The Influence of Context on Response to Spear-Phishing Attacks: an In-Situ Deception Study

Verena Distler
verena.distler@uni.lu
University of Luxembourg
Esch-sur-Alzette, Luxembourg

ABSTRACT

In today's digitized societies, phishing attacks are a security threat with damaging consequences. Organizations remain vulnerable to phishing attacks, and it is not clear how the work context influences people's perceptions and behaviors related to phishing attempts. I investigate (1) how contextual factors influence reactions to a spear-phishing attempt, (2) why people report or do not report phishing attempts, (3) which opportunities for security-enhancing interventions people identify. I use an in-situ deception methodology to observe participants (N=14) in their realistic work environment. I triangulate observational and self-reported data to obtain rich qualitative insights into participants' emotions, thoughts, and actions when receiving a targeted phishing email. I find that task, IT, internal and social context play an important role. The email's request being aligned with expectations and perceived time pressure when responding to emails were associated with insecure behavior. The social context positively influenced phishing detection, but "phished" participants did not tell anyone.

CCS CONCEPTS

• **Security and privacy → Human and societal aspects of security and privacy**; • **Human-centered computing → Empirical studies in HCI**.

KEYWORDS

Usable privacy and security, Human-computer interaction, Phishing, Empirical research, Qualitative research methods

ACM Reference Format:
Verena Distler. 2023. The Influence of Context on Response to Spear-Phishing Attacks: an In-Situ Deception Study. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23), April 23–28, 2023, Hamburg, Germany.* ACM, New York, NY, USA, 18 pages. https://doi.org/10.1145/3544548.3581170

1 INTRODUCTION

Many modern societies rely on digital infrastructure to function. Security attacks on this infrastructure can cause financial, societal and physical harm. Phishing attacks are a particularly compelling threat. They are the most commonly reported cyberattack [39], and accounted for 90% of data breaches in 2020 [8]. Phishing attacks are frequently the point of entry for far-reaching attacks that lead to personal and organizational harm. In the past, phishing attacks have been the root cause of cyberattacks on companies (e.g., Sony in 2014 [33]) as well as critical infrastructure such as power grids (e.g., in Ukraine in 2015 [33]). In 2021, members of the German Bundestag were also the victim of phishing attacks aiming to obtain lawmakers' login details [38]. A particularly damaging example of the problems caused by phishing are ransomware attacks (e.g., [35]). It is therefore both urgent and important to address phishing attacks.

Phishing is inherently a socio-technical problem, and solutions need to address both the human and the technical side, as technology on its own cannot solve organizations' vulnerability to phishing attacks. Technical solutions are mostly based on two general approaches: blocklists of known phishing URLS and taking down known phishing landing pages [51]. Other options include using heuristics to check whether a website has certain characteristics that might be linked to phishing, and page similarity detection, which checks whether a website is similar to a legitimate website [54]. These latter options can be relatively inaccurate and are used more rarely [54]. Unfortunately, most technical solutions work best when a large number of similar emails are sent to many potential victims. They are not helpful for the first person receiving the phishing message. Most importantly, these approaches do not detect customized, targeted phishing attacks (spear-phishing) that are sent to only a small number of individuals [51]. Technical approaches often do not succeed in filtering out the attack before it is seen by the potential victim.

In this paper, I focus on human responses to spear-phishing attacks. I use the term "response" to refer to both subjectively experienced and behavioral responses to phishing. I avoid the frequently used term "susceptibility to phishing", as it can lead to the impression that being (un)susceptible to phishing is a trait-like participant-level characteristic, rather than context-dependent behavior.

This study addresses three research questions.

- RQ1: What role do contextual factors play in successful phishing attacks?
- RQ2: How do people rationalize reporting (or not reporting) a phishing attack?
- RQ3: What opportunities for security-enhancing interventions do people identify after being exposed to an attack?

There are various valid views on what defines a "successful" phishing attack, which include clicking on a link, downloading an attachment, or communicating sensitive information to the attacker on a website or via other means. For the present study, I define a

This work is licensed under a Creative Commons Attribution International 4.0 License.

CHI '23, April 23–28, 2023, Hamburg, Germany
© 2023 Copyright held by the owner/author(s).
ACM ISBN 978-1-4503-9421-5/23/04.
https://doi.org/10.1145/3544548.3581170

successful phishing attack as one where the victim clicks on the link in the email, an action which can compromise the security of the victim and their organization. In terms of context, I focus on internal context, social context and task context. The term "reporting" in RQ2 refers to communicating a phishing attempt to the IT department. RQ3 builds on the idea that, immediately after being exposed to a spear-phishing attack, non-experts might have novel insights into how to help others in the same situation avoid falling for a phishing attempt.

To address these objectives, I observed 14 participants in their realistic work environment (in-situ methodology) and injected a simulated spear-phishing attack. Combined with qualitative interviews, this approach allowed me to gain in-depth insights into participants' thought processes, emotions and actions upon receiving a phishing email.

This study makes the following contributions.

- This study contributes to an improved understanding of user response (experienced and behavioral) to a spear-phishing attack, including reporting behavior. The study uses an in-situ methodology in a real-life in-office work context. This maximizes ecological validity and provides insights into how contextual factors influence responses to spear-phishing attacks.
- I derive recommendations to improve organizations' resistance to phishing and recovery after an employee has interacted with a phishing email.

I first introduce the relevant background for this research (section 2). I describe the methodology (section 3), the results (section 4), and discuss their meaning, limitations and practical implications (section 5), before concluding (section 6).

2 BACKGROUND

2.1 Challenges Related to the Study of Human Responses to Phishing Attacks

Phishing is a semantic attack that cons people into divulging sensitive information [13]. I adopt Wash's definition of phishing [51]: a message (email) that pretends to be something it is not, in order to get the user to do something they would not normally be willing to do. Spear-phishing is a more targeted version of phishing, which often addresses the victim by name [34]. Spear-phishing attacks often use "weapons of influence" (e.g., authority, commitment, scarcity, social proof) to encourage victims to take the intended action [15, 40]. Authority and urgency cues are associated with an increased likelihood of clicking on the link in the email [53].

People's reactions to phishing are a complex issue affected by both individual and context-related factors [53]. It is challenging to observe and study human reactions to spear-phishing, as such attacks are relatively rare and unpredictable. Researchers need to consider trade-offs between the realism of the exposure to risk, the practical feasibility, and ethical and legal concerns (a frequent challenge in usable privacy and security research [9]).

Role-playing approaches are sometimes used to study responses to phishing attacks [12, 13, 57]. Downs et al. [13] studied how non-expert users make decisions when confronted with suspicious emails. They identified three strategies used by their participants (this email appears to be for me, it is normal to hear from companies you do business with, reputable companies send emails). These strategies were not particularly helpful in recognizing phishing attempts in the role-playing exercise. The methodology used in this study (role-play in a controlled environment) has the limitation that it did not yield insight into how relevant these strategies would be in people's everyday context, using their usual inboxes. A later study [12] investigated how and why people fall for phishing, also using a role-play approach. The authors asked participants to play the role of an employee working for a company. Participants were asked how they would treat a number of emails. The results showed that a better understanding of the web (e.g., understanding URL and SSL/TLS indicators) helped participants recognize phishing attacks.

Other studies have used retrospective, qualitative methods to understand people's thought processes in the context of phishing. For instance, Wash [51] interviewed IT experts about instances where they successfully identified emails as phishing messages. The authors found that IT experts follow a three-stage process for identifying phishing messages: (1) make sense of email, (2) become suspicious, (3) deal with email, mostly by deleting the phishing email, with some of them reporting it. Wash et al. [52] surveyed non-experts and asked them to think about a time when they had received a "suspicious or potentially harmful email" and to recall features of the email that caused suspicion. The authors concluded that non-experts' strategies for detecting phishing emails are similar to experts' approaches.

Studies can also use click rates (i.e. whether a person clicked on a (simulated) phishing link) to investigate human responses to phishing [19]. Another approach is to recruit participants in-person and install a browser plugin that displays phishing-related warnings. After three weeks, participants were sent a simulated phishing email, and the researchers recorded whether participants provided relevant information on the simulated phishing website [56].

Research has investigated the possibility of "predicting" a person's likeliness of interacting with a phishing email by, for example, investigating how much variance in click behavior is predicted by respondent characteristics. Individual factors do not seem to be the determining factor in response to phishing, however. Previous work found some effects of age and gender [29, 46], but researchers suspected that these differences might be due to different technical training and knowledge, and lower risk aversion in young people [46]. Greitzer et al. [19] found that age or gender are not useful predictors, and that measurement data (e.g., firewall data, VPN data) was not useful to identify individuals who are more susceptible to phishing. This aligns with recent results from a representative UK sample, where *no* demographic characteristics, personality traits or privacy concerns reliably predicted phishing detection abilities [57]. The authors highlight the need for novel approaches to help users evaluate email authenticity, calling for interdisciplinary collaboration between software developers and social scientists. They also point to the current lack of knowledge of how context-related factors influence reactions to phishing [57].

Overall, studying human responses to phishing attacks is non-trivial. Recalling phishing attacks that lie in the past may be challenging for participants. When studies take into account real-life

behaviors, they typically rely on click rates, and often do not include self-reported data collected directly after the attack, and are thus unable to provide insights into thought processes when receiving (and perhaps falling for) a phishing attempt. Training approaches attempt to educate users to help them avoid taking any insecure actions when exposed to a spear-phishing attack.

2.2 Training Users to Avoid Interacting With Phishing Emails and Encourage Reporting

User-centered anti-phishing interventions include education, (awareness) training, and design (e.g., visual elements, redirecting a user's course of action) [16]. Others categorize defensive strategies into the dimensions of attitude and behavior change [45].

Training approaches can teach users to avoid interacting with phishing emails, but have shown several limitations. Individuals should avoid dangerous behaviors such as clicking on the link in the phishing email or downloading a malicious attachment. For example, Canova et al. [5] present a training to teach users how to check URLs for legitimacy, with varying levels of difficulty. Kumaraguru et al. [29] developed an embedded training system that delivers a training message when a user clicks on a simulated phishing email, finding that embedded training is superior to sending security notices. They later studied retention over time of training messages, showing that users retained knowledge even after 28 days, and that a second training message to reinforce the original training decreases the likelihood of people giving information to phishing websites [29].

These types of simulated, embedded phishing trainings have been criticized for potentially lowering organizational security [44, 50]. These authors argue that employees who are aware of the simulated phishing campaign might perceive an incentive to interact with the simulated phishing message (e.g., clicking on the link, downloading the attachment) out of curiosity, to learn more about the topic, or to boycott the training campaign. These campaigns might also be counterproductive to increasing the reporting of phishing messages, since employees might assume that a suspected phish is simulated, and thus already known to the IT department [50]. Phishing simulations might help users detect certain, especially non-sophisticated phishing attacks, but even experts struggle to detect sophisticated attacks. As responding to emails and clicking on links is a fundamental aspect of most working environments, it is unrealistic to expect users to remain vigilant all the time [7]. The UK Centre for the Protection of National Infrastructure encourages organisations to communicate to employees that it is ok to ask for help with phishing emails and to instil a culture that does not blame or punish those who have fallen for a phishing attack, instead encouraging them to report any mistakes they may have made. The use of click-rate metrics (measuring how many employees click on a phishing link) has also been criticized, highlighting how the design of the email and other factors can influence the click-rate metrics arbitrarily, and how such training is often related to a blame-based security culture [6]. Simulated phishing campaigns have led to public outcry when the simulation was perceived as unacceptable, and have negatively influenced employee perception of their employer [2].

