Developing and Evaluating a Five Minute Phishing Awareness Video

Melanie Volkamer[1], Karen Renaud[2], Benjamin Reinheimer[1], Philipp Rack[1], Marco Ghiglieri[1], Peter Mayer[1], Alexandra Kunz[1], and Nina Gerber[1]

[1] Karlsruhe Institute of Technology and Technische Universität Darmstadt (firstname.surname@kit.edu)

[2] Abertay University and University of South Africa (k.renaud@abertay.ac.uk)

Abstract. Confidence tricksters have always defrauded the unwary. The computer era has merely extended their range and made it possible for them to target anyone in the world who has an email address. Nowadays, they send phishing messages that are specially crafted to deceive. Improving user awareness has the potential to reduce their effectiveness. We have previously developed and empirically-validated phishing awareness programmes. Our programmes are specifically designed to neutralize common phish-related misconceptions and teach people how to detect phishes. Many companies and individuals are already using our programmes, but a persistent niggle has been the amount of time required to complete the awareness programme. This paper reports on how we responded by developing and evaluating a condensed phishing awareness video that delivered phishing awareness more efficiently. Having watched our video, participants in our evaluation were able to detect phishing messages significantly more reliably right after watching the video (compared to before watching the video). This ability was also demonstrated after a retention period of eight weeks after first watching the video.

Keywords: Phishing awareness, user study, retention study

1 Introduction

More than twenty years after its emergence, phishing still succeeds [1, 37, 43]. Phishing attacks are increasingly sophisticated. It used to be easy to spot phishing messages due to poor language use and incorrect spelling; nowadays phishers are far smarter, sending plausible-looking messages calculated to deceive. They have also migrated from exclusively using email to plying their trade on a range of messaging platforms including messages in social media and messaging apps. A very popular trick is to entice the target to follow a link that will install malware or visit a *doppelgänger* website. The latter will persuade victims to divulge sensitive information, such as their access credentials. Automated detection is a powerful defence, but far from 100% effective [3, 15]. To narrow the gap left by technical measures, we need to make recipients of online messages aware of how to detect phishing attempts.

Our research group has a long history of developing phishing awareness programmes (including apps, flyers, reading material, presentations for seminars) and have carried out several user studies verifying their effectiveness [5–7,26,28,33,39–41]. Our initial programmes required learners to spend between 20 to 45 minutes completing the awareness programmes. Evaluations showed that all programmes significantly increased phishing detection rates. However, companies are concerned about the amount of time employees have to commit to these programmes. In response, we developed a video, which made it possible for us to shorten the time people needed to commit, because we could benefit from the visualisation functionality videos offer. The video now lasts only 5 minutes. The video was developed iteratively, incorporating feedback from people with various backgrounds (such as lay users, video producers, psychologists and security experts). The final video was evaluated by 89 participants who detected phishing messages significantly more often after watching the video. Many were able to demonstrate a retained ability to do this eight weeks later. The video was improved even further based on the feedback provided during the evaluation and the result of the evaluation, i.e. for those attack types participants performed worst, the explanations in the video were improved. Thus, our contribution is twofold: (1) the developed video based on previous research on phishing awareness programmes, and (2) its evaluation both straight after watching the video and during a retention study eight weeks after watching the video. We published the video[3] under the Creative Commons licence CC BY-SA 4.0 to remove all barriers to its use.

2 Development Process

2.1 Identification of the Relevant Content

The content to be covered is the following. The video should make the watcher aware of commonly-used phisher *strategies*. For example, trustworthy-looking messages, with familiar design and language employing psychological tricks to entice victims to click on an embedded link. They should also be aware of the possible *consequences* of clicking on a link. For example, malware could be downloaded onto their device. The web page they visit could look authentic, but actually be owned by a phisher. If credentials are divulged on this *faux* site, they could be used to facilitate identity theft. The video also deliberately addresses common phish-related *misconceptions* identified in the literature [11, 12, 17, 18, 20]. These include the following: (1) Phishers only send emails. In fact, they also use other mediums such as short text and social networking messages; (2) Phishing always harvests online banking credentials. In fact, phishers can actually fake any arbitrary web site: credentials are what they want; (3) The displayed name of the sender can be trusted to reflect the actual sender. In fact, details are faked very easily. The displayed sender name cannot be relied upon to signal

