What Makes Phishing Simulation Campaigns (Un)Acceptable? A Vignette Experiment

Jasmin Schwab*, Alexander Nussbaum†, Anastasia Sergeeva‡, Florian Alt†§ and Verena Distler¶
*German Aerospace Center (DLR), Sankt Augustin, Germany, jasmin.schwab@dlr.de
†University of the Bundeswehr, Munich, Germany
‡University of Luxembourg, Esch-sur-Alzette, Luxembourg
§Ludwig-Maximilians-Universität, Munich, Germany
¶Aalto University, Espoo, Finland

***Abstract*—Organizations depend on their employees' long-term cooperation to help protect the organization from cybersecurity threats. The acceptance of training measures is thus crucial. Phishing attacks are the entry point for harmful follow-up attacks, and many organizations use simulated phishing campaigns to train employees to adopt secure behaviors. We conducted a pre-registered vignette experiment (*N*=793), investigating the factors that make a simulated phishing campaign seem (un)acceptable, and their influence on employees' intention to manipulate the campaign. In the experiment, we varied whether employees gave prior consent, whether the phishing email promised a financial incentive, and the consequences for employees who clicked on the phishing link. We found that employees' prior consent positively affected the acceptance of a simulated phishing campaign. The consequences "employee interview" and "termination of the work contract" negatively affected acceptance. We found no statistically significant effects of consent, monetary incentive, and consequences on manipulation probability. Our results shed light on the factors influencing the acceptance of simulated phishing campaigns. Based on our findings, we recommend that organizations prioritize obtaining informed consent from employees before including them in simulated phishing campaigns, and that they clearly describe their consequences. Organizations should carefully evaluate the acceptance of simulated phishing campaigns, and consider alternative anti-phishing measures.**

I. Introduction

Organizations rely heavily on email for both internal and external communication and are thus highly vulnerable to phishing attacks. Phishing describes an attack to obtain personal or confidential information from the victim through a message (e.g., an e-mail) by deliberately deceiving them and tricking them into harmful actions [1], [2]. Potential harms include personal damage to the person who interacted with a phishing email and their contacts, financial loss to individuals and organizations [3], harm to infrastructure such as power outages[4], [5], and subsequent societal impact such as problems with public infrastructure that relies on electricity to function (e.g., hospitals).

Symposium on Usable Security and Privacy (USEC) 2025
24 February 2025, San Diego, CA, USA
ISBN 979-8-9919276-5-9
https://dx.doi.org/10.14722/usec.2025.23010
www.ndss-symposium.org, https://www.usablesecurity.net/USEC/

Phishing detection inherently involves human and technical aspects. Technical phishing detection plays an important role in phishing protection, for instance, by preventing phishing emails from making it to a target's inbox. The main technical strategies for phishing detection include blocking known phishing URLs and removing known landing pages [1]. Other options include checking whether websites include certain characteristics that could be associated with phishing and page similarity detection [6]. A limitation of technical phishing mitigations is that they often work best when many similar emails are sent to many potential victims. Technical solutions often cannot protect the first person(s) who are exposed to a certain attack [7] especially in the case of highly targeted phishing attacks ("spear-phishing" [6]).

Many attacks cannot be filtered out through technical means. Organizations thus attempt to train employees to protect the organizations from phishing attacks, such as raising awareness of phishing and training [8] and cybersecurity incident reporting [9]. Simulated phishing campaigns are frequently used training opportunities. In such campaigns, organizations send "realistic" phishing emails to their own employees, often for training or evaluation purposes [8], [10]. Commercial vendors conduct such simulated phishing campaigns as a service, providing a user-friendly interface to conduct the campaign and analyze the results. Typically, if an employee interacts with the phishing email, they are re-directed to a training website explaining the indicators of phishing emails [10]. In some organizations, there are additional consequences for employees interacting with a simulated phishing email. These range from mandatory training to talking about cybersecurity, or in extreme cases, disciplinary consequences [11], [12].

Simulated phishing campaigns have faced serious criticism. It is unclear whether simulated phishing campaigns lead to the promised outcome of making organizations' employees adopt more secure email behaviors, and they might even negatively affect the security of organizations [13], [14], [15]. In addition to possibly not improving organizational security, these campaigns can have subtle negative effects, such as giving the impression that the IT department is trying to trick employees, of adding an additional burden on employees who are already often under pressure to perform in their job tasks [16] and might even lead to an insider threat through intentionally

manipulating phishing campaigns [17]. IT departments also depend on employees' long-term collaboration to keep their organization safe and to keep learning and adapting to new security threats. Recent work has highlighted the importance of enhancing perceived utility value of training interventions, prioritizing positive user experiences, and creating training interventions that are in line employees' motivations to engage them with phishing interventions [18], [19].

Despite criticism and possible lack of effectiveness, the current reality is that simulation campaigns are commonly employed in organizations. Organizations use various different implementations of phishing simulation campaigns. They might inform employees of the simulation in advance, or keep the campaign a secret. They might use certain pre-texts that are perceived as offensive, or avoid doing so. These decisions can influence how employees perceive the resulting campaign. Consider, for instance, the public outcry after a newspaper offered bonuses to employees in a simulated phishing email [20]. In the present paper, we explore the factors making a simulated phishing campaign seem (un)acceptable. This study should not be seen as an encouragement to conduct phishing simulation campaigns, but rather, a first investigation into the factors likely to influence employee acceptance. This paper does not investigate the effects of such campaigns on behavior when exposed to phishing, but we point the reader to insightful discussions and criticism [13], [14], [15].

We conducted a pre-registered between-subjects vignette experiment (N=793)[1] to investigate the factors affecting the acceptance of a simulated phishing campaign in an organization. Participants were asked to put themselves in the role of a caseworker who receives a draft for a possible phishing campaign to be carried out in their company. The participant should make an assessment about the acceptance of the proposed phishing campaign of the respective campaign on the basis of the information available by answering the questions in the questionnaire. Our results show that prior consent to being involved in a simulated phishing campaign positively affected the acceptance of a phishing campaign. In contrast, the content of the phishing email, including an incentive, had a negative effect on the acceptance of the campaign. We also found that the various consequences of interacting with the phishing campaign had different effects on the acceptance of the campaign. Varying these dimensions of a simulated campaign did not affect the manipulation probability (clicking on the phishing link despite knowledge of the simulated phishing campaign) in a statistically significant way. Based on our results, we advise organizations to prioritize obtaining informed consent from employees before participating in simulated phishing campaigns. The consequences of taking part in the simulated campaign should be clarified in advance. Furthermore, it's essential to carefully consider the pretext included in the phishing message to ensure they align with employee expectations and are acceptable to them.

[1]Pre-registration link: https://osf.io/vz9f2/?view_only=3f95e86f9a4743cba32c9877bb05338f

Contribution Statement. This paper makes the following contributions: (1) We explore the factors that influence the acceptance of simulated phishing campaigns using a vignette experiment, allowing us to make causal statements about factors influencing the acceptance and manipulation probability of simulated phishing campaigns. (2) We discuss the implications of these results for user-centered security research and cybersecurity practice and highlight how these could inform how organizations conduct anti-phishing training and future research.

II. Background and Related Work

A. Phishing and its Consequences

Phishing refers to the process in which sensitive information is elicited from the victim by pretending to be a trusted entity according to an automated pattern, typically via email [21]. In our connected world, phishing is an important concern in almost every company or government institution [22]. The frequency of phishing attacks is increasing, and their consequences are dire [23], [22]. Attackers predominantly conduct phishing campaigns using email but also use instant messaging or SMS [24].

Phishing attacks exploit human psychology to encourage potential victims to take certain actions. Attackers use a variety of tactics, such as threatening potential victims and time pressure [25]. These are often combined with tactics used in the field of Social Engineering [26], [27], such as distraction, authority and deception.

Various lenses have been applied to study phishing. A branch of studies aimed to link the success of phishing attacks to the personal characteristics of the targeted individual, but these approaches have been strongly criticized due to the absence of a solid psychological foundation [28].

Another set of studies examined the role of cognitive processes in phishing susceptibility. They suggest that phishing attacks activate the peripheral processing of information, thereby engaging the target user in lower levels of biased information processing. [29].

Phishing attacks can have severe consequences, including personal, financial, and societal harm. A notable example occurred in 2015 when spear phishing was used to target Ukraine's power supply, resulting in a six-hour power outage affecting around 80,000 people [4], [5]. Financial damage can also result from follow-up ransomware attacks, which demand payment for restoring access to stolen data. This time pressure can have devastating effects, such as disrupting production facilities or medical equipment in critical facilities, putting human lives at risk [3]. Successful phishing attacks can harm a company's reputation and customer trust [30].

B. Phishing Countermeasures

Companies and authorities attempt to implement effective countermeasures, which include intelligent anomaly detection through machine learning approaches, 2-factor authentication, or sandboxing [31]. But even with the combination of various

countermeasures, some risk remains, especially in organizations in which employees are expected to interact with actors outside of the organizations. Unfortunately, organizational vulnerability to phishing is difficult to mitigate completely with technical means [7]. There is a race between the defender and attacker: if the filtering rules for phishing attacks improve, the attackers adapt their emails to the target environment. Employees need to adapt flexibly. An increasing number of companies and government institutions focus on training people in addition to technical countermeasures. This is achieved both through training in general, but also simulated phishing campaigns.