In addition to phishing detection, many anti-phishing training approaches encourage employees to report suspicious emails to their IT department. Reporting phishing emails allows the organization to address an ongoing attack, warn employees and put other countermeasures into place [26]. Unfortunately, users seldom report phishing emails [30]. The likelihood of reporting a phishing attack is increased by self-efficacy (i.e. a person's confidence in performing a behavior [48]) , expected negative outcomes, and cyber security self-monitoring [30]. Authors have also investigated the use of gamification approaches to increase phishing reporting rates [22]. Previous work has argued that reporting "false positives" (reporting legitimate emails) could lead to an overload of reported emails [26]. Authors argue that there is a trade-off between encouraging employees to report more suspicious emails, while avoiding a high number of false positives [22].

2.3 The Influence of Context on Phishing Reactions

In phishing-related research, recent work has also pointed to the need to further understand attention and situational changes and their influence on reactions to phishing [53, 57]. A recent study asked participants to classify potential phishing emails, and found that increased time pressure lowered participants' ability to detect phishing emails [24]. As described in the theory of situated action, actions are necessarily influenced by their material and social circumstances [47]. Indeed, how a user experiences a situation is strongly influenced by context, and the majority of UX models include context as one of the main factors impacting UX [32]. Bradley and Dunlop [3, p. 424] define context as "anything that influences the process in which focal user actions are undertaken". Most models agree on the following important dimensions of context [3, 32].

- *physical context*: e.g., noise, temperature, familiarity of the location [32].
- *social context*: refers to the influence of the people surrounding the task (in-person or remote). Includes feelings of relatedness and social support [32].
- *internal context*: situational profile of a user's identity, e.g., mood, motivation, interest in the system, previous experiences [32].
- *technical context*: e.g., technical issues.
- *task context*: e.g., type of task, interruptions, competing tasks.
- *temporal context*: e.g., duration of the interaction, time of day.

I use these dimensions of context as a guide when investigating the role of context in the response to phishing. I adopt the term "IT context" to capture the technologies used by participants as well as any technical issues. To understand context, the researcher needs to immerse themselves into the relevant context [20]. A variety of methods encourage researchers to embrace the contextual nature of action (e.g., contextual inquiry [43, 55], contextual design [21] or practical ethnography [20, 31]).

The influence of context on responses to phishing is not well explored yet. As described in section 2.1, much of our understanding of how people recognize phishing attacks is based on click rates; self-reported information about a past phishing attack, which can

be difficult to recall accurately; or on role-play scenarios, which can lack realism. When receiving a phishing email, the context a person is currently in seems to be a highly relevant factor to investigate. For instance, feeling tired or stressed in a situation (internal context) could likely influence the perceptual and behavioral response to an ambiguous situation. Similarly, responding to emails on a mobile device vs. on a computer (technical context) can influence a user's response.

In summary, the influence of context on phishing-related behaviors is underexplored. This paper investigates the influence of participants' context on their response to phishing attempts and potential reporting behaviors. By analyzing the interplay between context and participants' responses, I derive practical improvements to counteract phishing attempts.

3 METHODOLOGY

3.1 Research Design Overview

This study used an in-situ deception methodology that allowed me to observe participants in their usual work environment (figure 1). While participants were aware of the observation, they were not aware that the study was about phishing (deception). During the observation, participants received a simulated spear-phishing email, providing a unique opportunity to observe a response to a phishing attack in a realistic context. After the observation, I interviewed the participants and asked them to complete a retrospective experience exercise (UX curve) enabling them to reflect on their experience. I triangulate these data sources with a summarizing content analysis.

3.2 Research Setting, Recruitment and Participants

3.2.1 Broader research setting. I set out to investigate phishing in an organizational setting. This study took place in a European university. The participating university used the tool "KnowBe4" to regularly simulate phishing attacks for training purposes, which all employees automatically take part in. This means that employees regularly receive simulated phishing attacks at their professional e-mail address. If an employee clicks on the link in the simulated phish, they are redirected to a website that seeks to inform them about red flags concerning phishing. I worked directly with the university's chief information security officer (CISO) to send the simulated phishing attempt.

3.2.2 Obtaining Managers' Agreement. Professors, in their role as group leaders, were asked for their agreement to (1) use the professor's name in the email to simulate a spear-phishing email and (2) allow interested staff members to participate during work hours. These professors were not associated with user-centered security and privacy or with human-computer interaction. The study objective and procedure were explained to them, including the deception. All five contacted professors agreed. I sent the professors a recruitment flyer (appendix A) to forward within their research group. The study invitation thereby reached approximately 70 potential participants, who could then volunteer to participate.

3.2.3 Recruiting Participants. Employees indicated their interest to participate by filling out a short form with their availability and e-mail address. It was challenging to recruit a diverse set of participants for this study. I discuss potential reasons in the limitations in section 5.5. I made an effort to include administrative staff members, who might have different strategies and views with respect to emails, and specifically asked them to participate. To include an "extreme" case in terms of technical skills, I also included one IT expert. After sending out multiple invitations over the span of six weeks, focusing on "under-represented" research or job groups, 14 participants agreed to participate. Participants were compensated with 60€ in gift vouchers (ca. two hours of participation). I recruited 14 participants (3 men, 11 women, 0 non-binary; mean age 35 years; SD: 8 years). The sample included one IT staff member, three administrative staff members, and ten research staff members. Participants had between 0 to 27 years of work experience (mean: 7 years, SD: 7 years). Apart from the IT expect, participants did not have a computing/security background. I did not exclude any participants from the study.

3.3 Material

3.3.1 Technical Equipment. I used a webcam (Logitech C925e) on a tripod to observe participants. The webcam was connected to a tablet from which I launched a video call. This setup was discrete and did not impede participants' ability to get up and move (figure 2). The setup allowed me to see what participants were doing on their computer during the observation, and to observe their reaction to the phishing email.

3.3.2 Phishing Email. The study objective was to observe a targeted phishing email sent by a motivated attacker. The email included an urgency cue and an authority cue, which have been found to be particularly likely to lead to a target clicking on a phishing link [53]. The email seemingly came from either a direct supervisor or a "principal investigator" (authority cue) in the case of administrative staff. While five professors agreed to being impersonated in the simulated phishing email, no staff members supervised by two of these professors volunteered to take part, despite multiple invites. Thus, the participants received authority cues from three professors. The phishing email is presented in figure 3. Note the similarity between the recipient's email address (a legitimate, internal address) and the sender address (the attacker's attempt to replicate the internal address). The latter included an additional country indicator ("de" in the figure) before the dot. As is the case for many research institutions and other organizations, the names of staff members and their supervisors are publicly available online and thus easily accessible for a motivated attacker. The link in the phishing email led to the default training website used by the university (appendix B), which presents some "red flags" to recognize phishing. I did not change the default website used, as I intended to test the simulation of a spear-phishing email as it might realistically happen in an organizational context.

3.4 Procedure

I iteratively improved the protocol via multiple pre-tests and feedback from HCI experts. Inspired by approaches from applied ethnography [31], which are considered the "gold standard" when investigating context [20], I initially planned to observe participants in person. However, I found that the offices left little space for an

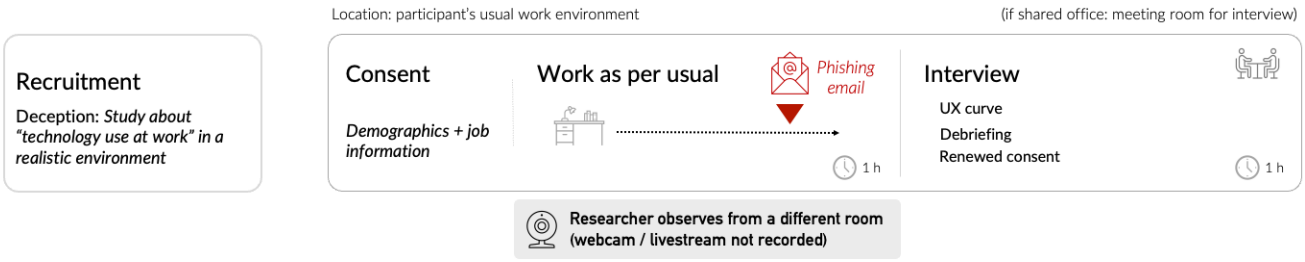

Figure 1: Summary of the methodology used in this study.

Figure 2: The study setup in a typical office. Note the webcam on a tripod and the tablet. The person shown is not a participant.

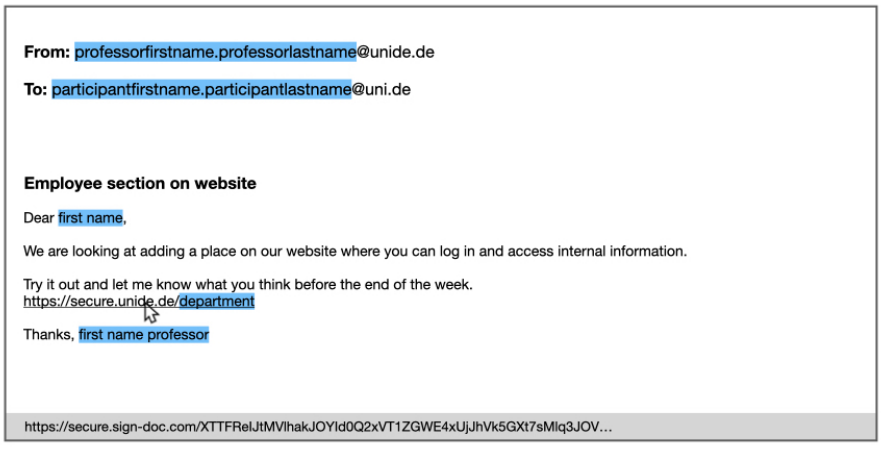

Figure 3: The simulated phishing email. Highlighted text was personalized to the participant.

observer and that in-person observation was perceived as too intrusive in the work environment. Thus, I switched to a remote setup. I provide the full study protocol in appendix C.

Consent (setup and interaction). At the agreed-upon time, the researcher came to the participant's office and explained the basic setup of the study. The study was introduced as intending to

"understand how people use technology at work". Participants provided informed consent, and any questions were answered. The participant filled out a short questionnaire with basic demographic and job information while the researcher set up the webcam in their office. The researcher left the office once the setup was complete, telling the participant that they would come back in approximately one hour for the interview. The participant was not assigned a task and was instructed to work as per usual. Participants were told that they did not have to change their behavior (incl. social interactions) and that they were not being evaluated on how well they worked. A typical office and the study setup are shown in figure 2.

