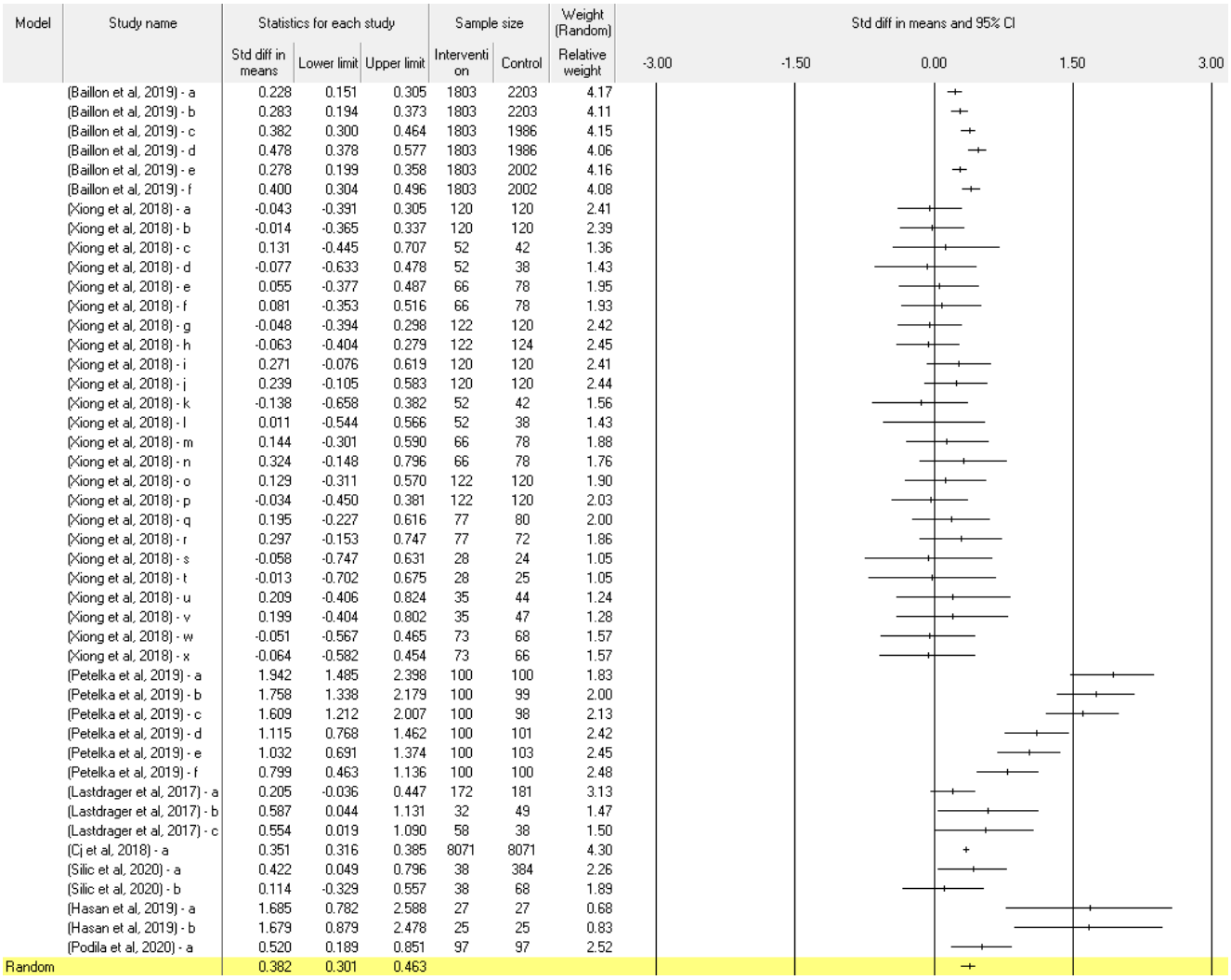

Figure 1: Meta-analysis results

3.2 Subgroup analysis

Table 2: Subgroup analysis

Characteristic	SMD	Lower limit	Upper limit	p-Value	I^2 (%)	n
Type						
Awareness	0.425	0.332	0.519		52.971	2
Design	1.317	1.014	1.621		69.628	7
Security ed.	0.254	0.198	0.309		0.000	6
Training	0.220	0.128	0. 312		69.087	30
Overall	0.301	0.343	0.259	0.000		45
Pre-Victimisation						
No	0.471	0.345	0.597		87.489	24
Yes	0.295	0.183	0.407		73.802	21
Overall	0.372	0.289	0.456	0.041		45
Priming						
No	0.731	0.584	0.878		92.423	17
Yes	0.148	0.058	0.239		36.643	28
Overall	0.308	0.231	0.385	0.000		45

(continued)

Characteristic	SMD	Lower limit	Upper limit	p-Value	I^2 (%)	n
Warning						
No	0.292	0.007	0.591		8.190	2
Warning + train	0.261	0.207	0.315		50.946	32
Warning only	1.150	0.770	1.530		88.986	11
Overall	0.279	0.226	0.332	0.000		45
Focus						
Both	0.343	0.294	0.393		57.971	12
Email	0.114	0.329	0.557		0.000	1
Site/URL	0.405	0.191	0.620		86.557	32
Overall	0.344	0.296	0.392	0.509		45
Retention						
Delayed	0.358	0.232	0.484		78.826	17
Instant	0.400	0.284	0.517		85.362	28
Overall	0.381	0.295	0.466	0.626		45
Awareness						
Fully	0.351	0.316	0.385		0.000	1
None	0.256	0.177	0.334		69.813	14
Partly	0.515	0.307	0.723		85.029	30
Overall	0.339	0.308	0.370	0.024		45
Intensity						
High	0.352	0.318	0.385		0.000	6
Low	0.404	0.251	0.558		85.823	34
Medium	0.290	0.224	0.355		51.499	5
Overall	0.341	0.312	0.370	0.179		45
Environment						
Lab	0.389	0.187	0.590		17.621	4
Real life	0.378	0.292	0.464		84.425	41
Overall	0.380	0.301	0.459	0.923		45
Modality						
Dynamic	0.353	0.319	0.387		0.000	3
Spoken	0.346	0.090	0.602		20.694	3
Static	0.380	0.276	0.483		85.185	39
Overall	0.355	0.323	0.388	0.887		45

3.2.1 Type of preventative measures