C. Phishing Simulation Campaigns

Phishing simulation campaigns are similar to a real phishing attack, but unlike a real attack, the adversary is a team of offensive forces who are not real attackers. These try to attack the organizational infrastructure via emails tailored to the organization without causing sustainable damage. Volkamer et al. recommend defining this procedure in advance with leadership to assess an organization's vulnerability [32].

Simulated campaigns can also be used to formally evaluate the security awareness of an organization's employees [32]. The simulated campaign often also attempts to instruct and thus train the employees who have clicked on the link of a simulated phishing email [32].

It is unclear whether simulated phishing campaigns lead to the promised outcome of making organizations' employees adopt more secure email behaviors. Researchers have argued that simulated phishing campaigns do not have the intended effects [13], [14], [15]. Simulated phishing campaigns are often costly and the benefits as well as the approach are controversial. In their analysis of hidden costs for phishing simulation campaigns, Brunken et al. showed how extensive and underestimated these costs often are, especially due to the numerous personnel hours involved [33]. Negative consequences of phishing simulation campaigns can als result from employee reactions, as was the case for a simulated phishing campaign at US-based Tribune Publishing Company. After years of layoffs and wage cuts, a simulated phishing campaign was conducted and employees were lured by fake financial bonuses worth $5{,}000 to 10{,}000$ [20]. The deception led to public outrage and a decline in trust among employees and journalists, ultimately harming the company's reputation [20].

A study with over 6,000 employees found that simulated phishing campaigns can actually increase the risk of negative behavioral outcomes [11]. In fact, those who have already fallen victim to a phishing attempt are even more likely to be targeted again by new attacks. Distler conducted an in-situ deception study showing that simulated phishing campaigns can have undesirable consequences, such as shame within a person who interacted with a phishing email, which can lead to inaction after interacting with a phishing email [17]. One explanation here could be a finding by Volkamer et al., which explains that there is a possibility of resignation or loss of motivation on the part of employees if simulations occur too frequently, and even real phishing emails could be mistaken for a simulation, or employees could click on any link as a form of protest [32]. Mihelic et al. showed through a study with 111 subjects that employees lower their attention to a second phishing attack after an attempted one [34]. This behavior could be deliberately provoked to carry out more successful phishing attacks by distracting the employees [34]. Wood also explained the serious consequences of fraud that can be associated with falling for a phishing attack [35]. Psychological factors such as anxiety, depression, shame, disrupted sleep, or even an increased risk of suicide can be a possible negative consequence, which can also apply to simulated phishing campaigns [35].

The recording of click numbers, i.e., the number of people who clicked on the phishing link, is often used as a performance indicator and is intended to provide a company's management with information about the organization's security awareness. However, click numbers do not capture the circumstances of why people clicked [34]. Also, according to Volkamer et al., the completion of a training or a simulated campaign should not be a mere ticking off of a necessary task and then blaming the employee if they interact with a phishing attack despite having completed the training [32].

D. Acceptance of Phishing Simulation Campaigns

It is important that a company's employees and stakeholders find phishing simulation campaigns acceptable to avoid negative consequences such as loss of trust in the organizational leadership or IT, disengagement with future training measures. We define employees' acceptance of phishing campaigns as the approval of a certain implementation. There are different types of acceptance. Acceptance can be defined as a construct in which many factors play an important role, including the extent to which a new measure, like a simulated phishing campaign to reduce the click rates of phishing emails, is accepted, the importance for the user of this campaign, the individual usefulness, personal attitude towards the measure, the intention to change one's behavior as a result of the simulated campaign and the actual use following the measure [36].

Reed et al. conducted a survey on people's view on the efficacy and ethics of punishment. Participants thought that punishment should be reserved for more dangerous behaviors. They viewed punishment procedures as less effective than positive reinforcement [37]. The question remains whether the type of consequence can also play an important role in simulated phishing campaigns. Volkamer et al. relate consequences to simulated phishing campaigns and indicate that consequences should always be discussed transparently with employees and not be too strict. Otherwise, employees will not report when they have been victims of an attack for fear of consequences [32]. Jampen et al. also point out the importance of an anti-phishing campaign being adapted to the employees to avoid putting additional pressure on the employees, who might not want to "fail". Additional pressure on employees

can lead to their health and work performance suffering [38]. This illustrates the importance of employee acceptance for a successful campaign.

While the simulated phishing campaign can be viewed as a way to educate employees, certain factors in its implementation can lead to the adverse effects, even provoking behavior opposite to the one the intervention planned to implement. As shown in different organizational contexts, the security measures proposed by organizations can provoke instances of non-compliance [39] and even computer abuse behavior [40]. This can be explained within the framework of Reactance Theory: when people perceive a threat to their freedom of choice and actions, they often start to act against the threat to restore their freedom [41]. In contrast to avoidance behavior, reactance is an active negative response [42]. It has also been shown that in these contexts, people can even adopt the exact behavior against which the restrictive actions were directed [43]. However, Reactance Theory also postulates that not all freedom restrictions provoke the same level of reactance. In the context of organizational security, previous works mentioned several factors provoking the reactive response of employees: communication failure from the organization (bad explanation of security measures, so they can be perceived by the employees as threat to their freedom)[42], lack of perceived organizational justice (where employees feel that they are treated unfairly), and general distrust to the organization [40]. It is possible that some parameters of simulated phishing campaigns (such as the absence of consent) can be interpreted in frames of these factors and, therefore, provoke reactive behavior.

E. Summary

Phishing is a threat to almost every organization and government institution and exploits human psychology by appealing to authority and distracting recipients, among other strategies.

Simulated phishing campaigns are controversial, and both phishing and simulated phishing campaigns can lead to a loss of motivation and trust, but also psychological harm such as shame and fear.

Acceptance can influence the effects of simulated phishing campaigns, including effects on employees' views and trust in their employer and adverse behavioral outcomes (e.g., intentionally boycotting campaigns). We do not currently know which factors influence employee acceptance of simulated anti-phishing campaigns.

This paper investigates the factors that make a simulated phishing campaign seem (un-)acceptable via a vignette study.

III. Research Objectives

We address two main research questions:

RQ1 What factors affect the acceptance of a simulated phishing campaign in an organization?

RQ2 What factors affect the likelihood of the participants to click on the link contained in the phishing email despite knowing that it is a simulated phishing campaign?

a) Hypotheses: Most research in the field of usable privacy and security collects informed consent [44] and deception studies are only rarely conducted [45]. In organizational contexts, different legislative and ethical protections are in place than in empirical research. Prior research emphasizes the significance of psychological contracts for employees' acceptance of an organization's cybersecurity policies [46]. Asking employees for consent before launching the simulated phishing campaign can be considered part of transparent organizational communication. Previous studies have shown that transparent communication of organizational actions that are relevant to employees can positively affect employee engagement in the organization's life and work processes [47], and their willingness to accept organizational changes [48]; it is also described as a necessary step to assure the organizational decisions are understood and accepted by employees [49]. Further research shows that the perceived usefulness and credibility can increase acceptance of a security measure [50]. This can be realized in the context of a simulated phishing campaign so that the company not only informs people about it in advance but also obtains the consent of the affected persons [50]. This ensures that people develop an understanding of the simulated campaign and are more likely to engage with it than with a purely informative announcement or no information at all [50]. Without consent, organizations risk violating this psychological contract, potentially eliciting negative emotional and behavioral reactions from employees [51].

We hypothesize that prior informed and free consent to participating in a simulated phishing campaign would influence employee acceptance of such campaigns.

H1: Obtaining employee consent in advance has a positive effect on the acceptance of the simulated phishing campaign.

The example of a newspaper company conducting a simulated phishing campaign leading to a public outcry [20] provides a negative example of how monetary promises in the phishing email can lead to a severe loss of trust and lasting negative consequences for both the employees and the company if the employees have money problems at the time of the campaign. Here, negative employee sentiment was shown to increase aversion and lack of understanding of the campaign [20]. Strong incentives, such as money, can cause a person to perform the desired behavior (falling for a phishing attempt), which is why monetary incentives are often used in phishing attacks [52]. The feeling of being tricked can lead to negative emotional consequences for the victim of a simulated phishing attack and can thus lead to a strong aversion [53], [54], [35], [55]. Hence, we hypothesize that monetary incentives in messages in the context of a phishing campaign might lead to lower acceptance:

H2: The promise of a monetary incentive in phishing email content has a negative effect on the acceptance of the

simulated phishing campaign.

We examine the effects of various types of consequences in the context of simulated phishing campaigns. Prior work has shown that punishment is perceived as acceptable only when reserved for severe transgressions [37]. Weinzimmer and Esken also highlighted the importance of organizations tolerating mistakes without fear of repercussion to influence psychological safety, and positively influence organizational learning and performance [56]. Wang et al. [57] found that error tolerance in organizational settings was linked to psychological wellbeing, arguing that constructive error management is essential for employee psychological wellbeing. We assume that any consequence which takes time from employees away from their primary work tasks is perceived as negative and will lead to lowered acceptance of a phishing campaign.