Observation (work as usual). The participants were instructed to work as usual for one hour. They received no instructions about their email software (open/close, notifications on/off). After 55 minutes, they received the spear-phishing email. Five minutes after the email, the researcher came back for the interview. I selected a one hour time frame because it allowed participants to sufficiently immerse themselves in their work (based on pre-tests that lasted only 45 minutes).

The observations were structured along the following dimensions of context.

- IT context (e.g., multiple screens, operating system, email program(s) used, devices used)
- Task context (e.g., interruptions, task switching, how does participant react to e-mails, e-mail notifications on/off)
- Social context (e.g., other people present in office, does participant talk to colleagues, does the participant receive/make phone calls, do people drop in)
- Temporal context (observations related to the timing of the observation)
- Physical context (e.g., workspace, temperature,...)
- Internal context (e.g., visible expressions of stress, but confirmed with participant during interview to avoid misinterpretation)
- Other observations

In addition, I observed and took note of what participants were doing when the phishing email arrived (incl. what they were doing on their computer screen), how they reacted, and whether they reported the phishing email. Note that while the observation structure included temporal and physical context for completeness and to allow me to discover factors related to these dimensions, I did not expect variations in physical context as all participants took part in the study from their offices. Although all participants' took the same amount of time to complete the study, I took note of temporal context, which might vary for participants taking part at different times of day.

For the purpose of this study, I defined a successful phishing attack as the victim clicking on the email link, which can compromise the indvidual's and the organization's security. I operationalised this measure by observing participant behavior.

Interview. I conducted semi-structured interviews in the form of a confidential conversation (in individual offices or reserved conference rooms). I followed the participant's train of thought as much as possible, and adapted the question order to their narrative. I first interviewed the participant about their general work environment and contextual factors. I then asked about typical interruptions during their work day (incl. emails), with follow-up prompts to understand how participants dealt with these interruptions.

I then debriefed participants fully, explaining the deception to them. I asked whether participants (a) had suspected the hidden objective of their study participation, (b) had any questions or concerns, (c) were still willing to participate in the study (renewed consent). None of the participants had suspected the objective of the study, indicating that the deception was successful. Participants were given the option to withdraw from the study, and were handed an information sheet to complement the consent form. All participants were still willing to participate. I encouraged participants to voice any questions or concerns. I explained that recognizing phishing emails is difficult, even for experts, to encourage them to speak more freely about their experience with the phishing email in the study and their general phishing experiences.

Participants were asked to explain their thought process when they received the phishing attempt and their reaction. If participants did not mention reporting the phish themselves, I asked them about reporting, introducing it as "Some people inform the university when they receive phishing emails. Have you ever done this?"

Finally, I asked participants to draw a UX curve representing their experience during the observation. The UX curve is a retrospective technique used to help participants reflect on meaningful aspects of their experience [27]. It facilitated the interview by encouraging the study participants to reflect on their thoughts and emotions during the data collection period, enhancing data quality. By asking participants to draw the UX curve, I was able to encourage detailed reflections on their thoughts and feelings at each point of the study. I provided participants with a template for the UX curve (figure 7 in appendix D). After drawing a curve representing their overall experience, I asked them to add the phishing email (if they hadn't already). I asked them to imagine that somebody else was in the same situation they had been, and to imagine how we might help them avoid clicking on the link in the phishing email. I asked them to draw these potential solutions on the curve.

Protecting participants from harm. This study was approved by the university's ethical review board (ERB). Going beyond ERB approval, I took measures to protect participants from potential harm. I collected the minimal data necessary. Employees' participation, their behaviors, and their statements were confidential. Participants' names and email addresses were needed to coordinate their participation, but this personal information was not linked to any of the notes or recordings. I did not obtain access to any past or future click or reporting behavior by the employees, as this was not relevant for the study objectives.

3.5 Data Analysis and Methodological Integrity

All interviews were transcribed, and the UX curves were digitized. I combined the observation data, interview data and UX curves. I used MAXQDA 2022 and created a document group for each of these types of data, with one document per participant. I used Mayring's process for a "summarizing content analysis" [36, p. 72] and adapted it to the material, following guidelines by Rädiker & Kuckartz [42, p. 146]. I use the term "code" to refer to the categories formed in the analysis to be consistent with the terminology used

by MAXQDA. I first paraphrased the interviews. I then formed codes based on the material (inductive), and used a-priori concepts to categorize these codes (deductive). The coding unit was defined as a meaningful statement about a topic ("Sinneinheit"), and I used thematic coding [42, p. 34].

Reporting of qualitative results. In qualitative research, it is not always meaningful to report quantified measures, such as the number of occurrences of a code [42, p. 54]. The more "freely" an interview was conducted and adapted to the participant, the less sense it makes to report the number of occurrences. In the results section, I therefore do not provide the number of occurrences per code, and instead focus on representing the range of thoughts brought forward by participants. Direct quotes are reported with the participant number and the paragraph from the interview transcript.

Inter-coder agreement. One of the quality criteria in qualitative research is the extent to which the codes can be applied by different coders, which ensures that the code definitions are sufficiently clear and distinct [42]. To test the categories, I organized a co-coding session (in-person, duration 2 hours). Two members of the HCI research group took part and were asked to read the codebook and ask any questions, using these inputs to clarify the code definitions. I then presented each coder with one of the interviews, including both the initial transcript and the paraphrases. I asked the coders to apply the codes to the paraphrases, using the transcript for context when necessary. I responded to disagreement and questions by discussing and clarifying the codebook (details in appendix E). The primary goal of the co-coding session was to yield clearer concepts. The study and coding was conducted by a single "expert researcher" with unique expertise related to the topic, making inter-rater reliability metrics less meaningful [37].

Detailed description of analysis and codebook. I provide a more detailed description of the analysis procedure in appendix E. I also provide the codebook in appendix F. As the topics discussed and observed are, in part, sensitive, I am unable to publish any raw data. Table 1 summarizes the data sources used to answer each research question.

4 RESULTS

4.1 The Role of Contextual Factors in Phishing Attacks (RQ1)

4.1.1 Overview: Response to a Spear-Phishing Attack. The objective of this study is to generate hypotheses about thought processes and reactions to phishing attacks inductively. It is thus not appropriate to make any claims about general phishing click rates. All but one of the participants saw the email immediately (due to notifications being turned on). 4 out of 14 (29%) participants clicked on the link in the phishing email. Note that all participants' data was used to answer RQ1, not only those who clicked on the phishing link. I asked participants to explain their thought processes when receiving the email and combined these insights with the observations. I present the results of the qualitative analysis visually in figure 4. The thought processes described varied between participants who did not click on the phishing link (process 1 in figure 4), and those who did (process 2 in figure 4).

Thought process for participants who clicked on the link. Participants who clicked on the phishing link had a relatively straightforward thought process compared to those who did not. They saw the name of the professor, and either immediately clicked on the link to complete the request (P6, P12, P13), or were confused at first but constructed sense about the request. P9 was not sure about the website the supervisor was referring to in the email, but then imagined that it might be a website where they needed to create a profile for themselves, so they clicked on the link. P6, P12, and P13 reported having an "automatic" reaction, without deliberate thought going into their action. Triangulating these self-reported insights with the observation of their behavior, I could see that their reaction to the email was fast, almost immediate. P6 was coincidentally waiting for an email of this type regarding a website that was being set up for the department. For P6, the email immediately made sense: "When I received the phish, I only thought: the website is here. That was it." (P6, 90)

The authority cue installed a sense of urgency (P9, P12, P13). This was combined with the pervasive impression that emails should be answered (almost) immediately (see section 5.1.2).

P13 saw that the email came from her boss and wanted to fulfill the request immediately. "It said, can you check if you can log in? I thought, okay, well, easy enough. I'll just do that now. I'll reply and get on with it. And then the thing popped up. I was like, oh no." (P13, 75)

Thought process for participants who did not click on the link. Participants who did not click on the phishing link had more complex thought processes that prevented them from clicking on the link, and there were many things that made them suspicious.

The first suspicious elements were the *language* and *writing style*. The participants used language as an informal verification of the identity of the sender. The studied university is a multilingual environment, and the participants expected the alleged sender of the email to communicate with them in a certain language. In addition to language, participants noticed when the writing style was different from their usual communication with the professor, including the name they addressed them as (nickname vs. full name).

Participants frequently mentioned that the *request* did not make sense to them at the moment of receiving the email. For instance, they would have expected the professor to mention the website in advance of sending an email or they did not expect to be contacted about a website.

The *communication channel* of email was also unexpected for some participants. For instance, when both the participant and the professor were present in the building, they would have expected the professor to talk to them about the request in person.

Most participants only checked the *sender email address* once another element made them suspicious. More rarely, participants reported looking at the *URL* to check if the email is legitimate. Amongst these participants, there were still misunderstandings about the visible and invisible parts of the URL. P4 first hovered over the URL, and then copy-pasted the visible part of the URL, searching for it in a search engine to check what came up. P7 mentioned in the interview that he knows that you should "hover" over a URL, but did not know what this meant: "What I've never

Table 1: The qualitative data sources used to answer each research question.

Section	Research question	Data sources
4.1	Which role do contextual factors play in successful phishing attacks? (RQ1)	Qualitative analysis of observation notes and interviews
4.2	How do people rationalize reporting (or not reporting) a phishing attack? (RQ2)	Qualitative analysis of observation notes and interviews
4.3	Which opportunities for security-enhancing interventions do people identify after being exposed to a phishing attack? (RQ3)	Qualitative analysis of UX Curve and interviews

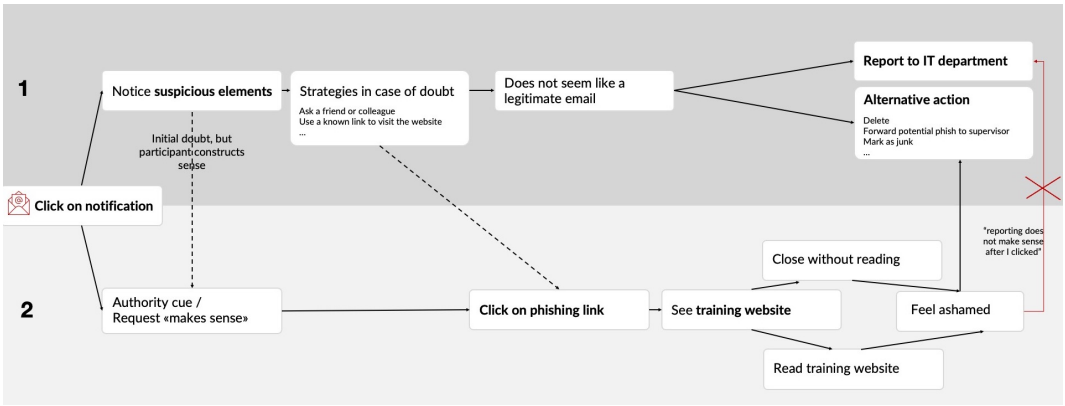

Figure 4: A visual representation of the qualitative results regarding the participants' thought processes. Process 1 is the process when participants did not click on the phishing link. Process 2 is the process that led to clicking on the link.

done before, you can somehow hover over it, I've never done that before. I don't really know how it works." (P7, 106-107).