[3] German Phishing Video: `https://www.youtube.com/watch?v=XeslAkZIuwY&t=9s`
English Phishing Video: `https://www.youtube.com/watch?v=F4y2wzYpIKw`

authenticity; (4) Only wealthy people are targeted by phishers. On the contrary, anyone can be targeted, independently of how well known they are, how wealthy or their status in an organisation. (5) Technical security mechanisms are able to catch and block all phish messages. Actually, sophisticated phishers design their messages in such a way that the technical measures do not catch them; (6) The 'S' in HTTPS is an infallible signal of integrity. In reality, many phishers use website certificates to ally fears [4]; and (7) Trustworthy phrases in a website URL are a signal of trustworthiness. In fact, these are merely tricks used by the unscrupulous to trap the unwary. Thus, the video should help people to distinguish between phishing and genuine messages. Similar to our more time-consuming awareness programmes, we focus the learner's attention on the difference between the URL's actual destination and the destination it seems to be. Only by examining the link can people reliably distinguish between phish and genuine messages. The following instructions, explanations, and hints were included in the video:

Instruction-1: Locate the Actual Destination of a Displayed Link: The first step in phish detection is to know how to identify the actual destination the link will send people to. It might be in a tooltip, a status bars or in a special dialogue. They also need to be aware of the nuances behind links. Sometimes the actual destination is concealed behind a button or image or text like 'click here'. The actual destination is often hidden unless the person knows how to look for it. In rare cases the actual URL is displayed in the clear. The displayed tooltip might be faked too, in order to lull people into a false sense of security.

Instruction-2: Identify the So-Called Who-Area of the URL: After people have identified the actual destination URL, they should know how to identify the domain, what we refer to as the *who-area*. In the video, we told people that this is the last two terms that are separated by a dot before the first stand-alone "/" of a URL [5].

We also tell them that phishers deceive people by embedding the genuine company name somewhere in the URL *rather than the who-area*. They could place it either before or after the who-area. They should not rely on the signal conveyed by the use of HTTPS. Examples of phishing URLs are provided:

`https://www.gmail.com.mail-nows.com/login`

`https://mail-nows.com/https://www.gmail.com/login`.

Instruction 3: Check Authenticity of the Who-Area: Having located the who-area, the final step is to verify its authenticity, basically by checking it character by character. They are made aware of the fact that phishers often (1) use trustworthy terms (e.g. 'secure-shop.com') in the who-area; (2) stealthily replace characters. For example, they might replace a 'd' with 'cI' or introduce typos such as '`mircosoft`'.

[4] To avoid confusion we used https in all our phishing examples

[5] The study was carried out in Germany which meant we focused on domains with two terms e.g. amazon.de and we did not consider other conventions followed by countries like the U.K. with three terms, e.g. amazon.co.uk.

2.2 Video Development

We developed a story and a text for the voice-over based on the content being communicated by the messages, together with someone professionally developing awareness videos. We used simple language and non-technical terms (e.g. use of the phrase "who-area" for domain). We labelled screenshots to direct their attention to the location of important information (e.g. the status bar). We asked people of different ages with varying backgrounds and levels of expertise with IT and security to provide feedback to help us to improve and refine the video. The professional video producer developed the video using our text and underlying story. The video was improved, based on feedback from a number of people who were representative of the anticipated participants.

3 Evaluation – Methodology

The evaluation focused on the video's effectiveness in order to reveal significant improvements, in terms of phish detection ability. The following hypotheses were formulated:

H1: Participants, having watched the video, correctly judge the legitimacy of messages more often i.e. identify more phishing messages, and identify legitimate messages more reliably.

H2: Eight weeks after watching the video, participants correctly judge the legitimacy of messages more often i.e. identify more phishing messages, and identify legitimate messages more reliably.

3.1 Study Design

We conducted an online between-subjects study in two sessions. Hypothesis 1 is evaluated with the data from the first session and hypothesis 2 is evaluated with the data from both sessions. The tasks in the first evaluation session were:

1. Judge screenshots of messages. Decide whether each is a phish or legitimate.
2. Watch the video.
3. Judge screenshots of messages. When participants were asked to judge messages, the question was: "*Is this a fraudulent message?*'. Possible answers were: '*Yes, it is a fraudulent message*'. '*No, it is not a fraudulent message*'.
4. Provide video feedback (free text answers).
5. Provide demographic information.
6. Grant permission for us to contact them to engage with a retention study. If they consented, we requested their email address and provided them with a random code to ensure an anonymous link between sessions.

During the second session (approximately eight weeks later), consenting participants were invited to participate in the retention study, which required them again to judge screenshots of phish and legitimate messages (i.e. purely step 3 from the first session). We used the SosciSurvey online platform. The study

was pre-tested and the feedback from the pre-test addressed and refinements effected. The changes were mainly related to the content of the messages participants were asked to judge. We decided to go for a quiz-like evaluation, with security being the participant's primary task. The alternative would have been a study design in which the participant's primary task is related to a cover story. This would theoretically not prime them to expect and detect phish messages. We had a number of reasons for choosing the former design. While one could argue that the second option would have more external validity, it would have been hard to maintain the deception in a lab study. As soon as participants watched the video they would have known what the study was about. We could have attempted a field study i.e. getting people to watch the video then sending phishing alike messages at some unpredictable time in the future and measure how many click on links or open an attachment. It is challenging to measure the participant's ability to identify phish at a distance though. If they do not click, the message might not have been delivered, or they might not click because they do not have an account with the "source". When they click, it might be because they know of the study setting and want to know what happens if they click. We would not be able to determine whether they actually inspected the URL or not, which is what the video trains them to do correctly. In particular, in a field study, we cannot control whether some receive the email on a smartphone with a more challenging setting than on a laptop. Thus, there are many uncontrollable factors that could confound the findings.