Design-based interventions were associated with an immense effect size (SMD= 1.317), followed by the moderate effect of awareness-raising interventions (SMD= 0.425), and lower moderate effect sizes were in training and security-based interventions (SMD= 0.254 and SMD= 0.220). The type of preventative measure was statistically significant (p= 0.000).

3.2.2 Pre-victimisation

The effects of victimisation or not were both associated with a moderate effect size. However, not victimising had a bit of a higher effect size (SMD= 0.471) than victimising (SMD= 0.295). The differences were statistically significant (p= 0.041).

3.2.3 Modality of intervention

All the different intervention modes also shared a similar effect size. Dynamic interventions had the highest effect size (SMD= 0.353), followed by static methods (SMD= 0.380), and the smallest moderate effect size was on spoken modality (SMD= 0.346). There was no statistically significant effect on the modality of intervention used (p= 0.887).

3.2.4 Subject priming

Interventions that did not use priming had a higher moderate effect size (SMD= 0.731), and the interventions that did prime their victims had a low effect size (SMD= 0.148). These differences were statistically significant (p= 0.000).

3.2.5 Warnings

Interventions that had just warnings had the most significant effect size (SMD= 1.150). Both warning+training and no warning had a similar lower moderate effect size (SMD= 0.261 and 0.292, respectively). There was a statistically significant effect on the use of warnings in preventing attacks (p= 0.000).

3.2.6 Focus of the intervention

Interventions focused on Site/URL and emails were with moderate effect size (SMD= 0.405 and SMD= 0.343), and interventions that focused on just emails had a small effect size (SMD= 0.114). There was no statistically significant effect (p= 0.509).

