Phishing awareness and education – When to best remind?

Benjamin Berens*, Katerina Dimitrova*, Mattia Mossano* and Melanie Volkamer*
*Karlsruhe Institute of Technology, Germany
Email: benjamin.berens@kit.edu
katerina.dimitrova@student.kit.edu
mattia.mossano@kit.edu
melanie.volkamer@kit.edu

***Abstract*—The use of security awareness and education programmes is very common in organisations. But how effective are they over time? Some initial research on this question is, among others, the extensive study of Reinheimer et al. [74] that measured effectiveness at several time intervals. Their research found still significantly better results than before the awareness program after four months, but no longer after six months. This left open a two months interval for the reminder. The contribution of our paper is to study whether the reminder should be closer to four or six months. Thus, we measured effectiveness after five months. With still significant better results than before the programme after five months, we conclude that it is recommended to remind users more towards six months rather than already after five. However, we kindly invite the community to conduct more long-term studies, in different contexts, to confirm these findings.**

I. Introduction

It is challenging for organisations to establish and maintain an adequate level of information security. This challenge is met with various measures, including rolling out *security awareness and education programmes* and asking the employees to participate on a regular basis – once a year, every other year or less often. But what is the basis to decide when to remind an employee's knowledge?

Although security awareness and education programmes are widely deployed, an evaluation of their effectiveness over an extended time period is often missing. However, it is crucial to know *when* awareness and knowledge levels should be reminded: Too early, the employees are not only less motivated to participate, but needlessly siphoning out working time, damaging the organisation itself. Too late, it generates a serious information security risk increase.

In order to gain more insights, Reinheimer et al. [74] adopted and evaluated over time the phishing awareness and education measure from [69]. They conducted a field study evaluation in a German organisation from the public administration sector. The programme was intended as on-site, mandatory, face-to-face tutorials, but due to the huge number of tutorials required for this organisation, they employed a train-the-instructor approach. The authors collected data before, right after, four, and six months after the tutorials to measure how long the improvement in distinguishing phishing from legitimate emails lasted. Note, although the tutorials were mandatory, participation in the study was optional. The results of Reinheimer et al. [74] show that the improvement wore off after six months.

Reinheimer et al. [74] leave it open when in the interval between four and six month it is recommended to remind the awareness and the knowledge. We aim to answer this question by conducting another retention study while measuring the effect after five months. We used the phishing awareness and education measure from [74] and adapted it to the university context. We integrated the content in our university's e-learning platform, extended it with small exercises after each topic and advertised it to students of our university. The exercises had to be passed in order to continue to the next topic. Participants were asked to fill out a quiz three times: before starting the programme, immediately after having passed it and after five months. Thus, we evaluated the participants' skills in distinguishing phishing from legitimate emails at three points in time. The quiz was similar to the one used in [74], but, instead of using artificial emails and service providers, we used real emails from well known service providers. Our results show that the significant improvement found immediately after passing the programme still lasts five months later.

Although our research follows previous research, the study does bring important further insights in several areas:

1) The timing seems to be in the 6-month range
2) Another form that provides better results by combining previous studies and findings
3) Using real organisations for the examples vs. artificially created ones does not make much difference.

Therefore, we conclude from both ours and [74] results that it is worth to wait until six month to remind awareness and knowledge. However, we would also like to encourage the community to conduct more such studies to both get further evidences and help all the organisations out there that try hard to maintain an adequate level of information security.

Usable Security and Privacy (USEC) Symposium 2022
28 April 2022, San Diego, CA, USA
ISBN 1-891562-79-7
https://dx.doi.org/10.14722/usec.2022.23075
www.ndss-symposium.org, https://www.usablesecurity.net/USEC/

II. Related Work

This section provides both phishing definitions and related works from the research literature. Related work is discussed with regard to the different types of security interventions, the study designs used and the user group types employed. We then present the research on the impact of phishing security awareness and education measures.

Phishing Definitions: There is no clear definition of phishing. However, the focus falls on two main aspects: (1) phishing where attackers deceive users in order to access sensitive information (e.g., passwords, personal data, bank details) using authentic-looking phishing emails or web pages [11], [12], [26], [47], [55], [55], [56], [73], [80] and (2) phishing where attackers spread malware through links or attachments [8], [25], [33], [34], [42], [51], [70], [78], [87], [91]. *We consider both (1) and (2) as phishing and our security awareness and education measure addresses them accordingly.*

Types of Interventions[1]*:* A range of tools (or UI designs for such tools) are created to provide anti-phishing support to users (e.g., additional security indicators or existing security indicators displayed in different ways) [1], [3], [13], [22], [27], [31], [48], [54], [55], [57], [59], [73], [82], [93], [94], [96], [97]. Various studies evaluated security awareness and education measures in different formats: videos [36], [89], games [6], [7], [16], [17], [40], [52], [80], various on-site instructor based tutorials [18], [84], [86], [95] and a multitude of text-based measures [2], [37], [50], [53], [69], [78], [84], [86], [88], [91], [95], [98]. An overview of phishing interventions is provided in [35]. *The security awareness and education measure we study is implemented as e-learning. We consider it between text-based measure and gaming approach.*

Furthermore, there are various researches on evaluating users' skills detecting phishing attacks without any interventions [4], [5], [8]–[10], [19], [23]–[25], [28], [30], [34], [38], [39], [42], [43], [45], [46], [49], [60], [61], [66], [67], [70]–[72], [75], [76], [81], [85], [87], [90], [92] (e.g., to understand decision making, to identify a baseline, or to motivate further research).

Study Designs: Various types of lab studies have been employed both with a cover story [2], [9], [10], [26], [27], [51], [57], [71], [77] and without one [5], [7], [28], [31], [84], [87], [95], i.e., having security as participants' primary goal by telling them the goal of the user study. A number of remote studies have been carried out, including various types of online surveys, with phishing messages sent to the study participants own email accounts (not study-specific) [23], [28], [31], [49], [67], [72], [86], [88], as well as to remotely accessible study-specific accounts [73], [76], [90], [91], [98]. Surveys are of two types: (1) showing screenshots to be judged either as phish or as a legitimate message [46], [58], [85] *as we did in our study.* In some cases real phishing emails were used; others used examples created by the researchers. Otherwise, (2) online surveys asking general questions such as the definition of phishing and the existing attack types [15], [43], [44], [65].

Types of User Groups: Studies have targeted different user groups, i.e., mixed groups on a variety of panels without deliberately isolating specific kinds of participants [11], [26], [31], [46], [58] such as employees [19], [33], [39] or students [4], [7], [9], [10], [24], [45], [71], [72]. *Our target users were students.*

Retention Periods of Security-Related Training: While most of the previously mentioned phishing studies evaluated the impact of the their interventions straight after roll-out, a few evaluated the effect after some time. These studies showed that the effect held, but did not systematically determined for how long. These retention studies were mainly conducted in the context of security awareness and education measures. In [21], [50], retention was evaluated after approximately a month. Authors could show that the effect still lasts. In [95], retention was evaluated after 45 days. Authors could show that the effect still lasts. In [89], the retention was evaluated after 8 weeks. Authors could show that the effect still lasts. In [17], retention was evaluated after 5 months. They also showed that the effect was still significant. *However, compared to our study they only focused on phishing emails with links and used less examples.* [64] examined the ability to judge insecure password-related behaviour. The participants received awareness-raising materials and were tested again after 6 months. The participants were able to retain significant knowledge. *In our case, we repeat the study after 5 months.*

III. Security Awareness and Education Measure

The measure we adapted from [74] consists of the following 12 chapters[2]:

1) **Introduction to phishing** - Phishing general introduction and its possible consequences.
2) **Plausibility checks** - How to determine the probability that the supposed sender has sent a certain email. This includes how to check for the sender's email address and not just the sender's name.
3) **Detect dangerous actions (sensitive data, transfers, calls)** - How to detect dangerous requests in an email content (e.g., money transfers, (expensive) phone calls).
4) **Find the web address** - How to determine the web address (also called URL) behind a link, a button or a QR code. In particular, the web address being placed in the status bar or in a tooltip (e.g., for Outlook users).
5) **Structure of web addresses** - Explanation of the web address structure. In particular, the important parts for distinguishing between phishing and legitimate web addresses.
6) **Easily recognisable tricks in web addresses** - Examples of obvious phishing web addresses and introduction to tricks increasing tempering detection difficulty. This includes sub domain based tricks (e.g., amazon.com.shopping-24.com) and path based tricks (e.g., shopping24hours.com/https://www.amazon.com).

[1] Interventions can be both tools or security awareness and education measures.

[2] In the final version, we will provide a link to the measure we used for this evaluation.

7) **Tricks: small deviations in the domain** - Information on easily overseen visual tricks that make a web address looking legitimate (e.g., arnazon.fr).
8) **Tricks in web addresses that can only be recognised with tools** - How to determine the legitimacy of a link with non-descriptive domain (e.g., a short URL).
9) **Recognisable only if the domain is known** - Attacks only recognisable if the user knows all legitimate domains (e.g,. whether amazon-shop24.es is from amazon or not).
10) **Dangerous Files - introduction** - Introduction to dangerous attachments in messages.
11) **Dangerous files - potentially (very) dangerous file formats** - How to differentiate between potentially dangerous and dangerous file extensions.
12) **Dangerous files - handling potentially dangerous file formats** - How to deal with potentially dangerous attachments.