Therefore, we decided to go for a study design in which security is participants' primary task. Note that improved awareness is only the first step towards taking action. In other words, if a user is not able to detect a phish when the primary task is security, it is unlikely that he/she will detect the phish when security is a secondary task. Thus, it is worth using a study design in which security is participants' primary task. In essence, this gives us an upper boundary for video effectiveness: the best we can hope to achieve.

3.2 Material

Messages were designed in such a way that a judgement could only be made based on the actual URL. We had to acknowledge that participants could consider a message as phish because they did not know the sender or did not have an account with the web service provider. Therefore, we asked participants to imagine the following scenario. They were Max Müller, who has an account with all web services used in the study, and who has a colleague named Jonas Schmidt. Furthermore, they were told that it was important for them to decide whether a message was legitimate or not because the fraudulent messages would harm them and ignored genuine messages could lead to negative consequences (we wanted to avoid their simply classifying all messages as phishes, just to be safe). This scenario was displayed when screenshots of messages had to be judged.

We used 16 messages in each task (pre, post, retention). All messages contained plausible content. They were displayed during the evaluation in a randomized order. Some more information about the messages:

Table 1. Overview of presented phishing messages (SB=Status Bar; TT=Tooltip)

Web address	Type	Inst.	Sender of message
`https://162.179.34.56`/login	TT	1+2	Service (DHL)
`https://www.secure-documents-online.com/...`	SB	1+2	Person (Colleague)
`https://control-center.1uncl1.de/...`	TT	1+3	Service (1 und 1)
`https://www.volksbanknig.de/...`	TT	1+3	Service (Volksbank)
`https://www.google.com.best-photos.com/...`	SB	1+2	Service (Google)
`https://www.zehrukol.com/ebay.com/...`	TT	1+2	Person (Colleague)
`https://www.bahncard.bahm.de/...`	SB	1+3	Service (DB)
`https://www.cognstar.de/...`	SB	1+3	Service (Congstar)

– One half contained suspicious links, the other half legitimate links.[6]
– We derived messages from messages received from web services and private contacts. Messages from web service providers were in the original design with original text (only the URL was replaced for the "phishing" messages).
– For all screenshots, the mouse was positioned so that the actual URL was displayed, depending on the software in place either in the tooltip (with Outlook) or in the status bar (with Thunderbird or a web browser). The usage of both types was equally distributed both for phishing messages as well as for legitimate messages. It was technically not possible to only show the URL when participants actually position the mouse on the link on SoSciSurvey.
– We designed phishing messages where ***instruction-1*** was enough to judge as well as those where ***instruction-2*** or ***instruction-3*** was needed (see Table 1)

3.3 Recruiting and Ethics

An attempt was made to recruit participants from a wide range of ages. Recruitment also took place via online platforms, social networks, flyers and personal invitations. Participants were not compensated for participating but we encouraged participation by telling them they learn how to avoid falling victim.

The requirements for research involving the human being, defined by the ethics committee of our university[7], were satisfied. This includes the fact that all data was collected independently of the identity of the participants. The email addresses they provided to permit us to contact them for the retention study were stored in a different database in a different order (as compared to their answers in the two sessions). The entries from the first session were linked to the one from the retention by asking participants to provide a random looking but well-defined code — well defined because they were told how to generate it based of names and birthdays from particular relatives. Furthermore, no third party

[6] Due to the fact that we used a quiz-like evaluation, we could present half-half although, in a realistic setting, half of people's messages would not usually be phish.

[7] `https://www.intern.tu-darmstadt.de/gremien/ethikkommisson/zustndigkeit/zustndigkeit.en.jsp`

(besides SosciSurvey) has a copy of the data and no third party was involved in the evaluation of the data.

4 Evaluation – Results

There are two groups of people in our sample: Those participating only in the first session (89: 39F/50M $\bar{x}$=36.1 years) and those who participated in both sessions (22: 12F/10M $\bar{x}$=38.09 years). There were no statistically significant differences between the groups; neither for age nor for gender. The distribution of the degree of education is as follows: from the 89 participants in the first session 50 have a university or university of applied science degree and 21 have an A-level qualification. The corresponding numbers for those 22 who participated in both sessions are: ten and five respectively. For the descriptive statistic see Table 2.