Strategies in case of doubt: When participants were doubtful about an email, they used two main strategies: Ask a friend or colleague about it, or go to the website via a link they had previously used (and thus knew was legitimate).

Weighing the risks: Participants weighed the risks of how dangerous phishing emails really were if you only clicked on them (rather than download or type in something). Some were worried about not answering legitimate emails: "There is a fear of it being a real e-mail, and I classified it as a phish because it's from an unusual e-mail address or something" (P2, 81-82).

4.1.2 *Contextual Factors.*

Internal context. When asked about their feelings, some participants reported stress and being overwhelmed, while others reported low stress and positive feelings. To deal with distractions, some participants reported intentionally turning off notifications for requests (email and chat). However, the participants felt a strong pressure to respond to requests (emails and chats) as quickly as possible, and most dealt with emails by immediately checking and answering them. "I immediately answer emails, otherwise they stay in my head" (P7, 14). Regardless of the number of emails participants reported typically receiving in a work day, all participants mentioned feelings of being overwhelmed and stress related to emails.

Social context. I observed a number of social interactions, both for participants who were in an office alone as well as participants in shared offices. Besides in-person communication, participants used chat and phone calls. Participants reported both positive and negative aspects of in-person interactions in the office environment. Positive factors of in-person work included feeling less isolated and in a better mood than when working from home, and more direct communication in the form of asking a colleague for information when needed. Negative aspects of the office social environment were numerous in-person distractions and thus loss of focus and productivity. Note that participants did not mention changing their phishing-related behavior (e.g., checking URLs) due to the social pressure of being observed.

The observations also allowed me to see the positive influence of social interactions when a participant received a phishing attack. Social interactions were used by some participants to immediately inform direct colleagues and their supervisor of the phishing attack. One participant immediately called the professor: "[name of professor], come have a look! You are sending spam!" (P8, observation, 22).

Of note, *none* of the participants who clicked on the phishing link talked about it with anyone. I describe these results in detail in section 4.2.

IT Context. In the office context I observed them in, participants used computers to complete their tasks and answer emails.

Participants used both Windows and Mac computers in approximately equal proportion. While phones and other devices were often present at their desks, these were not used to deal with work emails. Participants used a variety of tools to communicate with colleagues, mostly email, chat, and more rarely, the phone. All but one of the participants received immediate notifications from these tools. The number of emails participants reported receiving per day ranged between 3 and 150 emails per day. Administrative staff were on the higher end of the scale, whereas junior research staff were on the lower end of the scale. Regardless of the number of emails received, participants reported attempting to answer all emails as quickly as possible. Participants reported keeping notifications mostly activated and their email program open during work hours, which is line with what I observed during the study.

Task context. The task context varied most strongly along job categories. While research staff saw emails as a distraction from their other tasks, administrative staff saw emails as a high-priority task. Participants were working on a variety of tasks during the observation, including responding to emails, administrative work, programming, academic writing, reading, and data analysis. I found no clear pattern of typical tasks participants who clicked on the link were working on compared to those who did not click on the link. Of the four participants who clicked on the phishing link, one participant was responding to emails, one was categorizing literature, one was preparing a meeting, and one was creating a purchase order when the phishing email arrived. I observed the same variety of tasks for participants who did not click on the phishing link.

Temporal and physical context. Participants worked on different tasks in the morning and afternoon. Some participants tried to structure their days in a way that would minimize distractions from emails (e.g., "I work well from 8 to 10 am and then after 6 pm, when no more emails come in. That's why I try to do my meetings in the middle." (P7, 42)). Within the setting of the study, I observed no clear patterns in temporal context that seemed relevant to phishing. Similarly, physical context was not found to be an influential factor in the study setup.

Summary. To summarize the results regarding RQ1, important contextual factors influencing reaction to phishing included the phishing email's request being aligned with participant expectations, or at least not raising suspicion. I did not identify a different pattern of tasks for participants who clicked on the phishing link compared to those who did not, but participants' views of emails differed between being seen as a distraction from actual work (more common among research staff) and being seen as an integral work task (administrative staff). The perceived pressure of needing to react to emails as quickly as possible (internal context) had a negative influence, as participants wanted to react to an email from their supervisor immediately. The social context could play a positive role if the person receiving the phish warned colleagues, but did not have a positive influence once the person clicked on the phishing link, as they did not communicate the phish in this case. IT context plays a role in that it can make recognizing a phish easier or hard (e.g., mobile device), and encourage/discourage reporting through a "report a phish" button.

4.2 Reporting a Phishing Attack or Taking Alternative Behaviors (RQ2)

In the sample, 7 (50%) participants reported the phishing attempt to the IT department. I combine this observation of behavior with the insights from the interviews for in-depth insights into participants' thoughts and actions when falling for a phish. I describe participants' thought process on when to report, obstacles to reporting, and reasons to report.

4.2.1 Falling for a Phish – Emotions, Thoughts and Actions (incl. Reporting). I wanted to understand how participants felt after clicking on the simulated phishing email, and how their emotions and thoughts might influence their subsequent actions. Participants who clicked on the phishing link reported feeling shocked, upset, ashamed, embarrassed, angry, and mad at themselves. They felt like they should have known better.

"Interviewer: What was your first reaction when you saw the anti-phishing training?" – "Embarrassment, I was like, wow, how could you do this? How could this have happened? [...] My first thought was, oh, my gosh has actually been something, has there been like a data breach or something? Have I done something that is serious?" (P13, 79).

Note that, with only one exception, none of the participants actually looked at or read the website that pops up after clicking on the phishing link. Instead, they immediately closed the website and even deleted the email: "I saw the website for the first time and was so annoyed that I closed it right away. I didn't even read it at the time. After that I thought maybe you should have read what was written there, but I didn't want to click the link again. I was so annoyed that I said – go away." (P12, 49)

(Interviewer asks what the respondent would do with the information on the website) "I always just close the website and continue my day." (P9, 129) "I never read to see if there is anything more to do, to be honest." (P9, 147).

None of the participants who clicked on the phishing link told anybody about the attack. As described in the next section, they also did not report the attack to the IT department.

4.2.2 Thought Process on When to Report. None of the participants who had clicked on the link reported the simulated phish. P13 explained: "I thought about [reporting] because on Outlook there's the little like report a phish button, but I thought [...] what's the point [...] because I already clicked on the link and then I feel like it would look a bit guilty if I then tried to press on the report a phish button like I didn't just click on the link." P13, 85.

Participants had varied impressions of when to report a phishing email. Some said that they would always report in case of doubt about an email. Others thought that they should report only when it was relevant for the IT department, or in cases where somebody asked for money or personal information.

4.2.3 Obstacles to Reporting. An important obstacle to reporting was not being sure how to do it. One participant had previously reported a suspected phishing attempt to an administrative contact, who said they were not responsible for it, and the participant gave up. One participant mentioned being "too lazy" to do it, but in the subsequent conversation, explained being (1) unsure of how to

report and (2) unsure of the reasons why reporting is helpful (P3, 168-174).

Other reasons were not wanting to get anyone into trouble for sending unprofessional emails (that might thus be misidentified as phishing attempts) and being worried about getting more work after reporting a phishing email.

4.2.4 Reasons to Report. Participants would report phishing attempts to improve the email filters. They would expect the IT department to block the sender address and/or investigate. Participants also mentioned that positive feedback encourages them to keep reporting.

"Because I like the congratulations. Yeah, I like the reinforcement. So, I just mark it to get the kudos." (P5, 112)

4.2.5 Alternatives to Reporting. A common strategy (other than reporting) was to delete the suspected phishing email or move it to the junk folder. Another option was to forward the suspected phishing email to the supposed sender and ask whether it was legitimate. Other alternatives were to ask the administrative contact what to do about it or post it on the research group's Microsoft Teams channel.

4.2.6 Social interactions and Reporting Behavior. I observed that social interactions had a positive influence on participants' behavior after receiving a phishing link.

P7 immediately told their colleague about the phishing attempt: "I just got a phishing attempt from [name of professor]. That was really well done. Do you think I should write something to someone? To our administrative contact?" The colleague answers: "Isn't there some IT address?" The participant responds: "Yes, I reported it." (P7, observation, 19-23).

P5 immediately asked colleagues in the same office whether they also got a phishing email. P5 then reported the email. When the participant received the "Congratulations, you spotted a phish" response, they called out "So exciting!". The colleagues congratulated the participant and applauded them for reporting the phish: "Woo!" (P5, observation, 19-23).

Summary. To summarize results regarding RQ2, reasons for reporting a phishing attempt included improving filters and getting positive feedback. Obstacles included being unsure of how to report and for which reasons, and fear of getting colleagues into trouble. Participants also felt that reporting did not make sense after they had already clicked on the link.

4.3 Opportunities for Security-Enhancing Interventions (RQ3)

In this section, I draw upon participants' thoughts from the interviews and their UX curve drawings. Initially, the intention was to investigate friction design ideas that could be implemented technically. However, many of the opportunities highlighted by the participants take a broader stance. Therefore, I call these ideas "security-enhancing interventions". In summary, participants suggested solutions regarding training (e.g., avoiding shame, specific training for spear-phishing, how to report), improving phishing awareness (e.g., about ongoing attacks, re-branding of phishing reporting), and technical approaches (warnings).

4.3.1 Training and Raising Awareness. Solutions brought forward by participants addressed the following points:

- Multiple participants suggested training interventions. They suggested encouraging people to ask for help when they are uncertain about an email, and that the *training should communicate empathy and appreciation.* P11 remembers clicking on a simulated phishing email in the past. The training message implied that they had not paid attention to previous training, but really, they recounted falling for the phish because they felt stressed and in a rush. This caused feelings of shame: "I didn't tell anyone, my husband or my colleagues or whoever, because I was so ashamed and felt so stupid. Taking a step back, it's not a good way to act. I didn't report anything." (P11, 90)
  Specific suggestions related to training were:
  - Teach employees to recognize external email addresses.
  - Teach employees to recognize malicious links (incl. how to hover).
  - Teach employees to pay attention to the writing style. This type of informal, non-technical indicator was frequently used by the participants.
- *Create awareness for advanced phishing*: Participants felt that the usual phishing simulation training used emails that could be easily recognized as phishing attempts. They felt less prepared to deal with more targeted phishing attempts.
- *How to answer emails:* Participants suggested including general guidelines about how to answer emails, such as avoiding clicking automatically, taking the time to read emails and not responding when distracted.
- Participants wanted to be *informed of currently ongoing phishing attacks* that were particularly good or targeted, or if there was a specific topic that was currently being used in a phishing attack.
- To encourage reporting, participants suggested that other employees be taught *how to report* and told about the *positive message* they would get after reporting.
- Participants suggested *lowering the barrier to getting support with suspicious emails.* One idea was to rebrand the phishing reporting email address (currently using a similar style as: report-a-phish@uni.gov), as this email address implies that you should be certain that you are reporting an actual phishing attempt. Another suggestion was to introduce an award for helping the community for people who reported phishing emails, to highlight personal gain from reporting a suspicious email, and poster ads with clear steps.