3.2.7 Intensity of intervention

Effect sizes of intervention intensities were all moderate. Low-intensity interventions had the highest effect size (SMD= 0.404), followed by high intensity (SMD= 0.352), and the smallest effect size was medium intensity (SMD= 0.290). There was not a statistically significant effect on the intensity (p= 0.179).

3.2.8 Retention of knowledge

Retention of knowledge had similar moderate effect sizes, with instant (SMD= 0.400) having more effect than delayed retention (SMD= 0.358). However, there was not a statistically significant effect of retention (p= 0.626).

3.2.9 Awareness of study

Not being aware of participating in a study had a lower moderate effect size (SMD= 0.256), and being aware either fully or partially had a moderate effect size (SMD= 0.351 and SMD= 0.515). Awareness was statistically significant (p= 0.024).

3.2.10 Environment

Studies conducted in a lab were associated with a medium effect size in effectiveness (SMD= 0.389), and similarly, studies not

conducted in a lab also had a moderate effect size (SMD= 0.378). The effects were not statistically significant (p= 0.923).

4. DISCUSSION

4.1 Literature review

With the literature review, it was found that there is a big focus on computer-based preventions against social engineering since 2018. The interest in human-based preventions does not seem to be as active as it was, as most papers were not interested in the human aspect of such attacks. The papers that did research about the human side of this had four different training methods. They were training games, which educate people in a fun, playful way. Then warnings with guidelines on how to recognise dangerous links, sites or emails. Also, lectures to raise awareness about such technical attacks and to reduce the rate of falling for them. Finally, fake phishing was done with a debriefing about the attack to teach the victims what they should have done instead and what to look out for in the future.

4.2 Analysis

The main research question can be answered with the meta-analysis. This meta-analysis shows that recent social engineering preventative measures reduce victimisation in a statistically moderate effect, with a standardised mean difference of 0.382. This finding has a smaller effect size than was found by Bullee et al. [5] with their previous research into this topic. The difference can possibly be explained by people becoming more aware of the dangers of the internet with time, and thanks to that, they are not so easily victimised anymore.

The second research question can be answered with the subgroup analysis where not that many significant differences can be concluded from the data. The most effective *type of intervention* is design-based, which helps guide users to act correctly. It was found to be much more effective than dynamic, security education or training-based preventative methods, so design-based interventions would be recommended to make the most effective intervention.

Pre-victimisation did not significantly affect the final result of the studies, which does not align with the previously stated fact that people pay more attention after being attacked. The finding that pre-victimised victims have less of an effect than no previous victimisation seems to align with previous research [5].

With regards to *priming* subjects, a new observation was found with this study. Not priming subjects worked out more effectively than priming subjects about phishing attacks. This finding is probably because Xiong et al. [26] had quite an adequate control group compared to the intervention group and a large effect size where most of the primed study results were derived, so their effectiveness is not seen as influential.

The findings of *warning* seem to be opposite to the previous study [5]. Just warning people is a lot more effective than adding training to the warning or not warning them at all. The results from this might show that people are now more aware of the dangers of the internet so that they might turn more attention to warnings. The reason why warning + training seems to be less effective than just warning might also be explained like in the previous paragraph, with most of the data coming from Xiong et al. [26].

Focus on the interventions being on the URL/site designs were most effective, and the least effective was focus on just email. This result is also in line with the previous findings of Bullee et al. [5]. So we can possibly conclude that looking at the URL/site design is more straightforward than looking at specific elements in emails.

Regarding *retention of knowledge,* similar findings can be seen, with the effects of intervention being less effective with time. So we can conclude that awareness against such attacks needs to be spread constantly as people do not remember all they have been taught previously.

Awareness of participating in a study had a higher effect size than no awareness, so the analysed studies results might be biased due to letting people know they are participating in a study. This finding correlates with the findings of Bullee et al. [5].