We hypothesize:

H3: Consequences for the employee, resulting from clicking on the phishing link, have a negative effect on the acceptance of the simulated phishing campaign [2].

Volkamer et al. [32] have argued that employees might intentionally click a phishing link to protest a simulated phishing campaign. This argument is in line with several studies focused on reactance theory and its applications in explaining non-compliance [58] and computer abuse behavior [40] in organizational settings. Studies demonstrate that mistakes in implementing security measures (e.g., lack of transparency leading to the perception of organizational communication about security matters as "broken") can increase reactance levels and heightened non-compliance issues. Conversely, effective communication can be considered a factor that lowers reactance [59], [60]. Effectively communicating the upcoming simulated phishing campaign and obtaining employees' consent to participate could mitigate potential reactive impulses by showing that the organization clearly communicates information about the event and respects employees' freedom to participate or not. For this reason, we hypothesize:

H4: Obtaining employee consent in advance has a negative effect on the employees' intention to click on the phishing link despite knowledge of the simulated phishing campaign.

Monetary incentives have a long history of being the most commonly used phishing strategy, appearing even before the rise of current email-based communication [61]. From the user's perspective, they are also recognized as the most identifiable type of phishing attack, as users appear to be more cautious about emails mentioning money or banking alerts [62]. In an organizational context, it has been shown that one of the main factors affecting the evaluation of an email's legitimacy is the perceived likelihood of receiving such an email [62]. As monetary-incentive emails appear unusual in most organizational contexts, they have a higher chance of being perceived as obviously malicious or fake. In this case, we can expect the level of user frustration with these emails to be higher, and their presence can provoke more reactance-based actions similar to the reasons discussed in H4.

We hypothesize that promises of money in the phishing message trigger a higher dissonance and that employees feel tricked, which is why the following hypothesis is formulated:

H5: Campaigns in which a monetary incentive is promised have a positive effect on the intention to click on the phishing link despite knowledge of the simulated phishing campaign.

Prior work has recommended that consequences of simulated phishing campaigns should be discussed transparently with employees and should not be handled too strictly. Otherwise, employees will not report when they have been the victim of an attack for fear of serious consequences [32]. Coercive power from the authorities and organization's officials can trigger reactance [63], [40]. Measures relying on fear and severe consequences, such as tightening security policies and emphasizing the repercussions of non-compliance, can raise non-compliant behavior and provoke a reactive response [63], [40]; In a context beyond organizational security, it was shown that fear appeals strongly provoke psychological reactance [64].

Knowing the severe consequences of a phishing incident could reinforce the intention to protest. Thus, we hypothesize:

H6: More severe consequences for the employee resulting from clicking on the phishing link have a positive effect on the intention to click on the phishing link despite knowledge of the simulated phishing campaign.

Flores et al. [65] showed that individuals with computer experience exhibit higher phishing resilience. We hypothesize that higher IT affinity is associated to higher acceptance of phishing campaigns:

H7: Higher affinity for technology correlates with higher acceptance of the simulated phishing campaign.

IV. Methodology

A. Research Design

We conducted an online vignette experiment in July 2023 with a 2 4 2 (Consent Consequences Incentive) between-subjects design. The independent variables *Consent* (Yes vs. No), *Consequences* (No impact vs. Employee interview vs. Training vs. Termination after click on phishing email), and *Monetary Incentive* (Yes vs. No) were measured as independent within-subject factors. Participants were given the task of situating themselves in the role of a caseworker who receives a draft for a possible phishing campaign to be carried out in their company. The participant should make an assessment of the respective campaign on the basis of the information available by answering the questions in the questionnaire regarding the acceptance and manipulation probability. We systematically varied the vignettes with respect to

[2] In the pre-registration, our hypothesis was "*More severe organizational* consequences for the employee, resulting from clicking on the phishing link, have a negative effect on the acceptance of the simulated phishing campaign." This was misleading since we did not have a clear hypothesis regarding the order of severity of the consequences. We treat the different consequences as categorical variables. We thus adapted the hypothesis to reflect our view of consequences as categories.

the independent variables, resulting in 16 possible scenarios. Figure 1 shows the procedure. The study was approved by the ethics committee of the University of the Bundeswehr Munich in Germany .

Participants first provided informed consent to participate, followed by information about phishing. Participants were then presented with the background information needed and were then randomly assigned to *one* of the 16 vignettes. The between-subjects approach ensured that participants would not guess the exact purpose of the study and could not weigh between different scenarios. After reading through the vignette, participants were asked to assess (1) how likely they were to accept the scenario at hand and (2) how likely the participant would manipulate the scenario despite the knowledge that it was a simulated phishing campaign by their own organization. Last, we asked participants about their prior phishing campaign experience, IT affinity, and demographic variables. We provide the vignettes in Appendix A and the complete questionnaire as well as the entire data set and analysis syntax in the pre-registration[3].

a) Pre-tests: To ensure the understandability of the questionnaire, we first asked three experts in user-centered security and human-computer interaction to go through our questionnaire while thinking aloud while filling out the questionnaire. As a result, we refined our question items. We conducted a pre-test with N=35 subjects to detect possible comprehension problems, especially regarding the vignette scenarios. Based on the feedback from the pre-test, we were able to incorporate improvements that should increase the quality of the main study.

B. Vignettes

Each participant was exposed to one vignette out of 16 possible vignettes. Figure 2 shows the situation participants were asked to imagine themselves being in. Participants were asked to imagine that they were a caseworker in a company, asked to evaluate the design of a planned, simulated phishing campaign. The vignette first provided information on whether the employer in the scenario would obtain prior consent about the upcoming campaigns from each employee (*Consent: Yes vs. No*). The vignette then explained the content of the planned phishing email. The company's boss asked the employee to open the link and, depending on the scenario, promised a salary increase *(Monetary incentive: Yes vs. No)* if the information contained in the link was presented at the next meeting. Finally, the vignettes varied the consequences for employees if they fell for the phishing link *(Consequences: No impact vs. Employee interview vs. Training vs. Termination after clicking on a link in a phishing email)*.

C. Measurements

1) Acceptance (dependent variable): There is no one generally recommended way of measuring acceptance. A review from the field of driving automation found that there were eight major ways of measuring acceptance, which also varied depending on the study objective, and many studies used a one-item measure of acceptance [36]. Similarly, we evaluated acceptance of the vignette on a 10-point scale (1=not acceptable at all; 10=fully acceptable), asking "How acceptable would you find it if this campaign was conducted in this form in your company?".

2) Manipulation Probability (dependent variable): Manipulation probability was measured with the question "What is the likelihood that you would click on the phishing link if you already realized it was a phishing email from your employer?". Answers were recorded on a 10-point scale from 1=very unlikely to 10=very likely. In addition, we asked subjects to justify their decision regarding the likelihood of manipulation within an open response field to better understand the motivations for actively manipulating a simulated phishing campaign. This measure was purely exploratory as we did not find similar concepts being studied previously, but found an empirical investigation of the concept compelling.

3) Previous Experience of Phishing Campaigns: After the vignette experiment and independently of the condition a participant was assigned to, we asked the subjects whether they had already been part of a simulated phishing campaign as an employee in a company. This question was a filter question (answer format: yes/no). Participants with prior experience were then asked to describe the campaign in more detail. We chose an open response field and asked participants to provide information on the content, number of phishing emails, duration, and campaign scope.

We asked participants with prior experience whether clicking on the phishing link had any consequences and, if so, what these consequences were in an open-answer format. In addition, we surveyed whether and how the participants were informed about the campaign in advance and/or afterwards. Participants could choose from the following categories: Not at all (1), Verbally by the supervisor (2), Email (3), Work meeting (4), Training (5), Note during recruitment (6), Other (open response field; 7). Lastly, we recorded on a 6-point scale from 1=strongly disagree to 6=strongly agree how strongly participants agreed that the simulated phishing campaign improved the relationship between them and the employer and whether they rated phishing campaigns as very positive.

4) IT Affinity: We recorded the participants' technical affinity using the short version of the Affinity for Technology Interaction Scale (ATI Scale) according to Franke et al. [66].