4.3.2 Technical Solutions. Participants suggested that the email program (or browser) should warn the user when the sender email address is a slight variation of an official, internal address. They also suggested warning the user if the URL in a link is particularly long.

5 DISCUSSION

I will discuss the results regarding participant response to phishing attacks (section 5.1), rationalizations for (not) reporting the email (section 5.2), and participants' identified opportunities (section 5.3).

I also provide recommendations (section 5.4) and discuss limitations (section 5.5).

5.1 Response to a Phishing Attack and Contextual Factors

5.1.1 Receiving and Evaluating the Phishing Email. I distinguished two types of thought processes. The thought process that led to clicking on the phishing link was more automatic, whereas I identified a more deliberate thought process that led to *not* clicking on the phishing link. These two types of thought processes are similar to what is represented by dual process theories, which see information processing as a mix of systematic and non-systematic thinking [23]. This has also been operationalised as system 1 (automatic, fast) and system two (reflective, slow) thinking. Note that these processes do not occur one after the other, and almost all processes are a mix of both systems [25]. In this study, more reflective consideration of the email was triggered by "suspicious elements" in the email (e.g., language, writing style, unexpected request). On the other hand, if the email aligned with participants' expectations (corroborating findings by [18]), or the authority cue made it so that they did not even read the email, this deliberate process was not triggered. Instead, participants reported reacting to the email in an automatic manner, immediately clicking on the link. These results confirm previous findings on participants' thought process when faced with a phishing attack that raises suspicion [51, 52]. An important contribution of the present study is that it provides novel insights into thought processes that happen when this thoughtful, deliberate process is not triggered.

It seems unlikely that training, on its own, can modify the way a person processes an email that does not raise initial suspicion, even if they have the required knowledge in principle. This would imply that training interventions have their limits, as recognized, for instance, by Yang et al. [56], who argue that warnings and user training must be combined for maximal effectiveness. Second, one might think that changing employees' habitual response to emails might be promising. However, encouraging employees to *always* double-check the sender address, link URLs, and attachments is not realistic, as *most* emails are *not* phishing attempts. Responding to emails and clicking on links is a fundamental aspect of most working environments, and it is unrealistic to expect users to remain vigilant all the time [7].

Previous work in usable privacy and security has made an argument for introducing friction to encourage more secure behaviors and avoid purely automatic behaviors in security-critical situations [10]. Friction could help create space for the recipient of an email to double-check the indicators they learned about in training. Such frictional design could take the shape of traditional warnings, delayed link activation (as suggested, for example in [49]), or attracting users' attention to the security-relevant parts of an email (similar to what is suggested in [4]). Once the person's attention is attracted, they can then be encouraged to evaluate the email, and suggestions like making URLs more interpretable to non-expert users can help them investigate the e-mail's legitimacy [1]. It is important to carefully design such security-enhancing friction in a way that understands and counteracts habituation effects.

Based on the findings on these thought processes, it is likely that well-designed spear-phishing attacks will trick a non-negligible number of employees within an organization. It is important that the consequences of falling for a phish are dealt with quickly and constructively. To meet this objective, reporting phishing is essential. I discuss challenges and opportunities related to phishing reporting in section 5.2. I will now discuss how context influenced the responses to phishing.

5.1.2 Contextual Factors and Response to Phishing Attacks.

Internal context. I found that across all job categories, participants felt pressure to respond to emails quickly, making them less likely to become suspicious of the email and investigate further. I hypothesize that this could be related to an organizational norm. Norms are expectations or unwritten "rules" of behavior in a social context [41]. Recent work has started to investigate how social norms can influence an organization's vulnerability to phishing [41]. In future work, it seems relevant to integrate research from social psychology [23] to investigate how changing certain norms (e.g., decreasing the perceived urgency of emails, including those from hierarchical superiors) could help improve organizational security. Another example of a norm change could be to switch the communication channel for internal requests from emails (which are easily imitated by outsiders) to secure messaging services.

Social context. I found that a participant's social environment can help them evaluate suspicious emails and get information about reporting procedures. Participants often mentioned they would ask a friend or co-worker for help in case of doubt, and I observed multiple phishing-related social interactions. Sometimes, the phish sparked a conversation about prior experiences with phishing emails. In future work, it would be relevant to investigate how these types of social interactions can be leveraged even further to make organizations less vulnerable to phishing. It seems promising to shift the perspective and see phishing as a group-level problem. For instance, phishing training could aim to spark phishing-related conversations rather than merely attempting to change individual behavior. I concur with previous arguments to further extend collaborations between security experts and social scientists [57], such as when attempting to generate constructive social interactions to strengthen an organization's resistance against phishing. It seems promising, for instance, to conduct further research into the psychology of social influence as well as encourage pro-social behaviors [23, 41].

Administrators and supervisors could have a particularly positive influence on the security of an organization. Training these groups about phishing-related requests from employees seems important. Future work should also investigate how positive social effects can be leveraged remotely.

IT context. In this study, participants were in their office and used their computers to receive and respond to emails, but they reported also using mobile devices to respond to work emails. Phishing on mobile phones is more difficult to recognize and leads to a high proportion of targets being phished [17]. In addition to the smaller screen size [17], touch screen interaction makes certain recommendations inapplicable (e.g., hovering over a URL is not possible). Phishing detection on mobile devices is particularly difficult because the sender's email address is typically hidden by default,

and it is difficult and risky to investigate a link destination [11]. It is important to further investigate how employees use mobile phones to interact with emails and which vulnerabilities this causes.

Task context. For some participants (e.g., administrative staff), responding to large amounts of emails is an integral part of their work day. Others could, in principle, turn off email notifications and respond to them in batches. These contextual factors are likely to play a role in how they respond to phishing attacks. The importance of task context, such as number of emails received in a day, has been highlighted previously [40], but more work is needed to fully understand how task context relates to phishing.

Temporal and physical context. The observation lasted one hour, which might have increased perceived time pressure when receiving the phishing email shortly before the end. Note that participants read most emails immediately, including the phishing email. No participant mentioned postponing their reaction to the phishing email to a later point in time. The study was not designed to study long-term behavior. This is a relevant topic for future research, especially considering that employee fatigue is likely to influence responses to phishing attacks, as well as the "seasons" of phishing (for instance around the end-of-year holidays [8]). The study does not investigate physical context beyond the office environment, but future work should study employee responses to phishing in a variety of physical contexts.

These results show that organizations' vulnerability to phishing is far from resolved. Reducing harm once an employee has interacted with a phishing email should be the priority. But how can organizations encourage employees to report suspicious emails, regardless of whether they have already interacted with the phish?

5.2 Identifying and Acting on a Suspicious Phishing Email: Reporting and Alternative Behaviors

Originally, I set out to understand why people report or do not report (potential) phishing emails. However, this dichotomy does not align with employees' views about potential actions. The participants thought about reporting as one of many possible actions. Alternative actions included deleting the email, forwarding the phish to the alleged sender, or marking the email as junk. Measurement data (e.g., embedded in phishing training) should thus be complemented with self-report measures (e.g., sent briefly after a simulated phishing email) to better understand these alternative actions and obtain a more realistic and nuanced understanding of how employees react to phishing. Some of the alternative actions should be discouraged, such as forwarding the phish to the apparent sender, as this person might then interact with the email and cause a security breach.

Reporting false positives – a problem? Barriers to reporting included not wanting to get others in trouble (for sending unprofessional emails) and not having strong evidence for a report. These arguments follow the line of thought of previous work, which concerned increasing the number of reported phishing emails while keeping the number of false positives low [22, 26].

Our results thus show that expectations regarding employees' reporting behavior were not clear enough. It is not realistic for an employee to "know" in all instances whether a suspicious email is phishing, spam, or legitimate. IT departments should more clearly communicate (1) why employees should report suspicious emails, and (2) whether accuracy (avoiding false positives) is more important or whether employees should report any potential phishing emails.

Critically Reflecting on Embedded Phishing Training. The study used the participating university's existing embedded phishing training infrastructure to study spear-phishing. Previous work has criticized such interventions [6, 7, 44, 50]. Employees might purposefully interact with simulated phishing messages for various reasons (e.g., curiosity, to learn more, or to boycott the campaign) [50], and the campaigns might negatively affect reporting of phishing messages. This study provides empirical support to these concerns. In the context of the embedded training, the participants argued that it did not make sense to report a phish after clicking on it. One of the participants said, when recalling the training website displayed after clicking on the phishing link: "This? I always just close it." The person (1) did not read the information on the website and (2) did not report the phishing email. This makes sense in the context of the embedded phishing training but would be harmful if employees exhibit the same behavior when confronted with a real phishing email and do not report the phishing attack after clicking on it, perhaps out of habit, or perhaps assuming that the IT department already knows about this phishing attempt as well. The results also point to a certain over-confidence stemming from successfully recognizing the basic phishing emails sent in the embedded phishing training. This connects to work in social sciences, which has found that people tend to be overconfident in their knowledge [14]. It seems relevant to investigate further how embedded phishing training influences participants' self-efficacy with phishing and whether a certain over-confidence might be caused that is not conducive to dealing with more advanced phishing attempts.

Previous work has argued for the use of "teachable moments" in the context of phishing, presenting information to users when they interact with a (simulated) phish [28]. Based on the results, I question whether the embedded training created good conditions for learning. Participants reported feeling ashamed, angry at themselves, upset and worried. Indeed, participants did not pay particular attention to the information provided, which is likely due to the emotional response, as well as their primary task still being on their mind. It might be more promising to provide the information as a follow-up email after users have had a chance to calm down.

Note that embedded training, if used despite criticisms [6, 7, 50], should be transparent about its intentions and consequences. Even if the organization or IT department does not intend to dole out punishments, participants' internal context (previous experience) can lead to a different perception. For instance, P13 in had previously worked in a company where clicking on embedded phishing training emails was punished and, if repeated, could lead to termination.

Rethinking the Narrative of Who Falls for Phishing. The results showed that context played an important role in how participants reacted to phishing. It is thus important to avoid describing certain groups of employees as more "vulnerable" to phishing, when their work context might be vastly different from other employees.

Anecdotally, the IT expert participant reached out some weeks after study participation, telling me how they fell for a phishing attack when trying to sell a personal item online, even transferring money to the attacker. They explained being less suspicious in this context and wanting to get the sale done. Only after making the transfer did they become suspicious and notice they had fallen for a scam that used a phishing email. This highlights why the narrative of certain people being more susceptible to phishing can be so harmful and can lead to a false sense of security. Changing the narrative to "anyone can fall for phishing if caught at a bad time" is more constructive for improving organizational security.

5.3 Participants' Suggested Solutions to Counteract Phishing

Participants suggested some adaptations to training procedures, including reducing the shame induced by the training (also suggested by [7]), improved instructions on how to answer emails (e.g., avoid responding when distracted), how to recognize external email addresses, and paying attention to the writing style. Participants also suggested training to improve awareness of advanced phishing attacks (such as the spear-phishing used in this study), such as informing employees of ongoing phishing attacks that were particularly good or targeted (showing awareness of seasonal phishing trends [8]), and improving communication surrounding the reporting procedure to lower the barrier to reporting. These suggested solutions reflect the recommendations made by cybersecurity organizations [6, 7]. On the topic of technical solutions, participants suggested a warning when the sender email address or link URL is particularly long.