A new finding regarding the *intensity of intervention* is that low-intensity studies had a higher effect than medium and high-intensity studies. This result might be because people could start overthinking their teachings with higher intensity interventions, which made them make less correct choices. However, the difference is not big enough to make any specific conclusion from it.

The *environment* where the study was conducted had no significant differences in effectiveness. This finding is not in line with previous findings [5], which the small sample of laboratory studies in the current analysation might cause. Another point that could be drawn from this is that interventions now are as effective in real-life environments as they are in laboratory conditions.

Modality also had no significant differences between the effectiveness of different modes of preventative measures. This also contradicts previous findings, which might also be explained by the small example sizes of spoken and dynamic interventions. Another explanation could be that each different method of teaching is as effective as any other.

In conclusion, not all the findings were similar to those found in previous research into this topic, but this might be because people are more aware of the dangers of the internet, so even though the interventions are still effective, they are not as effective as they used to be. With further research into this topic, we recommend that at least two people gather the data to make no mistakes with data gathering. Going through such an amount of papers alone is a lot of work and time-consuming. Due to limited time constraints, the literature review results could not be coded twice to ensure that all the papers that need to be included were definitely included correctly. With no previous statistical analysis experience, it was not easy to plan everything before starting the whole process. For future work, we recommend that better scheduling and understanding of the time required for this needs to be taken into consideration.

4.3 Practical implications and recommendations

Practitioners can use the results from this study to plan an improvement to their current intervention or think of a new way to protect against social engineering attacks. We would recommend for a new "perfect" intervention to be design-based, with no pre-victimisation or priming. It should contain warnings and be conducted recurringly so the knowledge does not disappear with time.

5. REFERENCES

[1] Aldawood, H. and Skinner, G. 2019. Reviewing Cyber Security Social Engineering Training and Awareness Programs—Pitfalls and Ongoing Issues. Future Internet. 11, 3 (2019), 73. Available at: http://dx.doi.org/10.3390/fi11030073

[2] Baillon, A. et al. 2019. Informing, simulating experience, or both: A field experiment on phishing risks. PLOS ONE. 14, 12 (2019), e0224216. Available at: http://dx.doi.org/10.1371/journal.pone.0224216

[3] Borenstein, M. et al. 2009. Introduction to Meta-Analysis. (2009). Available at: https://doi.org/10.1002/9780470743386.ch10

[4] Bullée, J. and Junger, M. 2019. Social Engineering. The Palgrave Handbook of International Cybercrime and Cyberdeviance. (2019), 1-28 Available at: https://doi.org/10.1007/978-3-319-78440-3_38

[5] Bullee, J. and Junger, M. 2020. How effective are social engineering interventions? A meta-analysis. Information & Computer Security. 28, 5 (2020), 801-830 Available at: http://dx.doi.org/10.1108/ICS-07-2019-0078

[6] Bullee, J-W., Montoya, L., Junger, M., & Hartel, P. H. 2016. Telephone-based social engineering attacks: An experiment testing the success and time decay of an intervention. In A. Mathur, & A. Roychoudhury (Eds.), Proceedings of the inaugural Singapore Cyber Security R&D Conference (SG-CRC 2016) (pp. 107-114). (Cryptology and Information Security Series: Vol. 14) IOS Press. Available at: https://doi.org/10.3233/9781-61499-6170107

[7] Camerer, C. et al. 2016. Evaluating replicability of laboratory experiments in economics. Science. 351, 6280 (2016), 14331436. Available at: http://dx.doi.org/10.1126/science.aaf0918

[8] CJ, G. et al. 2018. PHISHY - A Serious Game to Train Enterprise Users on Phishing Awareness. Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts. (2018). Available at: http://dx.doi.org/10.1145/3270316.3273042

[9] Happ, C., Melzer, A. and Steffgen, G. 2016. Trick with treat – Reciprocity increases the willingness to communicate personal data. Computers in Human Behavior. 61, (2016), 372-377. Available at: http://dx.doi.org/10.1016/j.chb.2016.03.026