D. Recruitment and Participants

We recruited a sample of N=793 subjects from Prolific[4] in July 2023. Platform members receive a notification when they are eligible for a research study. For our study, we did not use any restrictions regarding gender or education. To participate in the study, participants had to meet the requirement of using technology at work of "about once a week, 2 or 3 times a week, 4 or 6 times a week, about once a day, more than once

[3]Pre-registration link: https://osf.io/vz9f2/?view_only=3f95e86f9a4743cba32c9877bb05338f

[4]Prolific Platform: https://www.prolific.co/

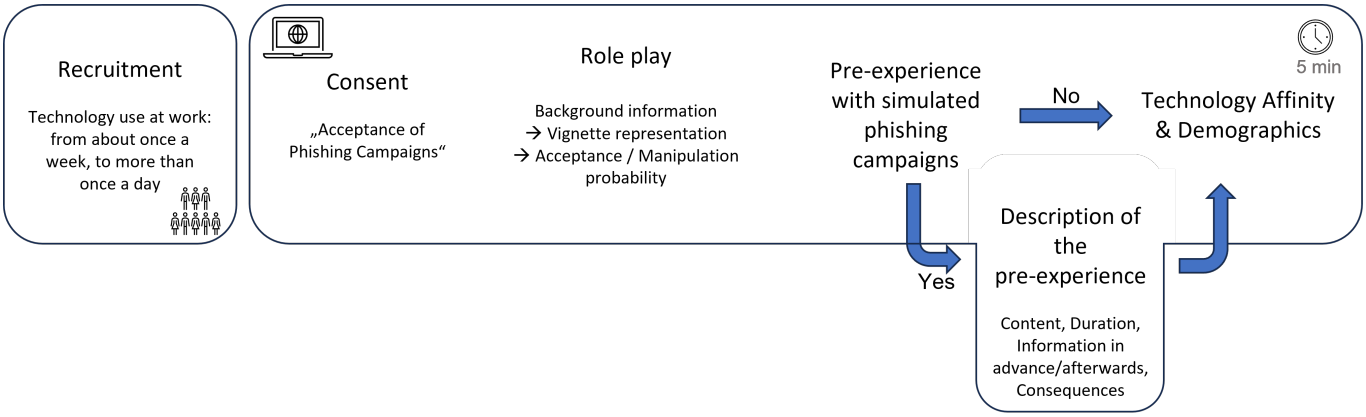

Fig. 1. An overview of the study procedure. Participants were recruited on an online platform and used technology at work at least once per week. Participants were asked to situate themselves in a role play, where they were first given information about phishing, and were then randomly shown one of 16 vignettes describing a simulated phishing campaign scenario. Participants answered questions about how acceptable they found the scenario and how likely the participant would manipulate the scenario despite the knowledge that it was a simulated phishing campaign by their own organization (manipulation probability).

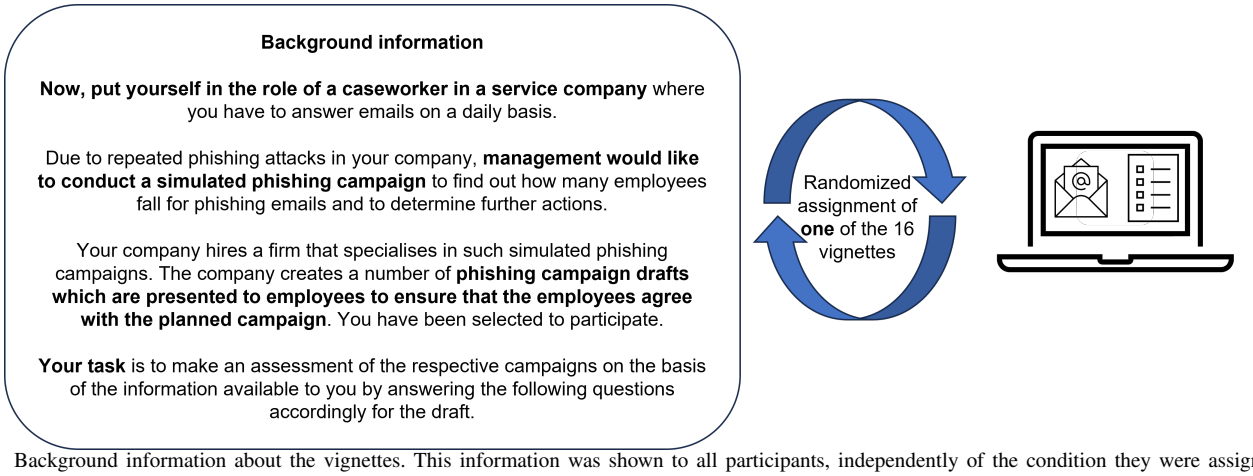

Fig. 2. Background information about the vignettes. This information was shown to all participants, independently of the condition they were assigned to. After this background information, participants were shown one of 16 vignettes.

a day". In addition, participants who had taken part in the pre-test of the study were excluded. Participants had to be UK residents. Our goal was to recruit 800 participants to obtain approximately 50 responses per vignette (50*16 = 800), as 50 responses per vignette are recommended in the literature as a rule of thumb to obtain sufficient statistical power [67].

1) Data Exclusion: We obtained data from 803 participants who completed the questionnaire in full. In line with our pre-registration, we excluded 10 participants because they reported an English language level of A1 or A2. In the analyses, only data from the study participants who had at least an advanced level of language proficiency at the time of the study would be analyzed (B1 or better). This is to ensure that the participants understand the vignettes, despite the complexity of the subject matter. A total sample size of N=793 was used in the analysis.

2) Description of Sample: Participants were 48.7% female, 50.2% male, 0.6% non-binary, and 0.5% did not indicate their gender. Participants were on average M=41 years old (SD=12.85, Range=18-78). Participants were relatively highly educated, with many holding bachelor's or master's degrees. Participants showed an average ATI score of technology affinity of M=14.62 (SD=4.59). The Cronbach's alpha is =.87.

E. Experimental Data

Each vignette was shown 49.56 times on average. For gender analysis, a t-test was calculated between the dependent and demographic variables. Spearman correlation was calculated for age and the Kruskal-Wallis tests were calculated for education to detect differences between groups. We found that males (M=5.76, SD=3.37) reported significantly higher acceptance scale scores than females (M=5.01, SD=3.28; t(782)=-3.20, p<.001, 95% CI [-1.22, -0.29]). With regard to age (p=.50), no significant relationship was found on vignette acceptance, nor were there significant differences with regard to education (p=.35) and acceptance. No significant gender difference was found on the manipulation probability scale (t(782)= -.93, p=.35, 95% CI [-0.15, 0.43]). We found no significant correlations between age (p=.46) and education (p=.31).

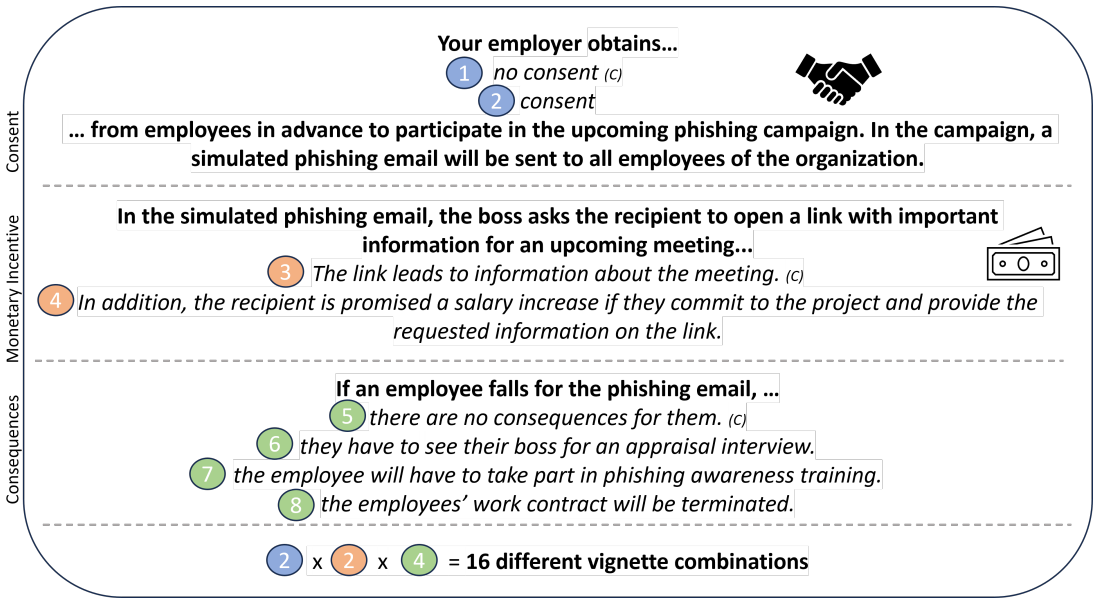

Fig. 3. Representation of the scenario described in the vignettes. The dimensions (consent, monetary incentive, consequences) are separated by a dashed line. The levels within each dimension are numbered. Baseline conditions are highlighted with a (C).

We inspected the response behavior across the vignettes. We found a symmetrical data distribution of the responses on the acceptance scale, where the extremes have the highest frequencies and the middle category values have lower frequencies, resulting in a U-shaped response distribution of acceptance (Appendix B.1). All response options were used. The manipulation probability scale shows a right-skewed distribution, with most responses being low values and a very small portion indicating higher values. All response options were used on this scale (Appendix B.2).

The Shaprio-Wilk test showed data not to be normally distributed. Schmidt and Finan's [68] work shows violations of the normal distribution to not noticeably affect the results for large samples (> 10 observations per variable).

F. Data Analysis

In separate models, we first estimated the overall effect of our independent variables (consent, monetary incentive, consequences) on the dependent variables' acceptance of the simulated phishing campaign and the manipulation probability. Then, we estimate the effect of the individual characteristics of the independent variables on the dependent variables. If we found a significant effect, we examined between which variable expression the effect is to be found.