5.4 Recommendations for User-Centered Anti-Phishing Interventions

- Contextual factors (particularly internal context, social context and task context) play an important role in how people deal with a suspicious email. Anti-phishing interventions should attempt to improve contextual factors to support employees. Examples include creating a social environment that encourages employees to talk about phishing, discouraging answering emails when tired or under time pressure, and changing social norms to lower the expectation that emails be answered immediately.
- Training administrative staff and team leaders to help them address phishing-related questions from other employees seems promising, as they can be the first point of contact for employees who are uncertain of the correct procedures.
- Future work should investigate how to best introduce friction design to give employees the space to consciously consider how to best respond to an email.
- The results lend empirical support to previous criticism of embedded phishing training. Future research should critically reflect on the objectives and effects of these training interventions. One potential shortcoming is that it might train employees not to report a phishing email after clicking on the link, in addition to encouraging interacting with the email.
- Reporting procedures should be as simple as possible (ideally, merely clicking a button). IT departments need to communicate clearly how and why to report phishing emails.
- Participants were unsure of whether falsely reporting a legitimate email would be problematic. If feasible, it seems promising to advertise the reporting procedure as a service for employees to check potentially suspicious emails, where submitting a false positive is not problematic.
- Communication should focus on the fact that reporting phishing is especially important after the dangerous action has already been taken (e.g., clicking on the link, downloading the file, providing personal information). Shaming participants who fell for a phish should be avoided to encourage reporting and open communication.

5.5 Limitations and Future Work

The present study is less conducive to observing the influence of certain dimensions of context compared to others. For instance, temporal context might play a more important role than I find in this paper. As the study period was during the work day and limited to one hour of observation, I was not able to observe the full range of possible influences of temporal context. Similarly, physical context (offices) was relatively similar across participants, meaning that I did not observe large variations in how participants perceived the physical context.

While I did not instruct participants to use their computers, all of the participants received and evaluated the phishing email on their computers, but reported using mobile devices outside of the office. Mobile devices can make it more difficult to follow guidelines [11, 17]. These issues should be investigated in future studies.

This study was not aimed at generalizability of results. All participants worked in a university environment and had advanced degrees. When participants take part in studies on a voluntary basis, this leads to self-selection bias. While I made substantial efforts to recruit participants, despite an attractive compensation, only a relatively small number of employees could be convinced to take part. In future work, it seems relevant to investigate how more employees might be motivated to participate and whether certain aspects of the study set-up are perceived as intrusive by certain employees. Employees who felt comfortable being watched while working were more likely to sign up for this study, and they might differ from other employees in terms of phishing-related behavior.

Note that the first three participants received the phishing attempt in English. The participating university is a multilingual work environment. Employees commonly use three languages. The first three employees received the phishing attempt in English, as is standard procedure in the simulated anti-phishing training. However, they immediately recognized the phish because they would expect the sender to address them in another language. They might have fallen for the phishing email if the language had not made them suspicious. I therefore adapted the language of the email for subsequent participants to be in line with their expected language by asking the professor which language they would use with a certain employee. While this iterative methodology is typical for qualitative studies, I flag this as a methodological challenge for phishing studies that are conducted in multilingual environments.

6 CONCLUSION

Phishing attacks remain a harmful and prevalent form of cyber-attack for both individuals and organizations. In organizational settings, phishing attacks are often only first point of entry for more far-reaching attacks (e.g., ransomware) and cause societal, financial and physical harm. A variety of user-oriented phishing interventions have been used in an attempt to address phishing, including education, training, awareness training, or design interventions [16]. Despite the variety of interventions, organizations remain vulnerable to phishing attacks. We lack knowledge of how the work context can influence employees' reactions to phishing attempts. Focusing on spear-phishing, I investigated (1) how contextual factors influence reactions to a spear-phishing attempt, (2) why people report or do not report phishing attempts, and (3) which opportunities for security-enhancing interventions people identify.

I observed participants in their realistic work environment and triangulated these observations with self-reported data collected directly after the observation (interviews, UX curves).

Important findings include that I observed two main thought processes among participants after receiving the phishing attempt. I found that contextual factors played an important role in how employees responded to spear-phishing. The social environment was particularly relevant in shaping the initial response to phishing and the reporting behavior. Participants reported phishing attempts to improve filters and get the positive feedback email. Obstacles to reporting included being unsure of how and why to report. Participants' suggested solutions focused on employee training (e.g., training for more advanced phishing) and raising awareness (e.g., ongoing phishing attacks, re-branding phishing reporting).

This work provides multiple avenues for future research, which should further investigate the influence of contextual factors on phishing. I suggest that organizations build on current knowledge about contextual influences on reactions to phishing and adapt the context wherever possible to support employees' responses to phishing attacks. When a person, inevitably, "falls" for a phish, organizations should not focus on punishing the employee, as this is unlikely to lead to more secure behavior in the future. Instead, organizations should focus on providing the appropriate support to ensure employees report the mistake and limit harmful consequences. In addition, I suggest that organizations should investigate how context might have played a role in the employees' actions, and make improvements for the future.

ACKNOWLEDGMENTS

I thank the research participants for taking the time to participate in this study and for sharing their experiences with me. I am grateful to my colleagues from the HCI Research Group at the University of Luxembourg for their feedback and insightful discussions, as well as their participation in early pre-tests of the study protocol. I thank Vincent Koenig for the support and feedback on the study methodology. I also thank the anonymous reviewers.

REFERENCES

[1] Kholoud Althobaiti, Nicole Meng, and Kami Vaniea. 2021. I Don't Need an Expert! Making URL Phishing Features Human Comprehensible. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. ACM, Yokohama Japan, 1–17. https://doi.org/10.1145/3411764.3445574

[2] Jeremy Barr. 2020. The company email promised bonuses. It was a hoax — and Tribune Publishing employees are furious. *Washington Post* (2020). https://www.washingtonpost.com/media/2020/09/23/tribune-bonus-email-phishing-hoax/

[3] Nicholas Bradley and Mark Dunlop. 2005. Toward a Multidisciplinary Model of Context to Support Context-Aware Computing. *Human-Computer Interaction* 20, 4 (Dec. 2005), 403–446. https://doi.org/10.1207/s15327051hci2004_2

[4] Cristian Bravo-Lillo, Saranga Komanduri, Lorrie Faith Cranor, Robert W. Reeder, Manya Sleeper, Julie Downs, and Stuart Schechter. 2013. Your attention please: designing security-decision UIs to make genuine risks harder to ignore. In *Proceedings of the Ninth Symposium on Usable Privacy and Security - SOUPS '13*. ACM Press, Newcastle, United Kingdom, 1. https://doi.org/10.1145/2501604.2501610

[5] Gamze Canova, Melanie Volkamer, Clemens Bergmann, and Roland Borza. 2014. NoPhish: An Anti-Phishing Education App. In *Security and Trust Management*, Sjouke Mauw and Christian Damsgaard Jensen (Eds.). Vol. 8743. Springer International Publishing, Cham, 188–192. https://doi.org/10.1007/978-3-319-11851-2_14 Series Title: Lecture Notes in Computer Science.

[6] National Cyber Security Centre. 2018. The Trouble with Phishing. https://www.ncsc.gov.uk/blog-post/trouble-phishing

[7] National Cyber Security Centre. 2019. Phishing attacks: defending your organisation. https://www.ncsc.gov.uk/guidance/phishing

[8] CISCO. 2021. *Cyber security threat trends*. Technical Report. https://learn-umbrella.cisco.com/ebook-library/2021-cyber-security-threat-trends-phishing-crypto-top-the-list

[9] Verena Distler, Matthias Fassl, Hana Habib, Katharina Krombholz, Gabriele Lenzini, Carine Lallemand, Lorrie Faith Cranor, and Vincent Koenig. 2021. A Systematic Literature Review of Empirical Methods and Risk Representation in Usable Privacy and Security Research. *ACM Transactions on Computer-Human Interaction* 28, 6 (Dec. 2021), 1–50. https://doi.org/10.1145/3469845

[10] Verena Distler, Gabriele Lenzini, Carine Lallemand, and Vincent Koenig. 2020. The Framework of Security-Enhancing Friction: How UX Can Help Users Behave More Securely. In *New Security Paradigms Workshop 2020*. ACM, Online USA, 45–58. https://doi.org/10.1145/3442167.3442173

[11] Matt Dixon, James Nicholson, Dawn Branley-Bell, Pam Briggs, and Lynne Coventry. 2022. Holding Your Hand on the Danger Button: Observing User Phish Detection Strategies Across Mobile and Desktop. *Proceedings of the ACM on Human-Computer Interaction* 6, MHCI (Sept. 2022), 1–22. https://doi.org/10.1145/3546730

[12] Julie S. Downs, Mandy Holbrook, and Lorrie Faith Cranor. 2007. Behavioral response to phishing risk. In *Proceedings of the anti-phishing working groups 2nd annual eCrime researchers summit on - eCrime '07*. ACM Press, Pittsburgh, Pennsylvania, 37–44. https://doi.org/10.1145/1299015.1299019

[13] Julie S. Downs, Mandy B. Holbrook, and Lorrie Faith Cranor. 2006. Decision strategies and susceptibility to phishing. In *Proceedings of the second symposium on Usable privacy and security - SOUPS '06*. ACM Press, Pittsburgh, Pennsylvania, 79. https://doi.org/10.1145/1143120.1143131

[14] Matthew Fisher and Frank C. Keil. 2016. The Curse of Expertise: When More Knowledge Leads to Miscalibrated Explanatory Insight. *Cognitive Science* 40, 5 (2016), 1251–1269. https://doi.org/10.1111/cogs.12280 _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/cogs.12280.

[15] Centre for the Protection of National Infrastructure. 2021. Don't take the bait! https://www.cpni.gov.uk/security-campaigns/don%E2%80%99t-take-bait-0

[16] Anjuli Franz, Verena Zimmermann, Gregor Albrecht, Katrin Hartwig, Christian Reuter, Alexander Benlian, and Joachim Vogt. 2021. 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, 339–358. https://www.usenix.org/conference/soups2021/presentation/franz

[17] Diksha Goel. 2018. Mobile phishing attacks and defence mechanisms: State of art and open research challenges. (2018), 26.

[18] Kristen Greene, Michelle Steves, Mary Theofanos, and Jennifer Kostick. 2018. User Context: An Explanatory Variable in Phishing Susceptibility. In *Proceedings 2018 Workshop on Usable Security*. Internet Society, San Diego, CA. https://doi.org/10.14722/usec.2018.23016

[19] Frank L. Greitzer, Wanru Li, Kathryn B. Laskey, James Lee, and Justin Purl. 2021. Experimental Investigation of Technical and Human Factors Related to Phishing Susceptibility. *ACM Transactions on Social Computing* 4, 2 (June 2021), 1–48. https://doi.org/10.1145/3461672

[20] Andrew Hinton. 2014. *Understanding context: Environment, language, and information architecture*. " O'Reilly Media, Inc.".