Table 2. Overview of detection rates in % and their standard deviation (SD) for all participants (all), those participating only in session 1 (G1) and those participating in both sessions (G2)

	Pre	**Post**	**Retention**
Phish G1	65.5 (SD 28.6)	83.8 (SD 20.5)	
Phish G2	42.6 (SD 29.3)	86.9 (SD 18.3)	81.3 (SD 16.3)
Phish all	59.8 (SD 30.3)	84.55 (SD 20.0)	81.3 (SD 16.3)
Legitimate G1	75.8 (SD 21.2)	87.7 (SD 17.3)	
Legitimate G2	75.0 (SD 21.1)	88.1 (SD 17.9)	83.0 (SD 20.3)
Legitimate all	75.6 (SD 21.1)	87.8 (SD 17.3)	83.0 (SD 20.3)

The performance change in detecting phishing and legitimate messages was measured in terms of correctly detected phish and legitimate messages. The difference in performance before and after watching the video H1 was analysed using a Repeated Measures ANOVA for both groups separately: (1) those participating only in session one and (2) those participating in both sessions. Furthermore we analysed the retention performance changes H2 using the Repeated Measures ANOVA considering only the answers from those participating in both sessions. The Mauchly Test indicates that there is a violation of Sphericity and therefore a Greenhouse-Geisser correction was needed for the comparison of pre- and post-performance. There was no violation of Sphericity to compare pre- and retention-performance.

4.1 Phishing Detection

Pre-Post for All Participants: We first report the Repeated Measures ANOVA with Greenhouse-Geisser correction for violated sphericity for the detection of phishing messages by all participants: The within-subject factor in time (pre and post performance) is significant with $p < .001$ and a η^2=.526, i.e.

the performance in detecting phishing messages changes significantly. In combination with the descriptive data (see Table 2), detection of phishing messages increases significantly after watching the video. Thus, H1 can be accepted.

Participants during Retention: A Repeated Measures ANOVA for fraudulent detection reveals a significant effect for the time (pre-, post- and retention-performance) with $p < .001$ and a η^2=.636. A post-hoc test with Bonferroni correction shows that there is a significant difference between pre- and post- with $p < .001$ and there is a significant difference for pre- and retention-performance with $p < .001$. Thus, H1 and H2 can be accepted for the group of participants taking part in both sessions.

4.2 Identifying Legitimate Messages

Pre-Post for All Participants: We first report the Repeated Measures ANOVA for identification rates. The within-subjects factor in time (pre- and post-performance) is significant with $p < .001$ and a η^2=.219, i.e. the performance in detecting legitimate messages changes significantly. In combination with the descriptive data (see Table 2), the identification of legitimate messages improves significantly after watching the video. Thus, H1 can be accepted.

Participants during Retention: A Repeated Measures ANOVA reveals a significant effect for the time (pre-, post- and retention-performance) with $p = .019$ and a η^2=.173. A post-hoc test with Bonferroni correction shows that there is a significant difference between pre- and post-performance with $p < .001$. Thus, H1 can be accepted.

4.3 Individual Messages

We also looked at the individual messages and their performance in order to improve the video. The corresponding mean values are depicted in Tables 3 and 4 respectively (note the number for pre and post are for all 89 participants).

Table 3. Detection rate, in %, for individual phishing URLs

Web address	Pre	Post	Ret.
`162.179.34.56/login`	55.1	93.3	81.8
`control-center.1uncl1.de/...`	63.6	93.3	90.9
`www.google.com.best-photos.com/...`	55.1	84.3	81.8
`www.zehrukol.com/ebay.com/...`	65.2	73.0	54.5
`www.secure-documents-online.com/...`	47.2	68.5	77.3
`www.bahncard.bahm.de/...`	66.3	91.0	90.9
`www.cognstar.de/...`	53.9	80.9	86.4
`www.volksbanknig.de/...`	73.0	92.1	86.4

Two particular phishing messages stand out where participants performed more poorly, as compared to the other phishing messages:

– Message using the legitimate URL (...docs.google.com/...) as HTML link in the text but "`https://www.secure-documents-online.com`" as the actual URL in the status bar. The message was identified correctly by 68.5% (lowest result) of participants after watching the video and 77.3% (second lowest result) in the retention session.
– Message using the actual URL "`https://www.zehrukol.com/ebay.com/software?id=12123213124`" in the text and in the tooltip: 73.0% (second lowest result) identified this message correctly after watching the video and 54.5% (lowest result) during the retention session.