In addition, we examined the relationship between individual affinity for technology and the acceptance of simulated phishing campaigns calculating the individual sum score for each person.

V. Results

A. Bivariate Correlations between Dependent Variables

We performed a correlation analysis of the dependent variables' acceptance and manipulation probability. We found a significant negative correlation (r=-.08, p=.02). This means that higher values in acceptance could tend to be associated with slightly lower values in manipulation probability. However, since the correlation coefficient is close to zero, this indicates that the relationship between these two variables is rather weak. The acceptance rating of the different vignettes was on average M=5.39 (SD=3.34) and the manipulation probability was M=2.12 (SD=2.06).

B. Experimental Evidence

1) Acceptance: We calculated a linear regression between the dependent variable acceptance and the independent variables using a significance level of =.05. To evaluate the hypotheses, we refer to the results of the single effects. However, the overall effects are also reported (table T.3 and figure B.3). We found a significant positive effect of obtaining prior consent from employees at the beginning of a phishing campaign on acceptance (r(789)=.28, p<.001). With regard to consequences, a significant negative effect was found on the acceptance (r(789)=.32, p<.001). We found no effect of the monetary incentive on acceptance in the overall effects (r(789)=-.05, p=.07). The correlation table of the independent variables can be found in Appendix B.

To further examine the results, we then calculated the single effects of the levels of the independent variables (see Table I). We found a statistically significant positive effect on acceptance (p<.001). Obtaining consent was found to increase the acceptability rating by almost one scale point (0.90, p<.001). The single effects of the consequences also showed a significant effect of the variable characteristics, whereby the consequence of an employee interview as a result of clicking on the phishing link caused the acceptance of the campaign to drop by more than one and a half scale points (-1.64, p<.001).

Termination of employment led to a significant drop in the acceptance rating of almost three and a half scale points (-3.37, p<.001). We found that a monetary incentive caused the acceptance of the phishing campaign to drop by almost half a scale point (-.47, p<.02). We did not obtain a significant result with regard to the acceptance rating of the training as a consequence of the simulated phishing campaign (.29, p<.32). Figure 4 shows a visual representation of the coefficients.

TABLE I
SINGLE EFFECTS OF THE INDEPENDENT VARIABLES ON THE ACCEPTANCE OF THE SIMULATED PHISHING CAMPAIGN. ACCEPTANCE WAS MEASURED ON A SCALE OF 1 TO 10 (1=NOT ACCEPTABLE AT ALL; 10=FULLY ACCEPTABLE).

Term	Estimate	SE	p-value
Intercept	6.4713***	0.2520	
Consent	0.9034***	0.2379	< .001
Incentive	-0.4741*	0.2075	0.023
Employee interview	-1.6394***	0.2845	< .001
Training	0.2923	0.2959	0.324
Termination of contract	-3.3688***	0.3121	< .001

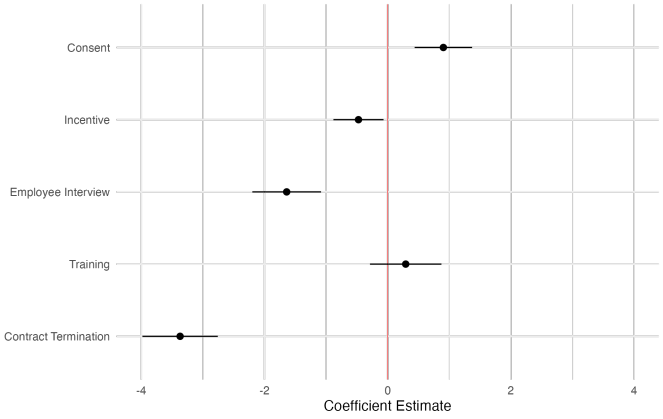

Fig. 4. Coefficient plot single effects of the independent variables on the acceptance of the simulated phishing campaign

2) Manipulation Probability: Overall, participants reported a very low likelihood of clicking on a phishing link if they already realized it was from their employer (see Appendix Figure B.2). We estimated the overall effect of the independent variables on the dependent variable at a significance level of =.05. No statistically significant relationship was found between the Consent, Incentive, and Consequences variables and the dependent variable (the overall regression model for manipulation probability can be found in the Appendix, Table T.4 and Figure B.4).

Nevertheless, we then looked at the single effects (see Figure 5) of the variable expressions, showing that obtaining consent and a monetary incentive in the context of the phishing email had no statistically significant effects on manipulation probability (p>.05).

In order to evaluate the open answers, we categorized the participants' answers. In this way, each answer could eventually be assigned to at least one category. The answers

TABLE II
SINGLE EFFECTS OF THE INDEPENDENT VARIABLES ON THE MANIPULATION PROBABILITY. MANIPULATION PROBABILITY WAS MEASURED ON A SCALE OF 1 TO 10 (1=VERY UNLIKELY, 10=VERY LIKELY).

Term	Estimate	SE	p-value
(Intercept)	1.9173	0.1784	
Consent	0.1116	0.1684	0.508
Incentive	0.1648	0.1469	0.262
Employee Interview	0.0714	0.2014	0.723
Training	0.2832	0.2095	0.177
Termination of contract	0.0196	0.2210	0.929

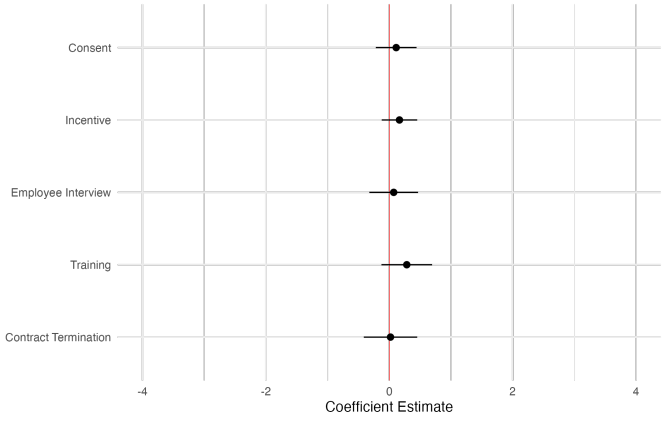

Fig. 5. Coefficient plot of single effects of the independent variables on manipulation probability

to the open question, in which participants were asked to explain reasons why they might click on a phishing link if they already know that it is a simulated phishing campaign, indicate that there were three main reasons for intentionally clicking a suspected phishing link from an employer: false trust in the email (n=44), protest (n=11) and curiosity (n=9). Answers such as "The boss has specifically asked me to open it, so I would think it is OK" (P399) or "I think I'm quite trusting and do as I'm told" (P47) showed that participants trusted phishing emails if they assume that it is a legitimate email from their boss. Protest was a reason for intentionally clicking on phishing links. For example, participants stated "they shouldn't be allowed to do it. I don't think it's morally right" (P639) or "I would still click on the link because I know there is no consequence for me"(P13). Answers such as "Just to read it" (P134) or "Just out of curiosity I guess.[...]" (P665) show that phishing messages can trigger curiosity in recipients leading to them interacting with the simulated phishing email.

3) ATI and Acceptance: Participants had an average ATI value of M=14.62 (SD=4.59). To investigate whether technology affinity influences the acceptance of phishing campaigns, we calculated a Spearman correlation due to the violation of the normality assumption. A significant positive correlation was found (r(791)=.12, p<.001). This suggests that people with a higher affinity for technology generally rate the acceptance of phishing campaigns higher. Hypothesis 7 can be accepted.

C. Prior Experience with Simulated Phishing Campaigns

Of the 793 participants, n=179 (23%) indicated prior experience with simulated phishing campaigns and completed the additional questions describing the prior experience. This subsection refers to the answers of this subset of participants. The purpose of this subsection is descriptive and independent of the experimental treatments. 64% of participants who had previous experience with simulated phishing campaigns were not informed about the campaign in advance. The participants were asked how they had been informed about the phishing campaign in advance. 17% reported being informed by email, 10% as part of training, 7% by the supervisor, and just under 3% each by a work meeting or during the application process. Regarding the announcement afterwards, 77% said they had been informed about the phishing campaign by email, 10% each in the context of training and/or during a work meeting, 8% verbally by the supervisor, and 5% not at all. A small number of the participants indicated, for example, an information message on the company's intranet site or conversations with the other employees.

44% of participants with prior experience indicated that falling for the phishing message resulted in consequences for that particular employee. Almost exclusively, participants indicated in the open response field the consequence was anti-phishing training.

D. Summary of the Results

We summarize the results of this study in Table III.

VI. Discussion

A. Limitations

We could not test every possible scenario or combination possible in reality. Thus, not every possible expression of the independent variables could be represented. However, using our methodological approach, we were able to isolate the effects of our variables and make a statement about the effects of the variables we collected. We acknowledge that the consequence "training" might be perceived differently by different employees. In future work, it would be insightful to provide further details regarding various consequences to understand their relative acceptance. Our sample consists exclusively of UK residents. There may also be cross-cultural differences with regard to the dependent variables of acceptance and manipulation probability. Lastly, vignettes did not have the same length, some vignettes were longer than others. However, each subject was presented with only one vignette, so the average duration of the study was 04:54 minutes. We assume that fatigue effects did not play a role.