[21] Karen Holtzblatt and Hugh Beyer. 1998. *Contextual Design: Defining Customer-Centered Systems*. Morgan Kaufmann Publishers.

[22] Matthew L. Jensen, Ryan T. Wright, Alexandra Durcikova, and Shamya Karumbaiah. 2022. Improving Phishing Reporting Using Security Gamification. *Journal of Management Information Systems* 39, 3 (July 2022), 793–823. https://doi.org/10.1080/07421222.2022.2096551

[23] Klaus Jonas, Wolfgang Stroebe, and Miles Hewstone. 2014. *Sozialpsychologie* (6 ed.). Springer, Berlin Heidelbert.

[24] Helen S. Jones, John N. Towse, Nicholas Race, and Timothy Harrison. 2019. Email fraud: The search for psychological predictors of susceptibility. *PLOS ONE* 14, 1 (Jan. 2019), e0209684. https://doi.org/10.1371/journal.pone.0209684

[25] Daniel Kahneman. 2011. *Thinking, fast and slow.* Farrar, Straus and Giroux, New York.

[26] Leon Kersten, Pavlo Burda, Luca Allodi, and Nicola Zannone. 2022. Investigating the Effect of Phishing Believability on Phishing Reporting. In *2022 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW)*. 117–128. https://doi.org/10.1109/EuroSPW55150.2022.00018 ISSN: 2768-0657.

[27] Sari Kujala, Virpi Roto, Kaisa Väänänen-Vainio-Mattila, Evangelos Karapanos, and Arto Sinnelä. 2011. UX Curve: A method for evaluating long-term user experience. *Interacting with Computers* 23, 5 (Sept. 2011), 473–483. https://doi.org/10.1016/j.intcom.2011.06.005

[28] Ponnurangam Kumaraguru, Lorrie Faith Cranor, and Laura Mather. 2009. Anti-phishing landing page: Turning a 404 into a teachable moment for end users. In *Conference on Email and Anti-Spam (CEAS)*. Citeseer.

[29] Ponnurangam Kumaraguru, Justin Cranshaw, Alessandro Acquisti, Lorrie Cranor, Jason Hong, Mary Ann Blair, and Theodore Pham. 2009. School of Phish: A Real-World Evaluation of Anti-Phishing Training. In *Proceedings of the 5th Symposium on Usable Privacy and Security (SOUPS '09)*. Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/1572532.1572536 event-place: Mountain View, California, USA.

[30] Youngsun Kwak, Seyoung Lee, Amanda Damiano, and Arun Vishwanath. 2020. Why do users not report spear phishing emails? *Telematics and Informatics* 48 (May 2020), 101343. https://doi.org/10.1016/j.tele.2020.101343

[31] Sam Ladner. 2014. *Practical Ethnography: A Guide to Doing Ethnography in the Private Sector.* Taylor & Francis Group, Walnut Creek, UNITED STATES. http://ebookcentral.proquest.com/lib/unilu-ebooks/detail.action?docID=1656740

[32] Carine Lallemand and Vincent Koenig. 2020. Measuring the Contextual Dimension of User Experience: Development of the User Experience Context Scale (UXCS). In *Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society.* ACM, Tallinn Estonia, 1–13. https://doi.org/10.1145/3419249.3420156

[33] Susan Landau. 2017. *Listening In: Cybersecurity in an Insecure Age.* Yale University Press, New Haven and London.

[34] Tian Lin, Daniel E. Capecci, Donovan M. Ellis, Harold A. Rocha, Sandeep Dommaraju, Daniela S. Oliveira, and Natalie C. Ebner. 2019. Susceptibility to Spear-Phishing Emails: Effects of Internet User Demographics and Email Content. *ACM Transactions on Computer-Human Interaction* 26, 5 (Sept. 2019), 1–28. https://doi.org/10.1145/3336141

[35] Sean Lyngaas. 2022. Ransomware attack hits New Jersey county. https://www.cnn.com/2022/05/26/politics/new-jersey-somerset-county-ransomware-attack/index.html

[36] Phillipp Mayring. 2015. *Qualitative Inhaltsanalyse* (12 ed.). Beltz, Weinheim und Basel.

[37] Nora McDonald, Sarita Schoenebeck, and Andrea Forte. 2019. Reliability and Inter-rater Reliability in Qualitative Research: Norms and Guidelines for CSCW and HCI Practice. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (Nov. 2019), 1–23. https://doi.org/10.1145/3359174

[38] Geir Moulson. 2021. Germany protests to Russia over pre-election cyberattacks. https://apnews.com/article/technology-europe-russia-elections-germany-26ea77a3b96b94d5760aab48c9dfc008 Section: Technology.

[39] Federal Bureau of Investigation. 2022. *2021 Internet Crime Report.* Technical Report.

[40] Daniela Oliveira, Harold Rocha, Huizi Yang, Donovan Ellis, Sandeep Dommaraju, Melis Muradoglu, Devon Weir, Adam Soliman, Tian Lin, and Natalie Ebner. 2017. 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 (CHI '17)*. Association for Computing Machinery, New York, NY, USA, 6412–6424. https://doi.org/10.1145/3025453.3025831 event-place: Denver, Colorado, USA.

[41] Gregor Petrič and Kai Roer. 2022. The impact of formal and informal organizational norms on susceptibility to phishing: Combining survey and field experiment data. *Telematics and Informatics* 67 (Feb. 2022), 101766. https://doi.org/10.1016/j.tele.2021.101766

[42] Stefan Rädiker and Udo Kuckartz. 2019. *Analyse qualitativer Daten mit MAXQDA.* Springer VS, Wiesbaden.

[43] Kim Salazar. 2020. Contextual Inquiry: Inspire Design by Observing and Interviewing Users in Their Context. https://www.nngroup.com/articles/contextual-inquiry/

[44] M. Angela Sasse, Jonas Hielscher, and Marco Gutfleisch. 2022. Human-Centred Security: Unfug Informationssicherheits-Sensibilisierung. *kma - Klinik Management aktuell* 27, 04 (Aug. 2022), 44–46. 44.

[45] Peter Schaab, Kristian Beckers, and Sebastian Pape. 2016. A Systematic Gap Analysis of Social Engineering Defence Mechanisms Considering Social Psychology.. In *HAISA*. 241–251.

[46] Steve Sheng, Mandy Holbrook, Ponnurangam Kumaraguru, Lorrie Faith Cranor, and Julie Downs. 2010. Who falls for phish?: a demographic analysis of phishing susceptibility and effectiveness of interventions. In *Proceedings of the 28th international conference on Human factors in computing systems - CHI '10*. ACM Press, Atlanta, Georgia, USA, 373. https://doi.org/10.1145/1753326.1753383

[47] Lucy A. Suchman. 1987. *Plans and Situated Actions: The Problem of Human-Machine Communication.* Cambridge University Press, USA.

[48] Arun Vishwanath, Tejaswini Herath, Rui Chen, Jingguo Wang, and H. Raghav Rao. 2011. Why do people get phished? Testing individual differences in phishing vulnerability within an integrated, information processing model. *Decision Support Systems* 51, 3 (June 2011), 576–586. https://doi.org/10.1016/j.dss.2011.03.002

[49] Melanie Volkamer, Karen Renaud, Benjamin Reinheimer, and Alexandra Kunz. 2017. User experiences of TORPEDO: TOoltip-poweRed Phishing Email DetectiOn. *Computers & Security* 71 (Nov. 2017), 100–113. https://doi.org/10.1016/j.cose.2017.02.004

[50] Melanie Volkamer, Martina A. Sasse, and Franziska Boehm. 2020. Phishing-Kampagnen zur Steigerung der Mitarbeiter-Awareness: Analyse aus verschiedenen Blickwinkeln – Security, Recht und Faktor Mensch. *Datenschutz und Datensicherheit - DuD* 44, 8 (Aug. 2020), 518–521. https://doi.org/10.1007/s11623-020-1317-x

[51] Rick Wash. 2020. How Experts Detect Phishing Scam Emails. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW2 (Oct. 2020). https://doi.org/10.1145/3415231 Place: New York, NY, USA Publisher: Association for Computing Machinery.

[52] Rick Wash, Norbert Nthala, and Emilee Rader. 2021. Knowledge and Capabilities that Non-Expert Users Bring to Phishing Detection. In *Seventeenth Symposium on Usable Privacy and Security (SOUPS 2021)*. USENIX Association, 377–396. https://www.usenix.org/conference/soups2021/presentation/wash

[53] Emma J. Williams, Joanne Hinds, and Adam N. Joinson. 2018. Exploring susceptibility to phishing in the workplace. *International Journal of Human-Computer Studies* 120 (Dec. 2018), 1–13. https://doi.org/10.1016/j.ijhcs.2018.06.004

[54] Carly Wilson and David Argles. 2011. The fight against phishing: Technology, the end user and legislation. In *International Conference on Information Society (i-Society 2011)*. 501–504. https://doi.org/10.1109/i-Society18435.2011.5978553

[55] Dennis Wixon, Alicia Flanders, and Minette A. Beabes. 1996. Contextual Inquiry: Grounding Your Design in User's Work. In *Conference Companion on Human Factors in Computing Systems (CHI '96)*. Association for Computing Machinery, New York, NY, USA, 354–355. https://doi.org/10.1145/257089.257365 event-place: Vancouver, British Columbia, Canada.

[56] Weining Yang, Aiping Xiong, Jing Chen, Robert W. Proctor, and Ninghui Li. 2017. Use of Phishing Training to Improve Security Warning Compliance: Evidence from a Field Experiment. In *Proceedings of the Hot Topics in Science of Security: Symposium and Bootcamp on - HoTSoS*. ACM Press, Hanover, MD, USA, 52–61. https://doi.org/10.1145/3055305.3055310

[57] Sarah Zheng and Ingolf Becker. 2022. Presenting Suspicious Details in User-Facing E-mail Headers Does Not Improve Phishing Detection. In *Eighteenth Symposium on Usable Privacy and Security (SOUPS 2022)*. USENIX Association, Boston, MA, 253–271. https://www.usenix.org/conference/soups2022/presentation/zheng

A RECRUITMENT FLYER

Figure 5 shows the information provided on the recruitment flyer.

Figure 5: The recruitment flyer sent to participants.

B WEBSITE SHOWN AFTER CLICKING ON SIMULATED PHISHING LINK

Note that I used the default website used by the university's security team in their phishing simulations, without making adaptations

to study phishing simulation trainings as they might happen in organizations.

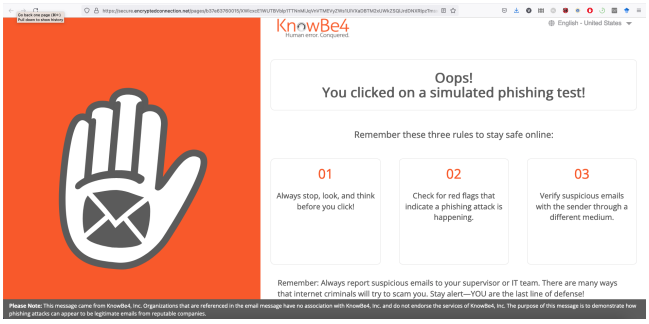

Figure 6: Screenshot of the website shown to participants who clicked on the phishing link.