Table 4. Detection rate in % for individual legitimate URLs

Web address	Pre	Post	Ret.
`marketresearch.apple.com/...`	61.8	84.3	81.8
`photos.google.com/...`	82.0	88.8	95.5
our university (anonymised)	94.4	97.8	100.0
`www.dropbox.com/...`	80.9	88.8	81.8
`www.gutefrage.net/...`	83.1	97.8	81.8
`buchung.lufthansa.com/...`	88.8	86.5	77.3
`www.vodafone.de/...`	51.7	71.9	59.1
`accout.wire.com/...`	62.5	86.5	86.4

Two legitimate messages stand out (the first one was particularly troublesome):

– "`https://www.vodafone.de/(...)`": 71.9% identified this correctly in the first session, with only 59.1% during the second session.
– "`https://buchung.lufthansa.com/servlet/cc?soDBYCTTDV YTEz0.26wa7uDU.261f7uuF.3df4D.2e.26EaEXEPNRTOOL_LINKEhttp:DVMD..`": The identification rate after watching the video was 86.5% with 77.3% correct identification during session 2.

4.4 Open Feedback

Positive comments mentioned the simplicity of the video, the clarity of the content and the general comprehensibility. In particular, they liked the fact that the video was not overloaded with information. Regarding the overall design, participants liked the idea of using this type of animated video for general knowledge transfer. Feedback for improving the video was: '*More examples of the different phishing tricks*' and '*Summary at the end of the video*'.

5 Discussion

The five-minute video significantly improved phish and legitimate message detection. In other words, after watching the video, participants were able to detect phishing URLs without becoming overly cautious.

The retention part of our study is of special interest since in real life people do not receive phishing messages on a daily basis due to improvements in technical measures that filter out these messages. It is unlikely that people will use their newly-acquired knowledge very often, so they are likely to forget instructions and hints from the video. Our participants improved significantly in terms of detecting phishing messages, whereas detection rates for legitimate messages stabilised. We suggest possible explanations for this observation:

- "`www.vodafone.de/(...)`": The mean detection rate, after watching the video, was 71.9% with 59.1% at retention. The message contained a telephone number and, instead of starting with 'Dear Martin ...', it started with 'Dear +1 121 34329'[8]. A paragraph in the email stated that Vodafone would always address their customers by their name. The issue, in this case, is that Vodafone does send emails to the phone number if the customer has not provided their name. We acknowledge that we did not spot this problem ourselves during the video refinement.
- "`https://buchung.lufthansa.com/s(...)/cc?soDBYCT(...)`": The mean detection rate, after watching the video, was 86.5% with 77.3% at retention. The problem here might be the length of the URL. The path contains HTTP twice, includes a number of dots and the term 'redirect'. This probably elicited suspicion. Again, the email was not altered from the original sent out by the company, besides changing the name of the customer.

The two phishing messages that evaded detection to the greatest extent were related to hints provided in ***instruction-2***. In the first case, participants did not detect the mismatch between the HTML-text-based URL displayed in the message and the actual destination URL displayed in the status bar. In the second case, participants are likely to have considered the path to be relevant in making their judgement. An improved video will have to explain these cases more clearly. The video did a great job with respect to ***instruction-3***. While these cases always performed worst in previous evaluations, they performed better after the video.

Most of the feedback regarding improving the video was related to extending it. This is interesting because we tried to keep it as short as possible while retaining efficacy. Two aspects might be worth considering in terms of improving the video: (1) make the fact that only the URL matters even more salient. (2) provide a summary at the end of the video to consolidate and reinforce concepts.

Note that, unlike studies reported by [29], we did not observe any age differences. This might be because security was their primary task. However, if the study had been carried out in the wild, our findings might well have coincided with those reported by [29].

Finally, it was interesting to observe that those participants who had many issues with detecting phishing messages in the pre-quiz were most likely to participate in the retention study. One possible explanation is that they really enjoyed

[8] The number we used in the message was randomly chosen, but realistic.

the video and were thankful for their improved awareness. One additional interpretation is that the video, in particular, addressed those who had very little pre-existing awareness of phishing.

5.1 Limitations

Almost half of the original 89 participants gave us their email addresses to contact them for the retention study. Half of these responded to our retention study request. We ended up with a sample of only 22 participants to participate in the retention session. This means that we cannot realistically generalise the results to the whole population. The participant sample, as a whole, is not representative, as most of our participants had an A-level certificate or university degree. Furthermore, due to the fact that we told participants, during recruiting, that they would learn how to protect themselves against online fraud if they took part, we probably attracted participants who were already interested in this topic. Thus, as future work, we should run the study with a different demographic.

Furthermore, participant performance should definitely be considered a "best-case" scenario, because security was their primary task. Their actual detection rates are likely to be poorer in the real world. However, an increased awareness of phishing detection is a necessary first step to resilience. Before watching the video, people were not able to detect phishes despite it being their primary task. This is why awareness programs are important.