After each chapter there are exercises deepening the newly acquired knowledge. Their correct completion is needed to continue to the next chapter, ensuring that users paid attention and understood the information. Another feature of our security awareness and education measure is that it can be interrupted and resumed at a later point. Since much information is given, it might be sensible to carry out the security awareness and education measure in multiple sessions.

IV. Methodology

We first introduce the hypotheses. Then, we discuss our study design decisions. Afterwards, we provide details about the study, i.e., recruitment, email examples used, and the study procedure, as well as the ethical considerations.

A. Hypotheses

There is one unavoidable pre-condition to study our main hypothesis: security awareness and education measure must significantly strengthens peoples' skills in distinguishing between phishing and legitimate emails. More precisely, the pre-condition hypothesis is:

Participants are better at distinguishing between phishing and legitimate emails directly after studying our security awareness and education measure than before.

If this pre-condition holds, then our main hypothesis is:

Participants are better at distinguishing between phishing and legitimate emails five months after studying our security awareness and education measure than before.

B. Design Decisions for Study Design

Due to the COVID-19 pandemic, we decided for a study design enabling remote participation. Similar to Reinheimer et al. in [74], we evaluated participants' skill in distinguishing between phishing and legitimate emails by asking them to judge a set of emails. The authors of [74] used email screenshots to assess their participants' ability to detect phishing emails. Their reason was that a phishing campaign would have required sending one email per day, requiring a long time to collect the data. Since the retention interval were very strict (4, 6, 8 and 12 months), the long data collection time would have made the results unreliable: the data would have been collected over weeks, not at the exact interval required. Moreover, [74]'s study participants were from a public organisation in the German Federal Government, hence, conducting a phishing campaign would have required further permissions from the work council. In our case, none of these restrictions applies. However, to make the two studies comparable, we decided to use screenshots as well.

One difference between Reinheimer et al.'s study design and ours is that they choose a between-subject design, while we opted for a within-subject one. The main reason was the potential low number of participants, given the usual dropout rates of retention studies. We consider the consequences of our decision in the limitation section.

C. Recruitment and Reimbursement

We took the following actions to recruit students of our university:

- Creation of posts in social media (Twitter and Facebook)
- Hand out of flyers in student dormitories and on campus
- Word-of-mouth

The security awareness and education measure completion time was measured in pre-studies with students, who required 108 minutes on average. We expected that filling out each of the three surveys took the participants between 15 and 20 minutes. Thus, overall, participation should take around 180 minutes / three hours. Participants who completed the first two phases (see SectionIV-D) received €20. Those who completed all three phases received a reimbursement of €40. The reimbursement was calculated to be above the German minimum wage at the time (€9,50/h).

D. Study Design

The study is structured over three phases. The first one consists of the introduction and the Pre-Knowledge-Quiz. The second phase see the security awareness and education measure delivery and the Post-Knowledge-Quiz. Ultimately, after five months the participants were contacted again and participated in the third phase of the study, consisting of the Retention-Quiz. The overall structure is depicted in Figure 1). Participants were compensated both after completing Phase 2 and Phase 3. We will now describe each step in greater detail.

- **Phase 1 - Introduction** - All participants received an introductory email with general information on the study and instructions on how to take part in it.

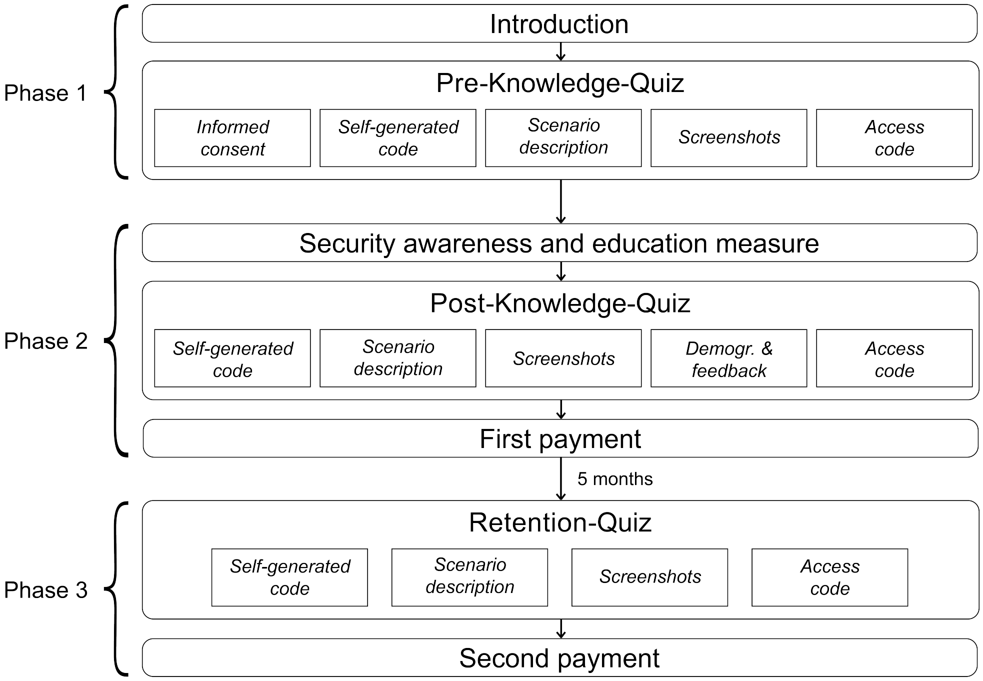


Fig. 1: Overview of the study design.

- **Phase 1 - Pre-Knowledge-Quiz** - The introductory email also contained a link to the SoSci Survey platform hosting the Pre-knowledge-Quiz. Once participants accessed the platform, they were asked to agree to participate in this study (for the *informed consent* text see Section IV-F).
  They were then asked to enter a *self-generated code* serving as pseudonym (for more information see Section IV-F). This code was used to link the three data sets of the same person to each other without knowing his/her identity.
  Afterwards, the participants saw a short description of the *scenario*. They should have pretended:
  - to be Martin Müller,
  - to have a colleague named Jonas Schmidt (sender of some emails),
  - to speak both English and German,
  - to use every service mentioned in the examples,
  - to use every operational system and device seen in the examples (i.e., Microsoft Windows, Apple OSX, Google Android, Apple iOS),
  - to use every email client seen in the examples (i.e., Thunderbird, Apple Mail, Google Mobile Mail, GMX Mobile Client).

  Participants then saw *email screenshots* in German in random order, in a quiz-like style (see Section IV-E for an overview of these screenshots). An example of the screenshot seen by the participants is in the AppendixA For each screenshot, we asked the participants to decide whether the displayed email was a phishing or a legitimate one.
  Each participant then received an *access code* to be sent via email to the study administrator. Note, the access code was the same for all participants to avoid linking survey entries and real identities.

- **Phase 2 - Security Awareness and Education Measure** - After sending the access code to the study administrator, the participants received a link to the actual security awareness and education measure. They were asked to complete it and to send via email to the study administrator the certificate obtained at the end of the measure. The awareness measure provided can be consulted at https://opencourses.kit.edu/goto.php?target=crs_849&client_id=opencourses.

- **Phase 2 - Post-Knowledge-Quiz** - After sending the certificate to the study administrator, the participants received a link to the the SoSci Survey platform containing the Post-Knowledge-Quiz. The quiz was similar to the Pre-Knowledge-Quiz, i.e., it started with the *self-generated code* step, then the *scenario description* step, and then the *email screenshots* to be judged. Note, the screenshots were the same, but displayed in random order, i.e., in a different order than in the first quiz. In addition, we asked them demographic questions (i.e., gender, age) and, if so desired, we gave them the possibility to provide feedback on the security awareness and education measure. Each participant then received an *access code* to be sent via email to the study administrator. While the code was again the same for everyone, it was a different one than the one used in the Pre-Knowledge-Quiz phase.

- **First Payment** - Those participants who sent the second access code were paid for their participation in the second phase.

- **Phase 3 - Retention-Quiz** - 5 Month later, all participants who sent the access code to the study administrator received another email requesting if they wanted to participate in the third phase of the study. If agreed, they received an email containing a link to the last SoSci Survey quiz. This quiz was again similar to both the Post-Knowledge-Quiz and the Pre-Knowledge-Quiz, i.e,. it started with the *self-generated code* step, then the *scenario description* step, and then the *email screenshots* to be judged. Each participant then received a third *access code* to be sent via email to the study administrator.
- **Second Payment** - Those participants who sent the last access code were then paid again for their participation.

E. Email Screenshots

We wanted to use the same types of email screenshots as those in [74]. However, to reduce the potential confusion caused by unfamiliar emails, we decided to switch from artificial services to real ones well known by Germans (i.e., Lufthansa, Paypal, DHL[3], Amazon, Vodafone, YouTube). Furthermore, all emails are in German, due to non-native speakers of the language used in a phishing email being more susceptible to it, as shown in [41].

The same number of phishing and legitimate emails were used in the study. For the creation process, we started with the legitimate emails, all original ones send from the services to one of the authors. We then took a copy of the legitimate email to generated a phishing one, by either changing the sender address, the text content, the URL behind a link or the file type of the attachment. Note, while we wanted to cover all chapters, some did not fit the quiz-like approach used in the study. Specifically: the general introduction (chapter 1), actions unrelated to links and attachments (chapter 3) and tricks recognisable only knowing the domain (chapter 9). In these cases, participants would need additional information other than just a screenshot. We discuss this decision in the limitation section.