C STUDY PROTOCOL

- On the day of the session, the researcher comes to the participant's office.
  - "The objective today is to understand how people use technology at work. This is intentionally a bit vague, and we will talk about all of this in a bit more detail later on."
  - "I can already tell you, however, that we will not evaluate your performance."
- Talk participant through informed consent sheet.
- Participant questions are answered.
- While the researcher sets up the camera, the participant fills out the short demographics questionnaire on their computer.
- Live stream with camera is set up. (The live stream is not recorded.)
- Participants are asked to work as they usually would for the next hour. After this hour, the researcher will come back for an interview.
  - "In this hour, I will ask you to work as you normally would. Feel free to get up, make calls,... Please try not to change your behavior. You also don't have to send colleagues away when they come ask you something or anything like that. Feel free to respond to phone calls, emails, messages, if that's normal for you, just so we can get a realistic impression."
- The researcher watches the live stream in a nearby conference room and takes notes. The participant works as usual.
- After 55 minutes, the simulated phishing from SIU email arrives.
- The study continues for about another 5 minutes.
- The researcher comes back to the office.
- If the participant does not read the email: Before leaving the participant's office, the participant is asked to check all their remaining emails in their inbox and to briefly state what they would do for each of the emails.
- The researcher and the participant go to the conference room (if shared office).
- Interview about general work environment and contextual factors (RQ1).
  - Can you explain what you were working on?
  - Would you say what I observed was somewhat typical for your work environment?
  - How does what is typical change throughout the day?
  - How many emails do you think that you get in a day approximately?
  - As how stressful do you perceive your daily work? And today?
  - Do you sometime get interrupted when working? For example colleagues, messaging apps, phone,...
  - Follow up prompts on the interruptions mentioned.
- Explaining the deception: the phishing attack actually came from the researcher.
  - "It is really hard to recognize phishing emails, and even highly educated and tech-savvy people and security experts can click on malicious links. This is why we wanted to conduct this study to understand more than just the numbers in the system, but instead how people feel about phishing emails."
  - Give them the information sheet.
  - "Did you already have a suspicion?"
  - "Do you understand why we did not tell you the full objective of the study in the beginning?"
  - "Do you have any concerns or questions about this?"
  - Renewed consent: "Are you still ok with participating? You can also withdraw from the study now."
- The phishing attack (RQ1)
  - Task context: "Let's think back to when the phishing attack happened. Can you explain to me what you were trying to do at that point?"
  - Participant state: "What were you thinking about?" "How were you feeling?"
  - Thought process: "So you told me that you were trying to do … and were feeling … What was your thought process then?"
- Reporting/not reporting (RQ2)
  - "So once the phishing email arrived, what did you do then?"
  - "Is this what you usually do when you get this type of email?"
  - "Some people actually write to inform the university of the email. Have you ever done this? Why? What was your experience?"
  - "What would you expect to happen?"
- Opportunities for intervention (RQ3)
  - The participant is asked to draw a UX curve thinking back about their experience.
  - "Can you explain what you drew here?"
  - The participant is asked to add the phishing email to the drawing.
  - "Imagine another person, for example a colleague that you like, received this phishing attack. How could we help them, so that they don't click on it? Can you draw these opportunities here? "
- Wrapping up
  - Do you have any other remarks?
  - Do you have any questions?

  - While data collection is ongoing: "Please don't tell your colleagues about the hidden objective of this study."
- Compensation. Participant is invited to ask any questions they might have about the study in the future.

D UX CURVE TEMPLATE

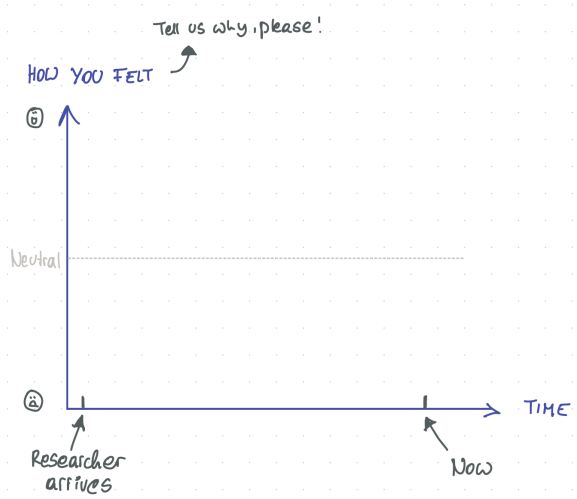

Figure 7: The template used when asking participants to draw the UX curve.

E DETAILED DESCRIPTION OF THE QUALITATIVE CONTENT ANALYSIS

All interviews were transcribed. I used Mayring's process for a "summarizing content analysis" as a guideline [36, p. 72] and adapted it to the study's specific material. I used MAXQDA 2022 and followed guidelines by Rädiker and Kuckartz to adapt Mayring's process [42, p. 146].

Based on approximately 30% of the material (4/14 interviews), the author created paraphrases for each element carrying a distinct meaning in the text ("Sinneinheit"), distilling the meaning of the text into short paraphrases (phase 1, paraphrasing).

In the second phase, the author generalized the content of the paraphrases. To do so, I did not delete any paraphrases, as cautioned against by [42, p. 146]. Instead, at this stage, paraphrases were generalized to a first level of abstraction using the "coding" function in MAXQDA. This resulted in a set of categories ("codes" in MAXQDA) that were more abstract than the paraphrases, but not yet abstract enough for further analysis. Statements that were not considered relevant for further treatment were not coded.

In the next phase, I further generalized the codes to higher levels of abstraction for further analysis. At this level, codes were subsumed into higher, more abstract categories. Redundant codes were merged. At this point, I then broadly categorized these codes deductively (based on the concepts of interest in the research questions) according to the topics "context" (RQ1), "what was happening when the email has arrived" (RQ2), and "opportunities for intervention" (RQ3).

I went on to paraphrase the remaining interviews by each element carrying a distinct meaning in the text. Elements that were irrelevant to the objectives of the study were not paraphrased. I then applied the previously created preliminary code system to the material. Throughout the paraphrasing and coding, I also wrote down any higher-level observations in the document memos.

In the next step, I merged the codes so that they would be appropriate to respond to the research questions. I used MAXQDA's creative coding function to gain an overview of the codes, merged redundant or overlapping codes, and refined the code definitions. I summarized the results for each participant per research question using the document summary function. I applied the same code system to the observations and UX curves. I iteratively improved and adapted the code system, merging codes or adding new codes when required.

After coding, I created case summaries for each research participant, combining the data sources described in table 1. I then used the MAXQDA summary grid function to create case summaries for each research question. I double-checked the coded transcripts to see whether I missed any information relevant for the RQs. I then created summary tables for each research question that included participant characteristics (using the MAXQDA document variables), such as whether the person had clicked on the link and had reported the phishing attempt. Based on these summary tables, I wrote the results section of this paper.

Changes made after co-coding session. As described in section 3, I organized a co-coding session and responded to disagreement and questions by clarifying the codebook. I found that it was unclear to coders whether mentions of emails should go into the IT context or internal context (as participants expressed their feelings about the emails). I decided to include mentions of how people experienced interactions with emails as part of internal context, and ensured that this was consistently applied through the analysis. I also found that the distinction between training and communication interventions was unclear and discussed this with a fourth HCI expert. I decided to merge the categories (resulting in the code "training and create awareness"). Another question that came up was whether the perceived urgency to respond to an email should be coded as temporal context. I decided to interpret this as a consistent feeling of pressure to answer immediately. Therefore, I categorized it within internal context.

F CODEBOOK

- 1. Receiving and evaluating the email
  - *1.1 elements that raise suspicion:* This code includes the elements that made participants suspicious of the email. *Examples:* "My supervisor would just come into the office to tell me this, instead of sending an email."
  - *1.2 strategies in case of doubt:* This code includes any statements about participants' strategies when they are not sure whether an email they have received is legit. *Examples:* "When I get an email telling me to log in somewhere, I type in the URL myself. Best not to follow links."
  - *1.3 subverting the system:* This code includes mentions of consciously subverting the simulated phishing trainings. *Examples:* "I programmed my email program to automatically highlight simulated phishing emails from the IT

department." "I sometimes click on simulated phishing emails on purpose."
  - *1.4 weighing the risks:* This code includes any weighing of risks by the participants, both of clicking on a phishing link and of falsely disregarding a legitimate email. *Examples:* "Clicking on a phishing link is not really that dangerous as long as you don't type in information" "I would be worried to think that a legitimate email from my supervisor is a phish"
  - *1.5 reporting the thought process:* This code includes descriptions of participants' thought process when they received the email. *Examples:* "When I saw my professor's name on the email, I thought it must be urgent."
  - *1.6 observation of attack:* This code includes observations describing the situation when participants received the simulated phishing attempt. *Example:* Observation note: "Participant immediately tells colleagues about the email."
- 2. Contextual factors
  - *2.1 internal context:* How the participant felt during the participation, their attitude to work, emails, distractions. *Examples:* "I feel pressure to answer emails immediately." "I felt overwhelmed."
  - *2.2 social context:* This code describes the influence of social aspects on the participant. *Examples:* Interview: "Usually, we are more people in the office. Today, I was on my own." Observation: The participant tells colleagues about the phishing attack.
  - *2.3 IT Context:* Describes technical aspects of the participant's work environment. *Examples:* "We use Microsoft Teams a lot, so I get a lot of notifications." Participant starts a phone call.
  - *2.4 task context:* Description of the participant's task during their study participation. *Examples:* "I was trying to find a reference online."
  - *2.5 temporal aspects (time of the day / month):* How participants' work changes throughout the day, week or month. *Examples:* "In the morning, I take care of emails. In the afternoon, I have meetings."
  - *2.6 physical context:* Describes the participant's physical context. *Examples:* door open, noisy environment
- **3. after the attack (incl. reporting / not reporting) (RQ2)**
  - *3.1 emotions, thoughts and actions after clicking on the phishing link:* Describes participants' emotions, thoughts and actions after clicking on a phishing link (both in the experiment or earlier). When not to use: Only use when the participant was talking about their own experience. Do not use when they talk about somebody else. *Examples:* "When I saw the IT warning website, I just thought "damn". It annoys me when I fall for it."
  - *3.2 reporting:* Describes participants' rationalizations of when they should or shouldn't report. Also includes social interactions that influence reporting behaviors. *Examples:* "always" "only when not stressed"
  - *3.3 alternatives to reporting:* Describes alternative actions participants take or have taken instead of reporting the simulated phishing email. *Example:* "I always delete phishing emails." "I would ask our secretary what to do."
- **4 interventions against phishing (RQ3)**
  - *4.1 training and create awareness:* Suggestions related to training employees to recognize phishing attempts or generate awareness about phishing-related issues. *Example:* "We could teach people to be careful about emails from their boss that don't come from the internal email address."
  - *4.2 technical solutions:* Suggested technical solutions. *Example:* "As soon as the email address includes a similar ending to our legitimate email addresses, it should give me a warning."