We used the same messages in all sessions. It may be argued that one explanation for post-video improvement was that they already knew what to look for. This might be a valid observation, but the chances seem small because legitimate message detection rates actually decreased. Furthermore, it is unlikely that after eight weeks they would still remember all the messages they had seen before, including the correct judgement. It is also worth mentioning that participants were not given feedback about which messages they had judged correctly, and which not.

Due to technical limitations of SosciSurvey, participants did not need to hover over the link deliberately in order to display the actual destination URL. It was automatically provided on the screenshot. It could be argued that we don't know whether people would hover over links because our study did not require this essential first step. On the other hand, being aware of the need to hover must be helpful.

5.2 Video Improvements

Based on the results, we identified a number of issues with the video, which we addressed in order to maximize performance. The new video lasts only 5:09 minutes and spends more time on examples. Previously, the who-areas were highlighted in green and only sub-domains highlighted in red. Now this highlighting is extended to the path. Moreover, the video now explicitly tells people when a URL is a phishing URL. We now conclude the video with a summary of the lessons learned, including tips and hints.

6 Related Work

A number of user studies were conducted to gain insights into the mental models of message recipients, or to evaluate the effectiveness of anti-phishing measures. For example, a game-based anti-phishing educational approach in was used by [30], [2, 4, 13, 14, 16, 32, 34–36, 42, 44]. The effectiveness of some of these games has been evaluated in user studies [4, 14, 16, 30, 32, 34, 35, 44] with [16] and [32] comparing the effectiveness of a game-based approach with text-based awareness.

Of these, only the proposal in [32] is further evaluated in a retention study a week later [25]. Usually the participants in user studies are adults, but some researchers have started studying phishing education for children [27].

Another approach to anti-phishing education, utilises the so-called *teachable moments*: participants were sent a simulated phishing email with a suspicious-looking link. If they clicked on it, they were directed to a landing page containing phishing-related information. Such an approach, in particular, has been used by Caputo *et al.* [8]. The authors also conducted a retention study of anti-phishing training in a corporate setting after a period of 90 days. The results of the study, however, did not indicate any significant improvement. A similar approach has been used in further research [19, 23, 24], both of which conducted retention studies after one week, that did show significant improvements in terms of reducing participants' susceptibility to phishing emails. A further study [22] built upon the evaluation in [23, 24] and tested the participants' retention via multiple simulated phishing emails sent over the course of 28 days, with the results showing no significant loss of retention by the end of study. A similar approach but with spear phishing messages was studied in [9, 38]. The study in [10] further evaluated the effectiveness of an anti-phishing training based on three simulated phishing trials over a two month period, showing significant improvements even after the end of the study.

Other anti-phishing educational approaches include training materials, educational videos and e-learning modules. As such, a study in [45] evaluated an anti-phishing training coupled with motivational videos. Participants in the control group watched cooking videos instead. The study found that the training increased participants' ability to detect phishing emails, but also increased the rate of false positives. The authors did not test the retention effect of the training. A study reported by [31] compared the performance of several anti-phishing educational approaches, including the game-based approach from [32], training materials from [24] and popular anti-phishing materials found on the web. All of these approaches significantly improved participant ability to detect phishing links. Retention was not tested. An anti-phishing e-learning module was developed by [21] but was only evaluated with a small sample of participants, and retention was also not tested.

7 Conclusion

Modern technology allows confidence tricksters to target a large number of people using phishing messages, at minimal cost. It is still desirable for people to

know how to detect these messages as technology is far away from detecting 100%. In this paper we report on the development of a video to raise phishing awareness without deterring confirmation of the legitimacy of genuine messages. We used our knowledge and experience from past research to develop a short yet effective phish-awareness video. Our main aim was to cover the most relevant content in a short video to address the companies' needs to raise awareness, but not necessarily to have the luxury of spending between 20 and 45 minutes to do so. The video was evaluated by 89 participants. Furthermore, a retention study, in which 22 of these 89 participants took part, was conducted after eight weeks. The results of the study show that the ability of the participants to distinguish between phishing and legitimate links increased significantly directly after watching the video and that, even after eight weeks, the participants were significantly better at detecting phishing links than before watching the video.

Acknowledgements

This work was supported by the German Federal Ministry of Education and Research (BMBF) within the Competence Center for Applied Security Technology (KASTEL) and within the Center for Research in Security and Privacy (CRISP). Thanks to Alexander Lehmann for creating the video; for more of his security and privacy related videos see: https://www.youtube.com/user/alexanderlehmann