[10] Hasan, Z. et al. 2021. A Multifactor Authentication Model to Mitigate the Phishing Attack of E-Service Systems from Bangladesh Perspective. Emerging Research in Computing, Information, Communication and Applications. Advances in Intelligent Systems and Computing. (2021), 75-86. Available at: http://dx.doi.org/10.1007/978-981-13-5953-8_7

[11] Jansen, J. and van Schaik, P. 2019. The design and evaluation of a theory-based intervention to promote security behaviour against phishing. International Journal of Human-Computer Studies. 123, (2019), 40-55. Available at: http://dx.doi.org/10.1016/j.ijhcs.2018.10.004

[12] Lallie, H., Shepherd, L. and Nurse, J. et al. Cyber security in the age of COVID19: A timeline and analysis of cyber-crime and cyber-attacks during the pandemic. Computers & Security 105, (2021), 102248. Available at: http://dx.doi.org/10.1016/j.cose.2021.102248

[13] Lastdrager, E. et al. 2017. How Effective is Anti-Phishing Training for Children?. Proceedings of the Thirteenth Symposium on Usable Privacy and Security. (2017), 229-239

[14] Newsletter 145: NCCMT Weekly Round-up - Introducing NEW video: Making Sense of a Standardised Mean Difference: 2015. https://www.nccmt.ca/e-newsletters/145. Accessed: 2021- 06- 08.

[15] Parsons, K. et al. 2015. The design of phishing studies: Challenges for researchers. Computers & Security. 52, (2015), 194-206. Available at: http://dx.doi.org/10.1016/j.cose.2015.02.008

[16] Petelka, J. et al. 2019. Put Your Warning Where Your Link Is. Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. (2019). Available at: http://dx.doi.org/10.1145/3290605.3300748

[17] Podila, L. et al. 2021. Practice-Oriented Smartphone Security Exercises for Developing Cybersecurity Mindset in High School Students. IEEE TALE2020 – An International Conference on Engineering, Technology and Education. (2021), 303-310. Available at: http://dx.doi.org/10.1109/TALE48869.2020.9368440

[18] Ramzan, Z. 2010. Phishing Attacks and Countermeasures. Handbook of Information and Communication Security. (2010), 433-448. Available at: http://dx.doi.org/10.1007/978-3-642-04117-4_23

[19] Sasse, M. et al. 2001. BT Technology Journal. 19, 3 (2001), 122-131. Available at: https://doi.org/10.1023/A:1011902718709

[20] Schmidt, R. and Bjork, R. 1992. New Conceptualisations of Practice: Common Principles in Three Paradigms Suggest New Concepts for Training. Psychological Science. 3, 4 (1992), 207-218. Available at: http://dx.doi.org/10.1111/j.1467-9280.1992.tb00029.x

[21] Silic, M. and Lowry, P. 2020. Using Design-Science Based Gamification to Improve Organizational Security Training and Compliance. Journal of Management Information Systems. 37, 1 (2020), 129-161. Available at: http://dx.doi.org/10.1080/07421222.2019.1705512

[22] van Schaik, P. et al. 2017. Risk perceptions of cyber-security and precautionary behaviour. Computers in Human Behavior. 75, (2017), 547-559. Available at: http://dx.doi.org/10.1016/j.chb.2017.05.038

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

[24] Workman, M. 2008. A test of interventions for security threats from social engineering. Information Management & Computer Security. 16, 5 (2008), 463-483. Available at: http://dx.doi.org/10.1108/09685220810920549

[25] Xiao, Y. and Watson, M. 2017. Guidance on Conducting a Systematic Literature Review. Journal of Planning Education and Research. 39, 1 (2017), 93-112. Available at: http://dx.doi.org/10.1177/0739456X17723971

[26] Xiong, A. et al. 2018. Embedding Training WithinWarnings Improves Skills of Identifying Phishing Webpages. Human Factors: The Journal of the Human Factors and Ergonomics Society. 61, 4 (2018), 577-595 Available at: http://dx.doi.org/10.1177/0018720818810942