B. Acceptance

a) Consent and simulated phishing campaigns: Consent and simulated phishing campaigns are a debated topic. In our study, consent positively influenced the acceptance rating, but in practice, simulated phishing campaigns typically involve deception [32]. For example, some organizations never disclose that a simulated campaign has been conducted and simply redirect the victim to a legitimate website. Others inform the victim once fallen for a simulated phish [32].

Securing consent before initiating simulated phishing training can be understood through the lens of the "psychological contract". The psychological contract outlines the implicit expectations between employees and employers regarding mutual responsibilities [69]. Prior research emphasizes the significance of psychological contracts for employees' acceptance of an organization's cybersecurity policies [46]. Studies also indicate that employees perceive justice and fairness as core components of these psychological contracts [70]. Essentially, employees expect organizations to act transparently and declare their intentions openly. Without clear consent for simulated phishing emails, organizations risk violating this psychological contract. Such violations can elicit negative emotional and behavioral reactions from employees [51], potentially diminishing their commitment to the organization's security measures [71].

Reactance theory offers an alternative perspective on the importance of consent in simulated phishing campaign acceptance. As described previously, reactance is described as a negative emotional response triggered by perceived threats or limitations on an individual's behavioral freedom [41]. Within organizations, this type of response frequently emerges in relation to security measures that aim to regulate employee behavior [72]. Our results confirmed that when an organization initiates a simulated phishing campaign without obtaining employee consent, it can be perceived as restricting their freedom to participate. This might provoke negative reactions, such as employees deliberately clicking on phishing links, which counters the organization's objectives.

Informed consent is considered an important ethical safeguard in most empirical studies in usable privacy and security [44], but the amount of information that should be provided to participants in research to qualify as informed consent is often unclear [73]. Lengthy documents are not necessarily informative for research participants or employees. Prospective research participants may not fully understand the information disclosed in the informed consent process [74], and similar issues could arise with employees. More investigation is needed to understand how employees might best be informed about simulated phishing campaigns in ways that both respect their time and provide all necessary information. Simulated phishing campaigns generate personal and potentially sensitive information about employees. An informed consent procedure should, at the very least, clarify who would gain access to information about employees' behavior, how long this information will be stored, how it will be secured, and how consent can be revoked (additional considerations can be found in [32]). Consent should be informed and freely given. For example, following this standard, employees who do not give consent to participate in a simulated phishing campaign should be given other options to learn about phishing countermeasures. Any security measures could also be conceptualized and refined in co-design sessions with employees, considering their thoughts and experiences in their daily work lives.

TABLE III
OVERVIEW OF THE RESULTS

	Hypothesis	Result	Explanation
1	Obtaining employee consent in advance has a positive effect on the acceptance of the simulated phishing campaign.	Confirmed	Prior consent had a positive effect on the acceptance rating of the phishing campaign.
2	The promise of a monetary incentive in phishing email content has a negative effect on the acceptance of the simulated phishing campaign.	Confirmed	The presence of a monetary incentive had a negative effect on acceptance.
3	More severe organizational consequences for the employee resulting from clicking on the phishing link have a negative effect on the acceptance of the simulated phishing campaign.	Partially confirmed	Training had statistically non-significant effect on acceptance. An employee interview or termination of the employment relationship had a statistically significant negative effect on acceptance.
4	Obtaining employee consent in advance has a negative effect on the employees' intention to click on the phishing link despite knowledge of the simulated phishing campaign.	Not confirmed	No statistically significant effect of prior consent on manipulation probability could be found.
5	Campaigns in which a monetary incentive is promised have a positive effect on the intention to click on the phishing link despite knowledge of the simulated phishing campaign.	Not confirmed	No statistically significant effect of monetary incentive on manipulation probability.
6	More severe consequences for the employee resulting from clicking on the phishing link have a positive effect on the intention to click on the phishing link despite knowledge of the simulated phishing campaign.	Not confirmed	No statistically significant effect of stronger consequences on manipulation probability could be found.
7	Higher affinity for technology correlates with higher acceptance of the simulated phishing campaign.	Confirmed	People with a higher IT affinity rated the acceptance of phishing campaigns higher.

b) False promises of monetary incentives: We found that the content of the email containing the promise of a monetary incentive had a small yet statistically significant negative effect. In a real-life phishing campaign, the effect of a promised incentive will depend on the organizational context, and it is important to consider that other pretexts may also have a negative effect on acceptance. Sensitive topics might include, in addition to financial incentives, vacation days, sick leave, organizational restructuring, politics.

c) Consequences of interacting with a phishing email: Depending on the consequences described in the vignette scenario, acceptance ratings differed (see Figure 4). While both contract termination (in line with [37]) and an employee interview had a negative effect, training, as a consequence, had a non-statistically significant negative effect. We hypothesize that the effect of training on acceptance will depend on a variety of factors, including the duration of the training measure, whether the training is perceived as "embarrassing" (e.g., if supervisors are informed) or as helpful.

C. Manipulation Probability

Previous work has mentioned the possibility of employees manipulating intentional clicking as a protest because they feel it is unreasonable for their organization to "trick" them in this way or out of curiosity [32]. In our study, the majority participants indicated that they would not knowingly click on a simulated phishing email from their employer, and none of the vignette factors showed a statistically significant effect on manipulation probability. In the open-ended answers, we did find some indication of clicking out of curiosity or because the participant would not expect any real consequences from clicking. There are multiple possible follow-up hypotheses. The intention to manipulate a simulated phishing campaign might not be very common in general, which explains why we rarely observed it in our sample. It is also possible that the intention to manipulate a simulated phishing campaign is bound to the real-life context of an organizational simulated phishing campaign and can not easily be replicated using a vignette scenario. For instance, [75] point to the tensions that arise in organizational contexts when time constraints, resource constraints, cognitive constraints, and incomplete information collide with contradictory approaches to secure behaviors in organizations, which can lead to employees making "good enough" decisions. Indeed, time, resource and time constraints cannot easily be replicated outside a realistic work context, and it is possible that clicking on a phish which somebody suspects to come from their employer is a phenomenon that appears only in the presence of such real-world tensions. There might also be a social desirability effect discouraging research participants from reporting what they may perceive as anti-social behavior in an attempt to present themselves in a positive light [76].

D. IT Affinity

Finally, the study focused on whether a higher IT affinity leads to a higher acceptance of phishing campaigns. Similarly,

Flores et al. [65] found that individuals with computer experience have a higher phishing resilience.

VII. Recommendations

Based on our findings, we offer practical recommendations for future simulated phishing campaigns. Note that our study does not assess whether these campaigns lead to improved security outcomes, but rather focus on understanding factors contributing to their acceptance.

Obtain consent from participants before including them in simulated phishing training. Our study points to a positive effect of obtaining employee consent before simulated phishing campaigns on acceptance. Since they have been warned, there is a trade-off between obtaining consent and potentially influencing employees' future behavior. However, the objective of any security measure must be to keep up long-term engagement with security measures, as well as trust in the security professionals of a company. Thus, it seems worthwhile to conduct simulated phishing campaigns after obtaining employee consent and to carefully measure the effects of such prior consent. Employees who do not provide informed consent should be allowed to participate in other types of security training. The knowledge of being part of a simulated phishing campaign might only influence behavior more in the short term.

Clarify the consequences of insecure behaviors before a simulated phishing campaign. Consequences (positive and negative) should be defined in collaboration with employees of an organization. We found that the consequences of a phishing campaign influence its acceptance. Combined with prior informed consent, we recommend clarifying the consequences of simulated anti-phishing campaigns. Organizations should clarify the simulated phishing campaign's intentions, which consequences can arise for employees (positive or negative), and exactly how their data is used (e.g., who can access it, is it used to evaluate performance). We found that employee interviews lead to a lower acceptance of the simulated phishing campaign, which might be relevant for smaller organizations in which such a measure might be more realistic. There is also no evidence that such an employee interview would positively affect phishing behavior.

Organizations should carefully consider whether certain pretexts are acceptable for their employees. Promising an incentive in the email had a negative effect on the acceptance of the campaign. While attackers might use any means necessary to trick victims, organizations should ensure employees' long-term commitment to security rather than tricking employees.

VIII. Future Research

Future work should investigate the real-life acceptance of employees who have been included in simulated phishing campaigns in their organizations. It would be especially insightful to interview employees who did not agree with the simulated phishing campaign to identify possible improvements. In addition, it would be relevant to understand how employees think these campaigns affected their behavior. Behavior change in organizational contexts is complex and multi-faceted. Future research could go into more detail regarding the characteristics of the simulated phishing campaign.

IX. Conclusion

Our results demonstrate the effects of varying certain factors (consent, monetary incentive, consequences) when designing simulated phishing campaigns. We found that the examined factors can have different influences on employee acceptance. Note that this study does not investigate or take a stance on the effectiveness of such campaigns to increase organizational security, which multiple studies have questioned. We hope that future work will investigate both how anti-phishing training (simulated or other) can be made more effective, as well as accepted by employees. Organizations depend on the long-term collaboration and motivation of their employees to stay safe from outside threats, and any security measure should be evaluated in terms of both behavioral effects and how acceptable the measure is perceived by employees. We hope to see more research investigating employee engagement, motivation, and acceptance of security measures.