References

1. Anti-Phishing Working Group: Phishing Activity Trends Report, 4th Quater 2016. `https://docs.apwg.org/reports/apwg_trends_report_q4_2016.pdf` (2016), accessed 18 May 2017
2. Arachchilage, N.A.G., Cole, M.: Design a mobile game for home computer users to prevent from "phishing attacks". In: i-Society 2011: International Conference on Information Society. pp. 485–489. IEEE, London, UK (2011)
3. Asudeh, O., Wright, M.: Poster: Phishing website detection with a multiphase framework to find visual similarity. In: CCS 2016. pp. 1790–1792. ACM (2016)
4. Baslyman, M., Chiasson, S.: "Smells Phishy?": An educational game about online phishing scams. In: eCrime 2016: APWG Symposium on Electronic Crime Research. pp. 1–11. IEEE, Toronto, ON, Canada (2016)
5. Canova, G., Volkamer, M., Bergmann, C., Borza, R.: Nophish: An anti-phishing education app. In: Security and Trust Management (STM). pp. 188–192. Springer (2014)
6. Canova, G., Volkamer, M., Bergmann, C., Borza, R., Reinheimer, B., Stockhardt, S., Tenberg, R.: Learn to spot phishing urls with the android nophish app. In: IFIP World Conference on Information Security Education. pp. 87–100. Springer (2015)
7. Canova, G., Volkamer, M., Bergmann, C., Reinheimer, B.: Nophish app evaluation: lab and retention study. USEC. Internet Society (2015)
8. Caputo, D.D., Pfleeger, S.L., Freeman, J.D., Johnson, M.E.: Going spear phishing: Exploring embedded training and awareness. IEEE S&P 12(1), 28–38 (2014)

9. Caputo, D.D., Pfleeger, S.L., Freeman, J.D., Johnson, M.E.: Going spear phishing: Exploring embedded training and awareness. IEEE Security & Privacy 12(1), 28–38 (2014), https://doi.org/10.1109/MSP.2013.106
10. Dodge, R., Coronges, K., Rovira, E.: Empirical benefits of training to phishing susceptibility. In: IFIP SEC 2012. pp. 457–464. Springer (2012)
11. Dong, X., Clark, J.A., Jacob, J.: Modelling user-phishing interaction. In: Human System Interactions. pp. 627–632. IEEE (2008)
12. Downs, J.S., Holbrook, M.B., Cranor, L.F.: Decision strategies and susceptibility to phishing. In: SOUPS. pp. 79–90. ACM, Pittsburgh, Pennsylvania, USA (2006)
13. Hale, M., Gamble, R.: Toward increasing awareness of suspicious content through game play. In: SERVICES 2014. pp. 113–120. IEEE (2014)
14. Hale, M.L., Gamble, R.F., Gamble, P.: Cyberphishing: a game-based platform for phishing awareness testing. In: Hawai'i International Conference on System Sciences. pp. 5260–5269. IEEE, Kauai, USA (2015)
15. Han, X., Kheir, N., Balzarotti, D.: Phisheye: Live monitoring of sandboxed phishing kits. In: CCS 2016. pp. 1402–1413. ACM (2016)
16. Helser, S.: Fit: Identity theft education: Study of text-based versus game-based learning. In: ISTAS 2015. pp. 1–4. IEEE, Dublin, Ireland (2015)
17. Jagatic, T.N., Johnson, N.A., Jakobsson, M., Menczer, F.: Social phishing. Communications of the ACM 50(10), 94–100 (2007)
18. Jakobsson, M., Tsow, A., Shah, A., Blevis, E., Lim, Y.K.: What instills trust? A qualitative study of phishing. In: Financial Crypto. pp. 356–361. Springer (2007)
19. Jansson, K., von Solms, R.: Phishing for phishing awareness. Behaviour & Information Technology 32(6), 584–593 (2013)
20. Kauer, M., Pfeiffer, T., Volkamer, M., Theuerling, H., Bruder, R.: It is not about the design — it is about the content! Making warnings more efficient by communicating risks appropriately. In: Sicherheit. vol. 195. GI (2012)
21. Kawakami, M., Yasuda, H., Sasaki, R.: Development of an e-learning content-making system for information security (elsec) and its application to anti-phishing education. In: International Conference on E-Education, E-Business, E-Management and E-Learning. pp. 7–11. IEEE, Sanya, China (2010)
22. Kumaraguru, P., Cranshaw, J., Acquisti, A., Cranor, L., Hong, J., Blair, M.A., Pham, T.: School of Phish: a real-world evaluation of anti-phishing training. In: SOUPS. p. 3. ACM (2009)
23. Kumaraguru, P., Rhee, Y., Sheng, S., Hasan, S., Acquisti, A., Cranor, L.F., Hong, J.: Getting users to pay attention to anti-phishing education: evaluation of retention and transfer. In: APWG: eCrime. pp. 70–81. ACM (2007)
24. Kumaraguru, P., Sheng, S., Acquisti, A., Cranor, L.F., Hong, J.: Lessons from a real world evaluation of anti-phishing training. In: APWG: eCrime. pp. 1–12. IEEE (2008)
25. Kumaraguru, P., Sheng, S., Acquisti, A., Cranor, L.F., Hong, J.: Teaching Johnny not to fall for phish. TOIT 10(2), 7 (2010)
26. Kunz, A., Volkamer, M., Stockhardt, S., Palberg, S., Lottermann, T., Piegert, E.: Nophish: evaluation of a web application that teaches people being aware of phishing attacks. Informatik (2016)
27. Lastdrager, E., Gallardo, I.C., Hartel, P.H., Junger, M.: How effective is anti-phishing training for children? In: SOUPS. pp. 229–239 (2017)
28. Neumann, S., Reinheimer, B., Volkamer, M.: Don't Be Deceived: The Message Might Be Fake. In: International Conference on Trust and Privacy in Digital Business. pp. 199–214. Springer (2017)