The phishing tricks we used are the same ones as those in [74]:

- Simple ones, with modified sender or text content (matching content in chapter 2);
- Tricks that change only the URL behind the link (matching content in chapter 4, 5, 6 and 7). Specifically, we distinguished between:
    - Non-brand related domain, e.g., https://www.hisoliajo.host547.com/ web3/HoEv /ksokGkd=ad3/kol45G for https://www.lufthansa.com/de/ miles-and-more/meilenvergabe;
    - Non-brand related domain + brand outside; e.g., https://amazon.de.kolwerg.com /596ksokGkd89=adweb3/HoEv for https:// amazon.de/596ksokGkd89=adweb3 /HoEv;
    - Small deviations in the domain; e.g., vodafon.de instead of vodafone.de;
    - Special link manipulation (either because the link text is an URL and it does not match the actual URL behind it or because a fake tooltip is programmed into the email to contain, a legitimate URL, while the actual URL behind the link is different);
- Tricks where only the file type of the attachment is changed (matching content in chapter 10, 11 and 12), e.g., from .pdf to either .exe or .pdf.exe.

We used a mixture of different contexts in which one can receive emails, i.e., both mobile and desktop, and for the latter, URL displayed both in the status bar and in a tooltip. An overview of the phishing tricks is provided in Table II.

F. Ethics and Data Protection

We used SoSci Survey to collect the data, as they are compliant with the European Data Protection Regulation (GDPR).

The ethical requirements of our university were met: We informed participants of the goal of the study – both during recruitment and in the introductory email. They were also informed about their rights both in the email and on the first page of the Pre-Knowledge-Quiz. Finally, they had to give consent in order to continue with the quizzes in SoSci Survey. We also informed them that they could withdraw their consent at any time, e.g., by not continuing the study (which includes not finishing any of the survey or not finishing the security awareness and education measure) or by informing the study administrator – without giving any reason.

In order *not* to use their real names, but still be able to link the three data sets, participants were asked to create a code following four steps:

1) Please name the first and last letter of the first name of your mother (e.g., Anne = AE).
2) Please name the first and last letter of the first name of your father (e.g., Thorsten = TN).
3) Please name the first and last letter of your first name (e.g., Michaela = MA).
4) Please name your mother's birthday day (e.g., 17 July 1950 = 17).

The resulting code from the example mentioned above would be then AETNMA17.

V. Results

We first provide information about our participants, then the analyses methods used and, afterwards, we present the results for the two hypotheses.

A. Participants

Of the original 46 participants that started the study only 20 finished, i.e., participated in the retention quiz and sent the third code. Some participants seemed not to have

[3] German post service

used the same code over the three phases, so we could only evaluate data from 16 of the 20 that finished. Those 16 participants had a mean age of 24.69, with 9 being male and 7 being female. The participants studied different subjects, ranging from sports to culture science. Only one participant studies computer science, while six more may have taken computer science courses, as their degree plan contains a few of them.

B. Analysis Methods

We use the *Signal Detection Theory* (SDT), as presented in [83], to test the hypothesises regarding the ability to distinguish between phishing and legitimate examples. According to SDT, participants' answers in any yes or no task are based on a *decision variable*. If the decision variable is high enough, the answer will be yes, otherwise no. This variable is represented by the *criterion* (C). A low criterion represents the tendency of answering "yes" regardless of the stimulus, while the optimum (i.e., a neutral orientation), is a criterion of zero. *Sensitivity* (d') describes instead the ability to distinguish a stimuli from the noise. The greater the sensitivity, the greater is a participant's skill to distinguish a specific stimulus from the noise.

This theory has been applied in various studies in the context of phishing and, by combining signal and noise, it adds value when looking only at correct answers [9]–[14], [29], [62], [63], [65], [68], [79], [80]. The signal are the phishing examples, whereas the noise are the legitimate ones. Criterion describes the tendency of participants to classify more examples as either phishing or legitimate. The larger C is, the more examples are classified as phishing ones. For sensitivity, instead, the larger d' is, the greater is a participant's skill to distinguish between phishing and legitimate examples.

We started the evaluation process with the sensitivity (d') and we then repeated it for the criterion (C). At the beginning of each one, we checked the necessary assumptions for repeated measure ANOVA [32].

Assumption 1: Dependence of the measurements. The measurement data is dependent, in the sense that the data was collected from one person over three measurement times.

Assumption 2: Interval scaling of the data. The measurement data is interval scaled, in the sense that it is about the number of correctly or incorrectly recognised examples (legitimate or phishing).

Assumption 3: The dependent variable is normally distributed. The measurement data was tested for normal distribution using the Shapiro-Wilk normality test. The test did not yield significant results for all 3 measurement time points, so we reject the hypothesis of a deviation from normal distribution for all 3 measurement time points ($p > 0.05$).

Assumption 4: Outlier correction. We checked for outliers in the data. For this purpose, we created boxplots to identify extreme values. No outliers were identified that would severely skew the results.

Assumption 5: Sphericity. The measurement data was tested using Mauchly's test of sphericity. In the course of calculating the ANOVA with repeated measures, a Greenhouse-Geisser sphericity correction is applied directly if sphericity is violated, and the data is reported directly with this.

Afterwards, we performed pairwise paired t-tests to check for differences between the measurements time points using the Bonferroni multiple testing correction. Note, the presence of difference is important *only* for the sensitivity, as the ideal case for the criterion is a neutral C (close to 0).

C. Results for Hypotheses

In addition to the SDT values, we also report the detection rate descriptive values in % of the legitimate, phishing and overall examples. The values achieved can be seen in Table I.

Quiz	Legitimate	Phishing	Overall
Pre-Knowledge	80.02%	63.24%	71.63%
Post-Knowledge	91.18%	90.44%	90.81%
Retention	93.75%	84.93%	89.34%

TABLE I: Descriptive detection rate values of Pre-Knowledge, Post-Knowledge and Retention quizzes.

1) Sensitivity: Sensitivity was statistically significantly different across the three measurement time points, $F(2,30) = 31.037$, $p < 0.0001$, eta2[p] = 0.674.

The t-tests revealed that both Post-Knowledge-Quiz ($p < 0.0001$) and Retention-Quiz ($p < 0.0001$) were significantly different from Pre-Knowledge-Quiz. In contrast, Post-Knowledge-Quiz and Retention-Quiz were not significantly different from each other ($p = 0.292$).

Based on the descriptive statistics, it can be seen that there is initially a low sensitivity ($d' = 1.28$, SD = 0.935). This increases towards the Post-Knowledge-Quiz ($d' = 2.66$, SD = 0.76) and then drops again slightly at the Retention ($d' = 2.47$, SD = 0.422).

Thus, the sensitivity of the participants is significantly higher directly after the security awareness and education measure as well as five months later. Participants therefore are significantly better in distinguishing between phishing and legitimate examples.

2) Criterion: Criterion was statistically significantly different across the three measurement time points, $F(2,30) = 31.037$, $p = 0.0018$, eta2[p] = 0.234.

The t-tests revealed that none of the comparison were significantly different from each other: Pre-Knowledge-Quiz to Post-Knowledge-Quiz ($p = 0.066$), Pre-Knowledge-Quiz to Retention-Quiz ($p = 1$) and Post-Knowledge-Quiz to Retention-Quiz ($p = 0.053$).

Again based on the descriptive statistics, we can see that initially there is a tendency to classify examples as phishing ($C = 0.27$). At Post-Knowledge-Quiz time, this tendency decreases to almost neutral ($C = 0.02$). In contrast, the tendency increases again at the Retention-Quiz time, even if it does not reach the level of Pre-Knowledge-Quiz (C = 0.247).

Thus, the criterion is not significantly different directly after the security awareness and education measure as well as five months later. Participants show no trend in either being overcautious or careless in deciding between phishing and legitimate examples.

D. Performance for Different Phishing Types

In this subsection, we report the performance of the different phishing trick types (see Table II). The one with the best performance are the "non-brand related domain" ones. Almost the same performance was shown for the "non-brand related domain + brand outside" trick type (i.e., either the brand is in the subdomain or the path) and the "special link manipulation" ones. Those based on the "content" and the "attachment" are with almost 85% recognition after five months. Note, the "content" ones already showed a very high recognition rate (81%) even before taking the security awareness and education measure. The one trick type participants struggled the most with are the "small deviations in the domain" ones.

E. Feedback

We explicitly asked for feedback during the Post-Knowledge-Quiz phase. Participants had at the end of the Retention-Quiz phase a text field for general comments, which some also used to provide feedback.

Participants, in general, liked the security awareness and education measure used, stating that they learned a lot and felt better prepared and protected. Some stated at the end of the Retention-Quiz phase that they are now more careful with emails than before taking part in the study.

We identified two main areas for improvements: (1) Each chapter of the security awareness and education measure ends with a quiz which currently can be either passed or failed. In case of failure, one is redirected to the beginning of the chapter to check the content again and then try the quiz again. Some participants recommended to provide feedback after the quiz on the specific answers/tasks performance, with mistake explanations. (2) Some participants mentioned shortcomings, albeit these were caused by the e-learning platform on which the security awareness and education measure is implemented: The resolution of the images is not very high, causing issues for some of those containing screenshots of emails with the URL in the statusbar. Moreover, moving from the successful quiz of one chapter to the content of the next chapter takes too many clicks.