Acknowledgements

This study is part of the project "Voice of Wisdom", funded by dtec.bw – Digitalization and Technology Research Center of the Bundeswehr. dtec.bw is funded by the European Union – NextGenerationEU.

References

[1] R. Wash, "How experts detect phishing scam emails," vol. 4, place: New York, NY, USA Publisher: Association for Computing Machinery. [Online]. Available: https://doi-org.proxy.bnl.lu/10.1145/3415231

[2] K. L. Chiew, K. S. C. Yong, and C. L. Tan, "A survey of phishing attacks: Their types, vectors and technical approaches," *Expert Systems with Applications*, vol. 106, pp. 1–20, 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0957417418302070

[3] M. House, "Attributing deaths to ransomware attacks on hospitals and medical care facilities," 2021, https://www.cyber.forum.yale.edu/blog/2021/7/20/attributing-deaths-to-ransomware-attacks-on-hospitals-and-medical-care-facilities.

[4] M. Choraś, R. Kozik, A. Flizikowski, W. Hołubowicz, and R. Renk, "Cyber threats impacting critical infrastructures," *Managing the complexity of critical infrastructures: A modelling and simulation approach*, pp. 139–161, 2016.

[5] G. Liang, S. R. Weller, J. Zhao, F. Luo, and Z. Y. Dong, "The 2015 ukraine blackout: Implications for false data injection attacks," *IEEE transactions on power systems*, vol. 32, no. 4, pp. 3317–3318, 2016.

[6] C. Wilson and D. Argles, "The fight against phishing: Technology, the end user and legislation," 06 2011, pp. 501–504.

[7] A. M. Shabut, K. T. Lwin, and M. A. Hossain, "Cyber attacks, countermeasures, and protection schemes — a state of the art survey," in *2016 10th International Conference on Software, Knowledge, Information Management Applications (SKIMA)*, 2016, pp. 37–44.

[8] D. D. Caputo, S. L. Pfleeger, J. D. Freeman, and M. E. Johnson, "Going spear phishing: Exploring embedded training and awareness," *IEEE security & privacy*, vol. 12, no. 1, pp. 28–38, 2013.

[9] F. L. Ballreich, M. Volkamer, D. Müllmann, B. M. Berens, E. M. Häußler, and K. V. Renaud, "Encouraging organisational information security incident reporting," in *Proceedings of the 2023 European Symposium on Usable Security*, ser. EuroUSEC '23. New York, NY, USA: Association for Computing Machinery, 2023, p. 224–236. [Online]. Available: https://doi-org.proxy.bnl.lu/10.1145/3617072.3617098

[10] R. C. Dodge Jr, C. Carver, and A. J. Ferguson, "Phishing for user security awareness," *computers & security*, vol. 26, no. 1, pp. 73–80, 2007.

[11] F. L. Greitzer, W. Li, K. B. Laskey, J. Lee, and J. Purl, "Experimental investigation of technical and human factors related to phishing susceptibility," *ACM Transactions on Social Computing*, vol. 4, no. 2, pp. 1–48, 2021.

[12] G. J. Homsma, C. Van Dyck, D. De Gilder, P. L. Koopman, and T. Elfring, "Learning from error: The influence of error incident characteristics," *Journal of Business Research*, vol. 62, no. 1, pp. 115–122, 2009.

[13] M. A. Sasse, J. Hielscher, and M. Gutfleisch, "Human-Centred Security: Unfug Informationssicherheits-Sensibilisierung," *kma - Klinik Management aktuell*, vol. 27, no. 04, pp. 44–46, Aug. 2022, 44.

[14] M. Volkamer, M. A. Sasse, and F. Boehm, "Phishing-Kampagnen zur Steigerung der Mitarbeiter-Awareness: Analyse aus verschiedenen Blickwinkeln – Security, Recht und Faktor Mensch," *Datenschutz und Datensicherheit - DuD*, vol. 44, no. 8, pp. 518–521, Aug. 2020. [Online]. Available: https://link.springer.com/10.1007/s11623-020-1317-x

[15] D. Lain, K. Kostiainen, and S. Čapkun, "Phishing in organizations: Findings from a large-scale and long-term study," in *2022 IEEE Symposium on Security and Privacy (SP)*. IEEE, 2022, pp. 842–859.

[16] I. Kirlappos, S. Parkin, and M. A. Sasse, "Learning from "shadow security": Why understanding non-compliance provides the basis for effective security," 2014.

[17] V. Distler, "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*, 2023, pp. 1–18.

[18] X. Chen, S. Doublet, A. Sergeeva, G. Lenzini, V. Koenig, and V. Distler, "What Motivates and Discourages Employees in Phishing Interventions: An Exploration of Expectancy-Value Theory," Philadelphia, PA, 2024.

[19] X. Chen, M. Sacre, G. Lenzini, S. Greiff, A. Sergeeva, and V. Distler, "The Effects of Group Discussion and Role-playing Training on Self-efficacy, Support-seeking, and Reporting Phishing Emails: Evidence from a Mixed-design Experiment," in *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*. Honolulu HI USA: ACM, 2024.

[20] J. Barr, "The company email promised bonuses. it was a hoax — and tribune publishing employees are furious." 2020. [Online]. Available: https://www.washingtonpost.com/media/2020/09/23/tribune-bonus-email-phishing-hoax/

[21] Z. Alkhalil, C. Hewage, L. Nawaf, and I. Khan, "Phishing attacks: A recent comprehensive study and a new anatomy," *Frontiers in Computer Science*, vol. 3, p. 563060, 2021.

[22] P. Sharma, B. Dash, and M. F. Ansari, "Anti-phishing techniques–a review of cyber defense mechanisms," *International Journal of Advanced Research in Computer and Communication Engineering ISO*, vol. 3297, p. 2007, 2022.

[23] T. APWG, "Phishing activity trends reports," 2022.

[24] P. Gupta, B. Srinivasan, V. Balasubramaniyan, and M. Ahamad, "Phoneypot: Data-driven understanding of telephony threats." in *NDSS*, vol. 107, 2015, p. 108.

[25] M. Jari, "An overview of phishing victimization: Human factors, training and the role of emotions," *arXiv preprint arXiv:2209.11197*, 2022.

[26] A. Ferreira, L. Coventry, and G. Lenzini, "Principles of persuasion in social engineering and their use in phishing," in *Human Aspects of Information Security, Privacy, and Trust: Third International Conference, HAS 2015, Held as Part of HCI International 2015, Los Angeles, CA, USA, August 2-7, 2015. Proceedings 3*. Springer, 2015, pp. 36–47.

[27] T. Stojnic, D. Vatsalan, and N. A. Arachchilage, "Phishing email strategies: understanding cybercriminals' strategies of crafting phishing emails," *Security and privacy*, vol. 4, no. 5, p. e165, 2021.

[28] P. López-Aguilar, C. Patsakis, and A. Solanas, "The role of extraversion in phishing victimisation: A systematic literature review," in *2022 APWG Symposium on Electronic Crime Research (eCrime)*. IEEE, 2022, pp. 1–10.

[29] P. M. Musuva, K. W. Getao, and C. K. Chepken, "A new approach to modelling the effects of cognitive processing and threat detection on phishing susceptibility," *Computers in Human Behavior*, vol. 94, pp. 154–175, 2019.

[30] J. W. Ragucci and S. A. Robila, "Societal aspects of phishing," in *2006 IEEE International Symposium on Technology and Society*, 2006, pp. 1–5.

[31] S. Purkait, "Phishing counter measures and their effectiveness–literature review," *Information Management & Computer Security*, vol. 20, no. 5, pp. 382–420, 2012.

[32] M. Volkamer, M. A. Sasse, and F. Boehm, "Analysing simulated phishing campaigns for staff," in *Computer Security: ESORICS 2020 International Workshops, DETIPS, DeSECSys, MPS, and SPOSE, Guildford, UK, September 17–18, 2020, Revised Selected Papers 25*. Springer, 2020, pp. 312–328.

[33] L. Brunken, A. Buckmann, J. Hielscher, and M. A. Sasse, ""To do this properly, you need more Resources": The hidden costs of introducing simulated phishing campaigns," in *32nd USENIX Security Symposium (USENIX Security 23)*. Anaheim, CA: USENIX Association, Aug. 2023, pp. 4105–4122. [Online]. Available: https://www.usenix.org/conference/usenixsecurity23/presentation/brunken

[34] A. Mihelič, M. Jevšček, S. Vrhovec, and I. Bernik, "Testing the human backdoor: Organizational response to a phishing campaign," *Journal of Universal Computer Science*, vol. 25, no. 11, pp. 1458–1477, 2019.