29. Oliveira, D., Rocha, H., Yang, H., Ellis, D., Dommaraju, S., Muradoglu, M., Weir, D., Soliman, A., Lin, T., Ebner, N.: Dissecting Spear Phishing Emails for Older vs. Young Adults: On the Interplay of Weapons of Influence and Life Domains in Predicting Susceptibility to Phishing. In: CHI '17. pp. 6412–6424. ACM (2017)
30. Scott, M.J., Ghinea, G., Arachchilage, N.A.G.: Assessing the role of conceptual knowledge in an anti-phishing educational game. In: ICALT. pp. 218–218. IEEE (2014)
31. Sheng, S., Holbrook, M., Kumaraguru, P., Cranor, L.F., Downs, J.: Who falls for phish?: a demographic analysis of phishing susceptibility and effectiveness of interventions. In: CHI. pp. 373–382. ACM (2010)
32. Sheng, S., Magnien, B., Kumaraguru, P., Acquisti, A., Cranor, L.F., Hong, J., Nunge, E.: Anti-phishing Phil: the design and evaluation of a game that teaches people not to fall for phish. In: SOUPS. pp. 88–99. ACM (2007)
33. Stockhardt, S., Reinheimer, B., Volkamer, M., Mayer, P., Kunz, A., Rack, P., Lehmann, D.: Teaching phishing-security: Which way is best? In: IFIP International Information Security and Privacy Conference. pp. 135–149. Springer (2016)
34. Sun, J.C.Y., Kuo, C.Y., Hou, H.T., Yu-Yan, L.: Exploring learners' sequential behavioral patterns, flow experience, and learning performance in an anti-phishing educational game. Journal of Educational Technology & Society 20(1), 45 (2017)
35. Sun, J.C.Y., Yeh, K.P.C.: The effects of attention monitoring with EEG biofeedback on university students' attention and self-efficacy: The case of anti-phishing instructional materials. Computers & Education 106, 73–82 (2017)
36. Tseng, S.S., Chen, K.Y., Lee, T.J., Weng, J.F.: Automatic content generation for anti-phishing education game. In: ICECE. pp. 6390–6394. IEEE (2011)
37. Verizon: Verizons. `http://www.verizonenterprise.com/verizon-insights-lab/dbir/2017/` (2017), accessed 18 May 2017
38. Volkamer, M., Stockhardt, S., Bartsch, S., Kauer, M.: Adopting the cmu/apwg anti-phishing landing page idea for germany. In: 2013 Third Workshop on Socio-Technical Aspects in Security and Trust(STAST). pp. 46–52. IEEE (2013)
39. Volkamer, M., Renaud, K., Gerber, P.: Spot the phish by checking the pruned url 24, 372–385 (10 2016)
40. Volkamer, M., Renaud, K., Reinheimer, B.: Torpedo: tooltip-powered phishing email detection. In: IFIP International Information Security and Privacy Conference. pp. 161–175. Springer (2016)
41. Volkamer, M., Renaud, K., Reinheimer, B., Kunz, A.: User experiences of torpedo: Tooltip-powered phishing email detection. Computers Security 71, 100 – 113 (2017)
42. Wen, Z.A., Li, Y., Wade, R., Huang, J., Wang, A.: What. hack: Learn phishing email defence the fun way. In: CHI EA 2017. pp. 234–237. ACM (2017)
43. Wombat Security Technologies: State of the Phish: Effectively Reducing Phishing and Malware Infections. `http://pittsburgh.issa.org/ISSA\%20Pittsburgh\%20Wombat\%20Security\%20May\%206\%202016.pdf` (2016), accessed 18 May 2017
44. Yang, C.C., Tseng, S.S., Lee, T.J., Weng, J.F., Chen, K.: Building an anti-phishing game to enhance network security literacy learning. In: ICALT. pp. 121–123 (2012)
45. Zielinska, O.A., Tembe, R., Hong, K.W., Ge, X., Murphy-Hill, E., Mayhorn, C.B.: One Phish, Two Phish, How to Avoid the Internet Phish. Human Factors and Ergonomics Society 58(1), 1466–1470 (2014)