One participant advised to use more videos, instead of reading text, and to modify the feedback given (currently it is e.g., "great this is correct", "good job", "that was great") as one feels like at grammar school.

VI. Discussion

Our study design was similar to the one in [74]: We studied participants' skill in distinguishing between phishing and legitimate emails in best case scenarios, i.e., with security being their primary goal and (in link-related screenshots) with the URL already being displayed in the statusbar or tooltip respectively[4]. Not being able to identify a particular type of phishing in such a scenario means that it is very likely not detected it in real life either. Our Pre-Knowledge-Quiz results shows that participants have serious issues with phishing detection even in this best case scenario, as the overall phishing detection rate is only 63%.

Participants' skills in distinguishing between phishing and legitimate emails improved significantly right after attending the security awareness and education measure. This result is inline with the results from [74].

Note, while the Pre-Knowledge-Quiz phishing detection rate in [74] (62%) was similar to our result (63%), in our study the Post-Knowledge-Quiz phishing detection increased to 90%, while it only increased to 80% in [74]. Thus, switching from an instructor-based tutorial approach to an e-learning based one seems not to have affected in a negative way the effectiveness of the content.

Note, however, that our sample consisted of students used to interact with the e-learning platform, hence the general population is likely to have a different performance. Moreover, while 90% seems to be a high number, it is worth mentioning that we used a quiz-based approach, i.e., in real life this number is likely to be smaller. Testing multiple examples in the quiz could also be some kind of training the previous attained knowledge compared to asking questions about definitions that themselves not actually train the knowledge. Finally, the final sample consisted of 16 participants, which is a small amount. This could lead to non-generalisable results. Nonetheless, it is still worth discussing the performance for the different tricks to determine how to improve the situation: The worst ones are those with small deviations in the domain. While we still believe that it is worth making people aware of these tricks, it is even more important for the domain owners to check whether someone tries to register a domain name similar to their own. As a consequence, it would get less likely for users to get into a situation where they have to notice small deviations in the domain. Furthermore, it may also be worth for the developers of email environments to reconsider how they display the domain in the URL. We leave it as future work whether there are alternative, more effective ways (e.g., a*r*n*a*z*o*n) to help people detecting small deviations in the domain.

As we were surprised that only 75% detected the phishing email with the nonsense content after five months, we had another look at this example. It actually informs about having received the information to end the contract with Vodafone. If this was a mistake, one should send back the following information via email: name, address, birthday, bank account details, name of others mentioned in the contract. In chapter 2 of the security awareness and education measure it is explained that one trick is to ask to send back information via email. It is also stated that avoiding this is critical, even if no password is requested. We are wondering whether those participants that did

[4] Participants did not need to move the mouse to the link. The mouse was already at the relevant link.

Phish Description	**Pre-Quiz**	**Post-Quiz**	**Retention-Quiz**
Content	*81,25%*	*90,63%*	*84,38%*
Sender	75,00%	100,00%	93,75%
Nonsense Content	87,50%	81,25%	75,00%
Non-Brand related Domain	*65,63%*	*100,00%*	*100,00%*
Random URL; 'click here'; Statusbar	50,00%	100,00%	100,00%
Random URL; 'click here'; Tooltip	81,25%	100,00%	100,00%
Non-Brand related Domain + Brand Outside	*73,44%*	*95,31%*	*98,44%*
Brand in Subdomain; Statusbar	62,50%	93,75%	100,00%
Brand in Subdomain; Tooltip	62,50%	93,75%	100,00%
Brand in Path; Statusbar	75,00%	100,00%	93,75%
Brand in Path; Tooltip	93,75%	93,75%	100,00%
Small Deviations in the Domain	*45,83%*	*70,83%*	*41,67%*
Typo in domain; Statusbar	18,75%	37,50%	12,50%
Typo in domain; Tooltip	56,25%	81,25%	37,50%
Swap letters in domain; Dialog	62,50%	93,75%	75,00%
Special Link Manipulation	*64,06%*	*93,75%*	*96,88%*
Random URL; Mismatch Statusbar	56,25%	100,00%	100,00%
Random URL; Mismatch; Tooltip	81,25%	93,75%	100,00%
Random URL; Mismatch; Dialog	87,50%	100,00%	100,00%
Random URL; Fake Tooltip; Statusbar	31,25%	81,25%	87,50%
Attachment	*46,88%*	*93,75%*	*84,38%*
File extension .exe	37,50%	93,75%	81,25%
File extension .pdf.exe	56,25%	93,75%	87,50%
Overall	**63,24%**	**90,44%**	**84,93%**

TABLE II: Percentages of correct answers per phishing email. Phishing types and their results are in italics.

not identified it as phishing simply have a different (you may call it *incomplete*) mental model of what phishing is. Another reason might be the absence of link or attachments in the phishing email, as this might increase its apparent legitimacy.

With respect to the performance after five months, we found that participants' skills in distinguishing between phishing and legitimate emails is still significantly better than before the security awareness and education measure. With 85% phishing detection, the detection rate in our study after five months is actually higher than it was in the study presented in [74] after four month (with 71%). This may have several reasons: (1) our sample is smaller and consisted only of students of a relatively young age group. (2) Our security awareness and education measure is slightly different, not in the content, but because participants have to pass a small mandatory exercise after each chapter to proceed. (3) In our study design, participants see the screenshots before, as they are the same between the three phases (albeit the order is random).

While the authors of [17] only studied phishing emails with links and studied less phishing tricks related to URLs, they showed for their security awareness and education measure that the effect was still significant after five months. We confirmed their finding that it might not yet be necessary to remind people after five months. Thus, while Reinheimer et al. in [74] left the reader with a time frame of two months for the reminder, from our results it seems better to remind users closer to six months, rather than after four. This could also be due to the more interactivity and personal involvement of our measure. Together with the better results of the post-knowledge quiz, we think that six months should be sufficiently supported to be the next refreshment step.

Also, the use of the SDT has shown that it has advantages over the pure report of the detection rate of fraudulent or legitimate messages. The values of SDT, especially the Criterion, show whether the measure influences the response tendency in one of the two directions. For example, the measure could cause participants to become overcautious, which would mean a tendency toward phishing messages. Also, the measure could make participants careless, which would mean a tendency toward legitimate. Neither of these is desirable, as it would be associated with adverse real-world effects. The results show that there is initially a tendency toward a neutral and thus optimal decision and that this then develops back in the direction of being overcautious. This is a clear starting point for future developments, not only to constantly improve better recognition but also to avoid the tendency toward being overcautious in the long run. One possibility here could be short refreshers of knowledge to demonstrate to people that an overcautious manner is not necessary.

1) Limitations.: Our study have security as the primary task of the participants and each example is shown in the best-case scenario, i.e., with the link already hovered and showing the URL behind it. As this is a best-case scenario, we considered that, if the participants were to fail in such conditions, it would have shown a deeper issue with phishing recognition. Our results show that, even in these conditions, it is clear that participants still show significant difficulties distinguishing phishing and legitimate emails.

Similarly to all the phishing studies mentioned in the related work section, we had to restrict the number of phishing emails covered in our user study. We based our selection on the one from Reinheimer et al. [74] to enable an easier comparison between the two evaluations. However, it might be that different combinations of services, email content and tricks result in a different performance. Nonetheless, we believe that the selection is representative enough in order to draw conclusions, as the examples are similar to those made by the related works.

Note, we excluded two types of phishing tricks explained in our security awareness and education measure from the email screenshots used in our quiz: 1) URLs used behind links that can only be recognised with tools (such as the short URLs explained in chapter 8) and 2) domains that can only be recognised as phishing if all legitimated domains are known (using tools that reveal, e.g., information regarding when the domain was registered, as explained in chapter 9). These tricks are actually applied in real life. Therefore it is important that people are aware that they exists. However, besides making people aware of these tricks, it is important to also support them with corresponding tools or extensions of email environments.

The student sample comes with some limitations. Moreover, the final number of participants (16) is somewhat small. A larger set of participants would give the chance to have a more diverse sample and therefore strengthen the generalizability of the results to other populations. Hence, we strongly suggest that studying the same security awareness and education measure with other user groups is highly recommend to get more evidence about the time recommended to remind people about the security awareness and education measure.

VII. Conclusion

Our research sheds more light on the question when to best remind people after having taken a security awareness and education measure. The research from Reinheimer et al. [74] concluded that it is best to remind (at least for the first reminder) between four and six months. With a similar security awareness and education measure and a similar study design, we showed that – at least for our setting – reminding after six months seems to be early enough. Our participants were still significantly better at distinguishing legitimate and phishing emails after five months, compared to before taking the security awareness and education measure, albeit our sample consisted of students and was somewhat small. As this is a really new field of research, more research is needed in this area to guide standards like PCI-DSS [20] to set their request when to remind people based on established research results.

There is one main difference in the security awareness and education measure of our study compared to the one presented in [74]: participants had to pass exercises after each of the 12 chapters to continue into the next one. In the tutor-based approach from [74], instead, it is possible that people were joining the event without being concentrated throughout. This may explain why the post-quiz performance in our study is higher than the one reported in [74]. Thus, a sub-finding of our research is that it is important to ensure that people are actually trying to follow the security awareness and education measure.

Acknowledgment

This work was supported by funding from the topic Engineering Secure Systems of the Helmholtz Association (HGF), by KASTEL Security Research Labs, by the Federal Ministry of Education and Research (BMBF) and by the Google Faculty Research Award 2019 for the Link-centric Phishing Warnings for Online Email Clients project.

References

[1] D. Akhawe and A. P. Felt, “Alice in warningland: A large-scale field study of browser security warning effectiveness,” in *22th USENIX Security CHI*, 2013.

[2] A. Alnajim and M. Munro, “An Anti-Phishing Approach that Uses Training Intervention for Phishing Websites Detection,” *2009 Sixth International Conference on Information Technology: New Generations*, pp. 405–410, 2009.

[3] ——, “An approach to the implementation of the anti-phishing tool for phishing websites detection,” in *2009 International Conference on Intelligent Networking and Collaborative Systems.* IEEE, 2009, pp. 105–112.

[4] I. Alseadoon, “The impact of users’ characteristics on their ability to detect phishing emails.” Ph.D. dissertation, (Doctoral dissertation, Queensland University of Technology), 2014.

[5] M. Alsharnouby, F. Alaca, and S. Chiasson, “Why phishing still works: User strategies for combating phishing attacks,” *International Journal of Human-Computer Studies*, vol. 82, 2015.

[6] N. A. Arachchilage, I. Flechais, and K. Beznosov, “A game storyboard design for avoiding phishing attacks,” in *Proceedings of the 11th Symposium On Usable Privacy and Security (SOUPS)*, 2014.

[7] N. A. G. Arachchilage, S. Love, and K. Beznosov, “Phishing threat avoidance behaviour: An empirical investigation,” *Computers in Human Behavior*, vol. 60, pp. 185–197, 2016.

[8] Z. Benenson, F. Gassmann, and R. Landwirth, “Unpacking spear phishing susceptibility,” in *Financial Cryptography and Data Security*, 2017.

[9] M. Butavicius, K. Parsons, M. Pattinson, A. McCormac, D. Calic, and M. Lillie, “Understanding Susceptibility to Phishing Emails: Assessing the Impact of Individual Differences and Culture,” *International Symposium on Human Aspects of Information Security & Assurance (HAISA)*, 2017.

[10] M. Butavicius, K. Parsons, M. Pattinson, and A. McCormac, “Breaching the human firewall: Social engineering in phishing and spear-phishing emails,” *Australasian Conference on Information Systems*, 2016.

[11] C. Canfield, A. Davis, B. Fischhoff, and F.-A. o. Usable . . . , “Replication: Challenges in using data logs to validate phishing detection ability metrics,” *Symposium on Usable Privacy and Security (SOUPS)*, 2017.

[12] C. Canfield, B. Fischhoff, and A. Davis, “Using Signal Detection Theory to Measure Phishing Detection Ability and Behavior,” in *SOUPS*, 2015.

[13] ——, “Quantifying Phishing Susceptibility for Detection and Behavior Decisions,” *Human Factors: The Journal of Human Factors and Ergonomics Society*, vol. 58, pp. 1158–1172, 2016.

[14] C. I. Canfield and B. Fischhoff, “Setting Priorities in Behavioral Interventions: An Application to Reducing Phishing Risk,” *Risk Analysis*, vol. 38, pp. 826–838, 2018.

[15] G. Canova, M. Volkamer, C. Bergmann, and R. Borza, “NoPhish: An Anti-Phishing Education App,” in *Security and Trust Management.* LNCS, 2014, pp. 188–192.

[16] G. Canova, M. Volkamer, C. Bergmann, R. Borza, B. Reinheimer, S. Stockhardt, and R. Tenberg, “Learn to Spot Phishing URLs with the Android NoPhish App,” in *WISE 9*. Springer, 2015, pp. 87–100.

[17] G. Canova, M. Volkamer, C. Bergmann, and B. Reinheimer, “NoPhish App Evaluation: Lab and Retention Study,” in *USEC*. Internet Society, 2015.

[18] L. Y. Chang and N. Coppel, “Building cyber security awareness in a developing country: lessons from myanmar,” *Computers & Security*, vol. 97, p. 101959, 2020.

[19] D. Conway, R. Taib, M. Harris, S. K. Berkovsky, K. Yu, and F. Chen, “A qualitative investigation of bank employee experiences of information security and phishing,” *Symposium on Usable Privacy and Security (SOUPS)*, 2017.

[20] P. S. S. Council, “PCI DSSv3.2.1 - May 2018,” 2018, https://www.pcisecuritystandards.org/document$_$library/ accessed: 27.02.2020.

[21] E. J. Custers, “Long-term retention of basic science knowledge: a review study,” *Advances in Health Sciences Education*, vol. 15, no. 1, pp. 109–128, 2010.

[22] R. Dhamija and J. D. Tygar, “The battle against phishing: Dynamic Security Skins,” in *Symposium on Usable Privacy and Security*. New York, NY, USA: ACM, 2005, pp. 77–88. [Online]. Available: http://doi.acm.org/10.1145/1073001.1073009

[23] R. Dhamija, J. D. Tygar, and M. Hearst, “Why phishing works,” in *Proceedings of CHI 2006 Human Factors in Computing Systems*. ACM, 2006, pp. 581–590.

[24] A. Diaz, A. T. Sherman, and A. Joshi, “Phishing in an academic community: A study of user susceptibility and behavior,” *Cryptologia*, pp. 1–15, 2019.

[25] J. S. Downs, M. Holbrook, and L. F. Cranor, “Behavioral response to phishing risk,” *Proceedings of the Anti-Phishing Working Groups 2nd Annual eCrime Researchers Summit*, pp. 37–44, 2007.

[26] J. S. Downs, M. B. Holbrook, and L. F. Cranor, “Decision strategies and susceptibility to phishing,” in *Proceedings of the Second Symposium on Usable Privacy and Security*, 2006, pp. 79–90.

[27] S. Egelman, L. F. Cranor, and J. Hong, “You’ve Been Warned: An Empirical Study of the Effectiveness of Web Browser Phishing Warnings,” in *CHI*. ACM, 2008, pp. 1065–1074.

[28] S. Egelman and E. Peer, “Scaling the security wall: Developing a security behavior intentions scale (sebis),” in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 2015, pp. 2873–2882.

[29] I. Embrey and K. Kaivanto, “Many Phish in the C: A Coexisting-Choice-Criteria Model of Security Behavior,” *arxiv*, 2018.

[30] J.-P. Erkkilä, “Why we fall for phishing,” in *Conference on Human Factors in Computer Systems*. ACM, 2011.

[31] A. P. Felt, R. W. Reeder, A. Ainslie, H. Harris, M. Walker, C. Thompson, M. E. Acer, E. Morant, and S. Consolvo, “Rethinking connection security indicators,” in *Twelfth Symposium on Usable Privacy and Security ({SOUPS} 2016)*, 2016, pp. 1–14.

[32] A. Field, *Discovering statistics using IBM SPSS statistics*. sage, 2013.

[33] D. Filipczuk, C. Mason, and S. Snow, “Using a Game to Explore Notions of Responsibility for Cyber Security in Organisations,” in *Proceedings of CHI 2019 Human Factors in Computing Systems*, 2019, pp. 1–6.

[34] W. R. Flores, H. Holm, M. Nohlberg, and M. Ekstedt, “Investigating personal determinants of phishing and the effect of national culture,” *Information and Computer Security*, vol. Volume 23, pp. 178–199, 2015.

[35] A. Franz, V. Zimmermann, G. Albrecht, K. Hartwig, C. Reuter, A. Benlian, and J. Vogt, “Sok: Still plenty of phish in the sea — a taxonomy of user-oriented phishing interventions and avenues for future research,” in *Seventeenth Symposium on Usable Privacy and Security (SOUPS 2021)*. USENIX Association, Aug. 2021, pp. 339–358. [Online]. Available: https://www.usenix.org/conference/soups2021/presentation/franz

[36] V. Garg, J. Camp, L. Mae, and K. Connelly, “Designing risk communication for older adults,” in *Symposium on Usable Privacy and Security (SOUPS)*. Citeseer, 2011.

[37] R. Gonzalez and M. E. Locasto, “An interdisciplinary study of phishing and spear-phishing attacks,” *SOUPS*, 2015.

[38] T. Halevi, J. Lewis, and N. Memon, “Phishing, Personality Traits and Facebook,” *arXiv*, 2013.

[39] T. Halevi, N. Memon, and O. Nov, “Spear-Phishing in the Wild: A Real-World Study of Personality, Phishing Self-Efficacy and Vulnerability to Spear-Phishing Attacks,” *SSRN Electronic Journal*, 2015.

[40] S. Hart, A. Margheri, F. Paci, and V. Sassone, “Riskio: A serious game for cyber security awareness and education,” *Computers & Security*, vol. 95, p. 101827, 2020.

[41] A. A. Hasegawa, N. Yamashita, M. Akiyama, and T. Mori, “Why they ignore english emails: The challenges of non-native speakers in identifying phishing emails,” in *Seventeenth Symposium on Usable Privacy and Security ({SOUPS} 2021)*, 2021, pp. 319–338.

[42] K. W. Hong, C. M. Kelley, R. Tembe, E. Murphy-Hill, and C. B. Mayhorn, “Keeping Up With The Joneses: Assessing phishing susceptibility in an email task,” *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 57, pp. 1012–1016, 2013.

[43] C. Iuga, J. R. Nurse, and A. Erola, “Baiting the hook: factors impacting susceptibility to phishing attacks,” *Human-Centric Computing and Information Sciences*, vol. 6, p. 8, 2016.

[44] A. Jain and B. Gupta, “Phishing Detection: Analysis of Visual Similarity Based Approaches,” *Security and Communication Networks*, vol. 2017, pp. 1–20, 2017.

[45] M. Jakobsson, A. Tsow, A. Shah, E. Blevis, and Y.-K. Lim, “What instills trust? A qualitative study of phishing,” in *Financial Cryptography*. LNCS, 2007, pp. 356–361.

[46] T. Kelley and B. I. Bertenthal, “Real-World Decision Making: Logging Into Secure vs. Insecure Websites,” *USEC*, 2016.

[47] I. Kirlappos and M. A. Sasse, “Security education against phishing:A modest proposal for a major rethink,” *IEEE Security & Privacy*, vol. 10, no. 2, pp. 24–32, 2012.

[48] Y. Kitamura, A. Quigley, K. Isbister, T. Igarashi, P. Bjørn, S. Drucker, K. Althobaiti, N. Meng, and K. Vaniea, “I Don’t Need an Expert! Making URL Phishing Features Human Comprehensible,” *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pp. 1–17, 2021.

[49] S. Kleitman, M. K. Law, and J. Kay, “It’s the deceiver and the receiver: Individual differences in phishing susceptibility and false positives with item profiling,” *PLOS ONE*, vol. 13, p. e0205089, 2018.

[50] P. Kumaraguru, J. Cranshaw, and A. Acquisti, “School of phish: a real-world evaluation of anti-phishing training,” *SOUPS*, 2009.

[51] P. Kumaraguru, Y. Rhee, A. Acquisti, L. F. Cranor, J. Hong, and E. Nunge, “Protecting People from Phishing: The Design and Evaluation of an Embedded Training Email System,” in *CHI*. ACM, 2007, pp. 905–914. [Online]. Available: http://doi.acm.org/10.1145/1240624.1240760

[52] A. Kunz, M. Volkamer, S. Stockhardt, S. Palberg, T. Lottermann, and E. Piegert, “Nophish: Evaluation of a web application that teaches people being aware of phishing attacks,” in *Informatik*, GI. GI, LNI, 2016, pp. 15–24.

[53] E. Lastdrager, I. C. Gallardo, P. Hartel, and M. Junger, “How effective is anti-phishing training for children?” *Symposium on Usable Privacy and Security (SOUPS)*, 2017.

[54] E. Lau and Z. N. Peterson, “A research framework and initial study of browser security for the visually impaired,” in *Symposium on Usable Privacy and Security (SOUPS)*, 2015.

[55] P. Likarish, D. E. Dunbar, J. P. Hourcade, and E. Jung, “Bayeshield: conversational anti-phishing user interface.” in *SOUPS*, vol. 9, 2009, pp. 1–1.

[56] E. Lin, S. Greenberg, E. Trotter, D. Ma, and J. Aycock, “Does domain highlighting help people identify phishing sites?” in *Conference on Human Factors in Computing Systems (CHI)*. ACM Press, 2011, p. 2075. [Online]. Available: http://dl.acm.org/citation.cfm?doid=1978942.1979244

[57] ——, “Does domain highlighting help people identify phishing sites?” in *Proceedings of CHI 2011 Human Factors in Computing Systems*. ACM, 2011, pp. 2075–2084.

[58] G. Liu, G. Xiang, B. A. Pendleton, J. I. Hong, and W. Liu, "Smartening the crowds: computational techniques for improving human verification to fight phishing scams," *Symposium on Usable Privacy and Security (SOUPS)*, 2011.

[59] S. Marchal, G. Armano, T. Grondahl, K. Saari, N. Singh, and N. Asokan, "Off-the-Hook: An Efficient and Usable Client-Side Phishing Prevention Application," *IEEE Transactions on Computers*, vol. 66, 2017.

[60] C. Marforio, R. Jayaram Masti, C. Soriente, K. Kostiainen, and S. Čapkun, "Evaluation of personalized security indicators as an anti-phishing mechanism for smartphone applications," in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, 2016, pp. 540–551.

[61] C. Marriott, "Through the Net Investigating How User Characteristics Influence Susceptibility to Phishing." Ph.D. dissertation, Dublin Institute of Technology, 2018.

[62] J. Martin, "Something Looks Phishy Here: Applications of Signal Detection Theory to Cyber-Security Behaviors in the Workplace," Ph.D. dissertation, University of South Florida, 2017.

[63] J. Martin, C. Dubé, and M. D. Coovert, "Signal Detection Theory (SDT) Is Effective for Modeling User Behavior Toward Phishing and Spear-Phishing Attacks," *Human Factors: The Journal of Human Factors and Ergonomics Society*, vol. 60, pp. 1179–1191, 2018.

[64] P. Mayer, C. Schwartz, and M. Volkamer, "On the systematic development and evaluation of password security awareness-raising materials," in *Proceedings of the 34th Annual Computer Security Applications Conference*. ACM, 2018, pp. 733–748.

[65] C. B. Mayhorn and P. G. Nyeste, "Training users to counteract phishing." *Work (Reading, Mass.)*, vol. 41 Suppl 1, pp. 3549–52, 2012.

[66] J. G. Mohebzada, A. El Zarka, A. H. Bhojani, and A. Darwish, "Phishing in a university community: Two large scale phishing experiments," in *2012 International Conference on Innovations in Information Technology (IIT)*. IEEE, 2012, pp. 249–254.

[67] G. D. Moody, D. F. Galletta, and B. Dunn, "Which phish get caught? An exploratory study of individuals' susceptibility to phishing," *European Journal of Information Systems*, vol. 26, pp. 564–584, 2017.

[68] M. M. Moreno-Fernández, F. Blanco, P. Garaizar, and H. Matute, "Fishing for phishers. Improving Internet users' sensitivity to visual deception cues to prevent electronic fraud," *Computers in Human Behavior*, vol. 69, pp. 421–436, 2017.

[69] S. Neumann, B. Reinheimer, and M. Volkamer, "Don't be deceived: the message might be fake," in *International Conference on Trust and Privacy in Digital Business*. Springer, 2017, pp. 199–214.

[70] D. Oliveira, H. Rocha, H. Yang, D. Ellis, S. Dommaraju, M. Muradoglu, D. Weir, A. Soliman, T. Lin, and N. Ebner, "Dissecting spear phishing emails for older vs young adults: On the interplay of weapons of influence and life domains in predicting susceptibility to phishing," in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 2017, p. 6412–6424.

[71] K. Parsons, A. McCormac, M. Pattinson, M. Butavicius, and C. Jerram, "Phishing for the Truth: A Scenario-Based Experiment of Users' Behavioural Response to Emails," *SEC 2013: Security and Privacy Protection in Information Processing Systems*, pp. 366–378, 2013.

[72] M. Pattinson, C. Jerram, K. Parsons, A. McCormac, and M. Butavicius, "Managing Phishing Emails: A Scenario-Based Experiment," in *Symposium on Human Aspects of Information Security & Assurance (HAISA 2011)*, 2011, pp. 74–85.

[73] J. Petelka, Y. Zou, and F. Schaub, "Put Your Warning Where Your Link Is: Improving and Evaluating Email Phishing Warnings ," in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 2019.

[74] B. Reinheimer, L. Aldag, P. Mayer, M. Mossano, R. Duezguen, B. Lofthouse, T. von Landesberger, and M. Volkamer, "An investigation of phishing awareness and education over time: When and how to best remind users," in *Symposium on Usable Privacy and Security (SOUPS)*, 2020, pp. 259–284.

[75] R. Roberts, Y. Goldschlag, R. Walter, T. Chung, A. Mislove, and D. Levin, "You are who you appear to be: A longitudinal study of domain impersonation in tls certificates," in *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security*, 2019, pp. 2489–2504.

[76] D. M. Sarno, J. E. Lewis, C. J. Bohil, M. K. Shoss, and M. B. Neider, "Who are Phishers luring?: A Demographic Analysis of Those Susceptible to Fake Emails," *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 61, pp. 1735–1739, 2017.

[77] S. E. Schechter, R. Dhamija, A. Ozment, and I. Fischer, "The emperor's new security indicators," in *2007 IEEE Symposium on Security and Privacy (SP'07)*. IEEE, 2007, pp. 51–65.

[78] S. Sheng, M. Holbrook, P. Kumaraguru, L. F. Cranor, and J. Downs, "Who falls for phish?: a demographic analysis of phishing susceptibility and effectiveness of interventions," in *Proceedings of CHI 2010 Human Factors in Computing Systems*. ACM, 2010, pp. 373–382.

[79] S. Sheng, B. Magnien, P. Kumaraguru, A. Acquisti, L. Cranor, J. Hong, and E. Nunge, "Anti-phishing phil: the design and evaluation of a game that teaches people not to fall for phish," *Symposium on Usable Privacy and Security (SOUPS)*, 2007.

[80] S. Sheng, B. Magnien, P. Kumaraguru, A. Acquisti, L. F. Cranor, J. Hong, and E. Nunge, "Anti-Phishing Phil: The Design and Evaluation of a Game That Teaches People Not to Fall for Phish," in *Symposium on Usable Privacy and Security (SOUPS)*. ACM, 2007, pp. 88–99.

[81] A. Siami Namin, R. Hewett, K. S. Jones, and R. Pogrund, "Sonifying internet security threats," in *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. ACM, 2016, pp. 2306–2313.

[82] G. Sonowal, K. Kuppusamy, and A. Kumar, "Usability evaluation of active anti-phishing browser extensions for persons with visual impairments," *2017 4th International Conference on Advanced Computing and Communication Systems (ICACCS)*, pp. 1–6, 2017.

[83] H. Stanislaw and N. Todorov, "Calculation of signal detection theory measures," *Behavior Research Methods, Instruments, & Computers*, vol. 31, no. 1, pp. 137–149, 1999.

[84] S. Stockhardt, B. Reinheimer, M. Volkamer, P. Mayer, A. Kunz, P. Rack, and D. Lehmann, "Teaching Phishing-Security: Which Way is Best?" *IFIP SEC*, pp. 135–149, 2016.

[85] C. Thompson, M. Shelton, E. Stark, M. Walker, E. Schechter, and A. P. Felt, "The web's identity crisis: understanding the effectiveness of website identity indicators," in *28th {USENIX} Security Symposium ({USENIX} Security 19)*, 2019, pp. 1715–1732.

[86] K. F. Tschakert and S. Ngamsuriyaroj, "Effectiveness of and user preferences for security awareness training methodologies," *Heliyon*, vol. 5, p. e02010, 2019.

[87] A. Vance, B. Kirwan, D. K. Bjornn, J. Jenkins, and B. Briton Anderson, "What Do We Really Know about How Habituation to Warnings Occurs Over Time?: A Longitudinal fMRI Study of Habituation and Polymorphic Warnings," in *Conference on Human Factors in Computing Systems (CHI)*, 2017, pp. 2215–2227.

[88] M. Volkamer, K. Renaud, and P. Gerber, "Spot the phish by checking the pruned URL," *Information and Computer Security*, vol. Volume 24, pp. 372–385, 2016.

[89] M. Volkamer, K. Renaud, B. Reinheimer, P. Rack, M. Ghiglieri, P. Mayer, A. Kunz, and N. Gerber, "Developing and Evaluating a Five Minute Phishing Awareness Video," *Trust, Privacy and Security in Digital Business (TrustBus)*, pp. 119–134, 2018.

[90] J. Wang, Y. Li, C. College, H. R. Rao, and T. U. o. T. a. S. Antonio, "Overconfidence in Phishing Email Detection," *Journal of the Association for Information Systems*, vol. 17, pp. 759–783, 2016.

[91] R. Wash and M. Cooper, "Who Provides Phishing Training?" in *Proceedings of CHI 2018 Human Factors in Computing Systems*, 2018.

[92] R. Wash, N. Nthala, and E. Rader, "Knowledge and capabilities that non-expert users bring to phishing detection," in *Seventeenth Symposium on Usable Privacy and Security (SOUPS 2021)*. USENIX Association, Aug. 2021, pp. 377–396. [Online]. Available: https://www.usenix.org/conference/soups2021/presentation/wash

[93] W. Yang, J. Chen, A. Xiong, R. W. Proctor, and N. Li, "Effectiveness of a phishing warning in field settings," *the 2015 Symposium and Bootcamp*, p. 14, 2015.

[94] K.-P. Yee, "Designing and evaluating a petname anti-phishing tool,"

in *Poster presented at Symposium on usable Privacy and Security (SOUPS)*, 2005, pp. 6–8.

[95] T. Zhang, "Knowledge Expiration in Security Awareness Training," *Conference on Digital Forensics, Security and Law (ADFSL)*, 2018.

[96] Y. Zhang, S. Egelman, L. F. Cranor, and J. Hong, "Phinding Phish: Evaluating Anti-Phishing Tools," in *NDSS*. School of Computer Science, Internet Society, 2007.

[97] Y. Zhang, J. I. Hong, and L. F. Cranor, "CANTINA: A Content-Based Approach to Detecting Phishing Web Sites," in *16th International World Wide Web Conference*, 2007, pp. 639–648.

[98] O. A. Zielinska, R. Tembe, K. Hong, X. Ge, E. Murphy-Hill, and C. B. Mayhorn, "One Phish, Two Phish, How to Avoid the Internet Phish," *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 58, pp. 1466–1470, 2014.

Appendix

A. Survey

1) Introduction: Sehr geehrte/r Teilnehmer/in,

wir freuen uns, dass Sie an unserer Studie zur Erkennung von betrügerischen Nachrichten teilnehmen. Ihre Meinung ist für uns sehr wertvoll.
Sie nehmen im Folgenden an einer Studie teil, bei der Sie zunächst Ihre Fähigkeiten zur Erkennung von betrügerische Nachrichten zeigen können. Bitte verwenden Sie dazu keine Hilfsmittel, ihr derzeitige Stand ist wichtig und Fehler sind erlaubt. Die Studie hat das Ziel, herauszufinden, ob die ILIAS Schulung Sie dabei unterstützt betrügerische Nachrichten erfolgreich zu erkennen.
An manchen Stellen werden Sie gebeten Text in ein Feld einzugeben. Um das Quiz auszuwerten, ist es für uns sehr wichtig, dass Sie diese Felder alle ausfüllen.

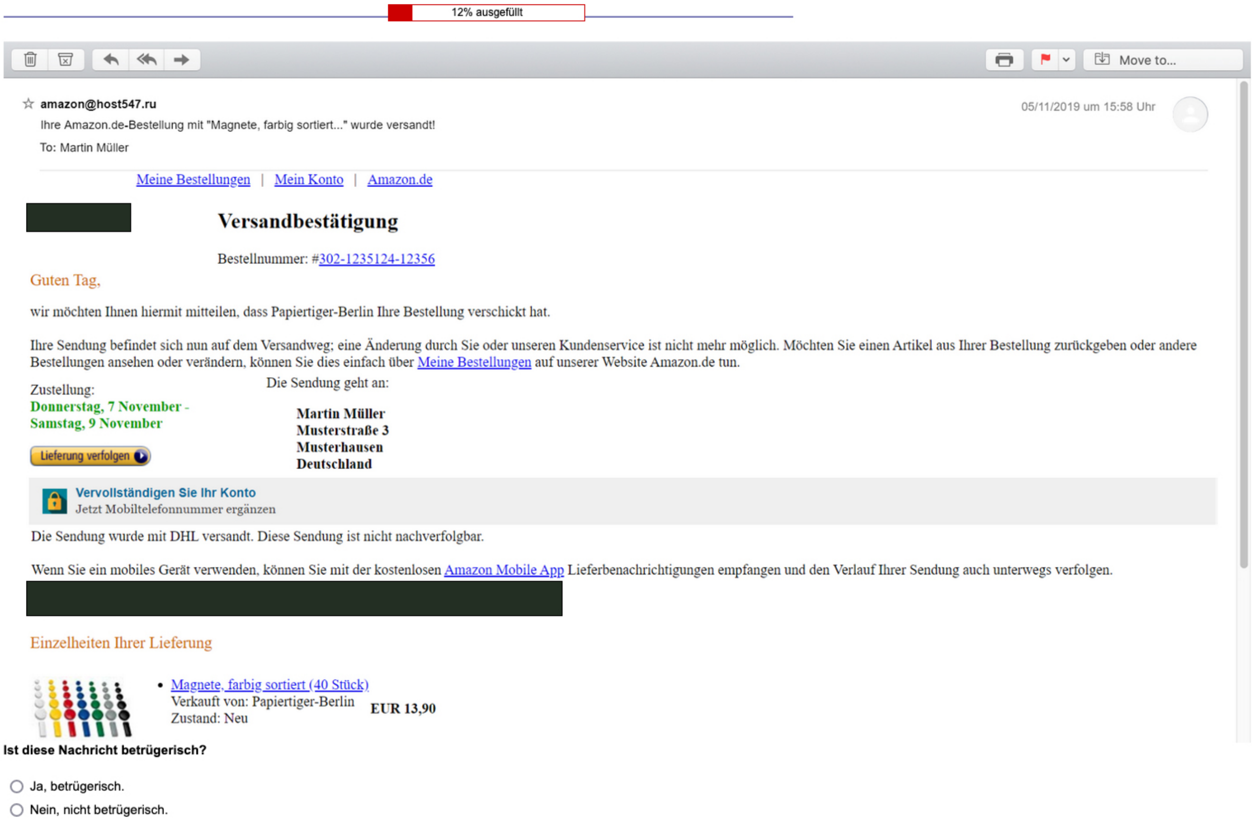
2) Self-generated code: Lieber Teilnehmer/innen,
da wir Ihren Fragebogen anonym zuordnen wollen, ist es wichtig, dass Sie sich Ihren persönlichen Code generieren.
Denn nur so können Ihre Fragebögen einander zugeordnet werden, ohne dass jemand herausfinden kann, wer diese Fragebögen ausgefüllt hat.
Wichtig ist also, dass Sie denselben Code noch wissen, wenn Sie beim nächsten Mal gefragt werden.
Aus diesem Grund haben wir die nachfolgenden Fragen formuliert, die Ihnen helfen sollen, sich an Ihre persönliche Kombination zu erinnern.
Bitte nennen Sie den ersten und letzten Buchstaben des Vornamens Ihrer Mutter (z.B. Anne = AE)
Bitte nennen Sie den ersten und letzten Buchstaben des Vornamens Ihres Vaters (z.B. Thorsten = TN)
Bitte nennen Sie den ersten und letzten Buchstaben Ihres Vornamens (z.B. Hannah = HH)
Bitte nennen Sie den Geburtstag Ihrer Mutter (z.B. 17. Juli 1950 = 17)

3) Scenario description: Um festzustellen, wie gut Sie Nachrichten mit gefährlichem Inhalt von echten Nachrichten unterscheiden können, sind im Folgenden Beispiele mit Nachrichten in unterschiedlichen Kontexten dargestellt. Wohl wissend, dass Sie nicht alle Absender, Diensteanbieter, Betriebssysteme und Programme kennen bzw. nutzen, stellen wir empfangene Nachrichten für verschiedenste Absender, Diensteanbieter, Betriebssysteme und Programme dar.

Um die Absender, Diensteanbieter, Betriebssysteme und Programme, zu denen Sie in der Realität keinen Bezug haben, nicht direkt für unplausibel zu erklären, gehen Sie im Folgenden bitte davon aus, dass...

- sie Martin Müller sind.
- Sie die Sprachen Deutsch und Englisch sprechen.
- Ihr Arbeitskollege Jonas Schmidt ist.
- Sie alle Dienste nutzen, die in diesem Fragebogen verwendet werden.
- Sie die verschiedenen Betriebssysteme (z.B. Microsoft Windows, Apple OSX, Google Android, Apple iOS) und Programme (Thunderbird, Apple Mail, Google Mobile Mail, GMX Mobil Client) nutzen, die in diesem Fragebogen verwendet werden.

4) Screenshots: Example of an email screenshot as seen by the participants. Every page shared the same structure, only changing the screenshot shown. The logos are hidden for publication, but they were present during the evaluations.



12% ausgefüllt

☆ amazon@host547.ru
Ihre Amazon.de-Bestellung mit "Magnete, farbig sortiert..." wurde versandt!
To: Martin Müller

05/11/2019 um 15:58 Uhr

Meine Bestellungen | Mein Konto | Amazon.de

[redacted]

Versandbestätigung

Bestellnummer: #302-1235124-12356

Guten Tag,

wir möchten Ihnen hiermit mitteilen, dass Papiertiger-Berlin Ihre Bestellung verschickt hat.

Ihre Sendung befindet sich nun auf dem Versandweg; eine Änderung durch Sie oder unseren Kundenservice ist nicht mehr möglich. Möchten Sie einen Artikel aus Ihrer Bestellung zurückgeben oder andere Bestellungen ansehen oder verändern, können Sie dies einfach über Meine Bestellungen auf unserer Website Amazon.de tun.

Zustellung:
Donnerstag, 7 November - Samstag, 9 November

Lieferung verfolgen

Die Sendung geht an:

Martin Müller
Musterstraße 3
Musterhausen
Deutschland

Vervollständigen Sie Ihr Konto
Jetzt Mobiltelefonnummer ergänzen

Die Sendung wurde mit DHL versandt. Diese Sendung ist nicht nachverfolgbar.

Wenn Sie ein mobiles Gerät verwenden, können Sie mit der kostenlosen Amazon Mobile App Lieferbenachrichtigungen empfangen und den Verlauf Ihrer Sendung auch unterwegs verfolgen.

[redacted]

Einzelheiten Ihrer Lieferung

- Magnete, farbig sortiert (40 Stück)
  Verkauft von: Papiertiger-Berlin
  Zustand: Neu

EUR 13,90

Ist diese Nachricht betrügerisch?

○ Ja, betrügerisch.
○ Nein, nicht betrügerisch.

5) Feedback & Socio-Demographics:

- Haben Sie allgemeines Feedback zur Schulung?
- Geschlecht: Männlich/Weiblich/Sonstiges
- Alter: Ich bin im Jahr geboren.
- Beruf / Studiengang:

B. Email examples

Sender	Subject	Phishing Type	URL
amazon@host547.ru	Your Amazon.de order with "Magnets, assorted colors..." has been shipped	wrong e-mail address	
Amazon (versandbestaetigung@amazon.de)	Your Amazon.de order with "Magnets, assorted colors..." has been shipped		
Vodafone Team(nichtantworten@kundenservice.vodafone.com)	Confirmation of your cancellation	Social engineering, implausible content	
Vodafone Team(nichtantworten@kundenservice.vodafone.com)	Confirmation of your cancellation		
Lufthansa(newsletter@lufthansa.com)	As of March 2019, we are introducing a new system for awarding bonus miles	Missmatch + Random URL	https://www.hisoliajo.host547.com/web3/HoEv/ksokGkd=ad3/kol45G
Lufthansa(newsletter@lufthansa.com)	From March 2019, we will introduce a new system for awarding bonus miles		https://www.lufthansa.com/de/miles-and-more/meilenvergabe
Lufthansa(newsletter@lufthansa.de)	Starting in March 2019, we will introduce a new system for awarding bonus miles	Random URL + Missmatch	https://www.dtrdtcbj.com/de/en#blade
Paypal (paypal@mail.paypal.de)	PayPal account overview for August		https://www.paypal.de/webapps/mpp/aq?utdled-notificamai&s=ci&mail=sys
DHL Paket (info@dhl.de)	New password for dhl.de	Random URL + Missmatch	https://www.host547.com/verify/ßkey=kw2RtU_5dsh
DHL Paket (info@dhl.de)	New password for dhl.de		https://dhl.de/account
Paypal (paypal@mail.paypal.de)	PayPal account overview for August	Random URL + Missmatch	https://www.hisoliajo.host547.com/web3/HoEv/596ksokFkd89=ad3/kol45G5Hwerg?32
Paypal (paypal@mail.paypal.de)	PayPal account overview for August		https://www.paypal.de/konto-uebersicht.php
Paypal (paypal@mail.paypal.de)	PayPal account overview for August	Random URL + Missmatch + Fake Tooltip	https://www.hisoliajio.host547.com/web3/HoEv/596ksokFkd89=ad3/kol45G5Hwerg?32
Paypal (paypal@mail.paypal.de)	PayPal account overview for August		https://www.paypal.de/konto-uerbersicht
Amazon (versandbestaetigung@amazon.de)	Your Amazon.de order with "Magnets, assorted colors..." has been shipped	Deception area after the who area	https://amazon.de.kolwerg.com/596ksokGkd89=adweb3/HoEv
Amazon (versandbestaetigung@amazon.de)	Your Amazon.de order with "Magnets, assorted colors..." has been shipped		https://amazon.de/596ksokGkd89=adweb3/HoEv
Amazon (versandbestaetigung@amazon.de)	Your Amazon.de order with "Magnets, assorted colors..." has been shipped	Deception area after the who area	https://amazon.de.kolwerg.com/596ksokGkd89=adweb3/HoEv
Amazon (versandbestaetigung@amazon.de)	Your Amazon.de order with "Magnets, assorted colors..." has been shipped		https://amazon.de/596ksokGkd89=adweb3/HoEv
Paypal (paypal@mail.paypal.de)	PayPal account overview for August	Random URL + Missmatch + Fake Toolip before the Who area	https://www.qpglljhjotqgg.com/paypal.de
Lufthansa (newsletter@lufthansa.com)	As of March 2019, we are introducing a new system for awarding bonus miles		https://www.lufthansa.com/meinlufthansa/service
Vodafone Team (nichtantworten@kundenservice.vodafone.com)	Confirmation of your cancellation	Random URL + Missmatch + Deception attempt before Who area	https://www.qpglljhotg.com/vodafone.com.596ksokGkd89=adweb3/HoEv
Vodafone Team (nichtantworten@kundenservice.vodafone.com)	Confirmation of your cancellation		https://www.vodafone.com/meinvodafone/services/Ihre-rechnungen
Lufthansa (newsletter@lufthansa.com)	Starting March 2019, we will introduce a new system for awarding bonus miles	Missmatch + Spelling error	https://www.luftthansa.com/de/newsletter-information
Lufthansa (newsletter@lufthansa.com)	Starting March 2019, we will introduce a new system for awarding bonus miles		https://www.lufthansa.com/de/en/newsletter-information
Vodafone Team (nichtantworten@kundenservice.vodafone.com)	Confirmation of your cancellation	Missmatch + spelling error	https://vodafon.de/fwlink/?LinkID=462932
DHL Paket (info@dhl.de)	New password for dhl.de		https://dhl.de/go/12/3ETWV5HW-2PSB2P6G-2RFZRFKW-17YA1VS-o.html
Vodafone Team (nichtantworten@kundenservice.vodafone.com)	Confirmation of your cancellation	Exe. File in the attachment (Directly executable at the attachment)	
Vodafone Team (nichtantworten@kundenservice.vodafone.com)	Confirmation of your cancellation		
DHL Paket (info@dhl.de)	Convenient parcel receipt with the DHL parcel box	Exe - file in the attachment, which is to be covered up by the "Pdf"	
DHL Paket (info@dhl.de)	Convenient parcel receipt with the DHL parcel box		
Jonas Schmidt	Notes	Random URL + Missmatch	https://husjukuila-torgibut/com/join
Jonas Schmidt	Gifts		https://www.amazon.de/dp/B00Li0KY52/ref=sr_1_4?keywords=sohn+kriminalroman&qid=1581586169&sr=8
Jonas Schmidt	Job offer farmers market	misspelling + missmatch	https://www.baurenmarkt.de/job-angebot/1eqds321h1w34UHA
Jonas Schmidt	Video KickOff		https://www.youtube.com/watch?VGhijvA4XTU&list=RDA44i5rzysu

TABLE III: Sender, subject, phishing type and URLs used for the examples of the study.