[35] S. Wood, "How does fraud impact emotional well-being?" *Psychology Today*, 2021. [Online]. Available: https://www.psychologytoday.com/us/blog/the-fraud-crisis/202101/how-does-fraud-impact-emotional-well-being

[36] E. Adell, V. András, and L. Nilsson, "Definition of acceptance and acceptability," in *Handbook of Research on Advanced Concepts in E-Collaboration*, P. Zhang and R. T. Watson, Eds. Taylor & Francis, 2012, ch. 2, pp. 15–30. [Online]. Available: https://www.taylorfrancis.com/chapters/edit/10.1201/9781315578132-2/definition-acceptance-acceptability-emeli-adell-andr%C3%A1s-v%C3%A1rhelyi-lena-nilsson

[37] F. D. D. Reed and B. J. Lovett, "Views on the efficacy and ethics of punishment: Results from a national survey." *International Journal of Behavioral Consultation and Therapy*, vol. 4, no. 1, p. 61, 2007.

[38] D. Jampen, G. Gür, T. Sutter, and B. Tellenbach, "Don't click: towards an effective anti-phishing training. a comparative literature review," *Human-centric Computing and Information Sciences*, vol. 10, no. 1, pp. 1–41, 2020.

[39] A. Hovav and F. F. Putri, "This is my device! why should i follow your rules? employees' compliance with byod security policy," *Pervasive and Mobile Computing*, vol. 32, pp. 35–49, 2016.

[40] P. B. Lowry, C. Posey, R. B. J. Bennett, and T. L. Roberts, "Leveraging fairness and reactance theories to deter reactive computer abuse following enhanced organisational information security policies: An empirical study of the influence of counterfactual reasoning and organisational trust," *Information Systems Journal*, vol. 25, no. 3, pp. 193–273, 2015.

[41] S. S. Brehm and J. W. Brehm, *Psychological reactance: A theory of freedom and control*. Academic Press, 2013.

[42] P. B. Lowry and G. D. Moody, "Explaining opposing compliance motivations towards organizational information security policies," in *2013 46th Hawaii International Conference on System Sciences*. IEEE, 2013, pp. 2998–3007.

[43] S. Byrne and P. S. Hart, "The boomerang effect a synthesis of findings and a preliminary theoretical framework," *Annals of the International Communication Association*, vol. 33, no. 1, pp. 3–37, 2009.

[44] V. Distler, M. Fassl, H. Habib, K. Krombholz, G. Lenzini, C. Lallemand, V. Koenig, and L. F. Cranor, *Empirical Research Methods in Usable Privacy and Security*. Cham: Springer International Publishing, 2023, pp. 29–53. [Online]. Available: https://doi.org/10.1007/978-3-031-28643-8_3

[45] V. Distler, M. Fassl, H. Habib, K. Krombholz, G. Lenzini, C. Lallemand, L. F. Cranor, and V. Koenig, "A systematic literature review of empirical methods and risk representation in usable privacy and security research," *ACM Trans. Comput.-Hum. Interact.*, vol. 28, no. 6, dec 2021. [Online]. Available: https://doi-org.proxy.bnl.lu/10.1145/3469845

[46] J. Han, Y. J. Kim, and H. Kim, "An integrative model of information security policy compliance with psychological contract: Examining a bilateral perspective," *Computers & Security*, vol. 66, pp. 52–65, 2017.

[47] H. Jiang and R. L. Men, "Creating an engaged workforce: The impact of authentic leadership, transparent organizational communication, and work-life enrichment," *Communication research*, vol. 44, no. 2, pp. 225–243, 2017.

[48] V. D. Miller, J. R. Johnson, and J. Grau, "Antecedents to willingness to participate in a planned organizational change," 1994.

[49] K. Zorlu and F. Korkmaz, "Organizational communication as an effective communication strategy in organizations and the role of the leader," in *Management Strategies to Survive in a Competitive Environment: How to Improve Company Performance*. Springer, 2021, pp. 305–320.

[50] C. M. Jones, *Utilizing the technology acceptance model to assess employee adoption of information systems security measures*. Nova Southeastern University, 2009.

[51] E. W. Morrison and S. L. Robinson, "When employees feel betrayed: A model of how psychological contract violation develops," *Academy of management Review*, vol. 22, no. 1, pp. 226–256, 1997.

[52] S. Goel, K. J. Williams, J. Huang, and M. Warkentin, "Can financial incentives help with the struggle for security policy compliance?" *Information & management*, vol. 58, no. 4, p. 103447, 2021.

[53] L. Ganzini, B. McFarland, and J. Bloom, "Victims of fraud: Comparing victims of white collar and violent crime," *Journal of the American Academy of Psychiatry and the Law Online*, vol. 18, no. 1, pp. 55–63, 1990.

[54] D. K. Sechrest, D. Shichor, J. H. Doocy, and G. Geis, "A research note: Women's response to a telemarketing scam," *Women & Criminal Justice*, vol. 10, no. 1, pp. 75–89, 1998.

[55] A. Castel, "Fool me once: Why scams leave people feeling foolish." 2021. [Online]. Available: https://www.psychologytoday.com/us/blog/metacognition-and-themind/202104/fool-me-once-why-scams-leave-people-feeling-foolish

[56] L. G. Weinzimmer and C. A. Esken, "Learning From Mistakes: How Mistake Tolerance Positively Affects Organizational Learning and Performance," *The Journal of Applied Behavioral Science*, vol. 53, no. 3, pp. 322–348, Sep. 2017. [Online]. Available: http://journals.sagepub.com/doi/10.1177/0021886316688658

[57] X. Wang, P. Guchait, and A. Paşamehmetoğlu, "Why should errors be tolerated? Perceived organizational support, organization-based self-esteem and psychological well-being," *International Journal of Contemporary Hospitality Management*, vol. 32, no. 5, pp. 1987–2006, May 2020. [Online]. Available: https://www.emerald.com/insight/content/doi/10.1108/IJCHM-10-2019-0869/full/html

[58] F. F. Putri and A. Hovav, "Employees compliance with byod security policy: Insights from reactance, organizational justice, and protection motivation theory," 2014.

[59] M. N. Alraja, U. J. Butt, and M. Abbod, "Information security policies compliance in a global setting: An employee's perspective," *Computers & Security*, vol. 129, p. 103208, 2023.

[60] P. B. Lowry, N. Teh, B. Molyneux, and S. N. Bui, "Using theories of formal control, mandatoriness, and reactance to explain working professionals' intent to comply with new it security policies," in *Roode Workshop on IS Security Research, Boston, MA, USA*, 2010.

[61] T. Neuhaus, "A (nudge) psychology reading of the" nigerian scam"," *Brolly*, vol. 3, no. 3, pp. 7–28, 2020.

[62] A. Jayatilaka, N. A. G. Arachchilage, and A. Babar, "Falling for phishing: An empirical investigation into people's email response behaviors."

[63] G. Bansal, J. Thatcher, and S. W. Schuetz, "Where authorities fail and experts excel: Influencing internet users' compliance intentions," *Computers & Security*, vol. 128, p. 103164, 2023.

[64] L. Shen, "The effectiveness of empathy-versus fear-arousing antismoking psas," *Health communication*, vol. 26, no. 5, pp. 404–415, 2011.

[65] W. Rocha Flores, H. Holm, M. Nohlberg, and M. Ekstedt, "Investigating personal determinants of phishing and the effect of national culture," *Information & Computer Security*, vol. 23, no. 2, pp. 178–199, 2015.

[66] T. Franke, C. Attig, and D. Wessel, "A personal resource for technology interaction: development and validation of the affinity for technology interaction (ati) scale," *International Journal of Human–Computer Interaction*, vol. 35, no. 6, pp. 456–467, 2019.

[67] K. Auspurg and T. Hinz, "Multifactorial experiments in surveys," in *Experimente in den Sozialwissenschaften*. Nomos Verlagsgesellschaft mbH & Co. KG, 2015, pp. 294–320.

[68] A. F. Schmidt and C. Finan, "Linear regression and the normality assumption," *Journal of clinical epidemiology*, vol. 98, pp. 146–151, 2018.

[69] D. M. Rousseau, "Psychological and implied contracts in organizations," *Employee responsibilities and rights journal*, vol. 2, pp. 121–139, 1989.

[70] P. Herriot, W. Manning, and J. M. Kidd, "The content of the psychological contract," *British Journal of management*, vol. 8, no. 2, pp. 151–162, 1997.

[71] D. Lee, H. S. Lallie, and N. Michaelides, "The impact of an employee's psychological contract breach on compliance with information security policies: intrinsic and extrinsic motivation," *Cognition, Technology & Work*, pp. 1–17, 2023.

[72] A. B. Yost, T. S. Behrend, G. Howardson, J. Badger Darrow, and J. M. Jensen, "Reactance to electronic surveillance: a test of antecedents and outcomes," *Journal of Business and Psychology*, vol. 34, pp. 71–86, 2019.

[73] L. A. Bazzano, J. Durant, and P. R. Brantley, "A modern history of informed consent and the role of key information," *Ochsner Journal*, vol. 21, no. 1, pp. 81–85, 2021.

[74] J. Flory and E. Emmanuel, "Interventions to improve research participants' understanding in informed consent for research. a systematic review," vol. 139, no. 2, p. 399. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0002939404015727

[75] A. Demjaha, S. Parkin, D. Pym, T. Groß, and L. Viganò, "The boundedly rational employee: Security economics for behaviour intervention support in organizations1," *J. Comput. Secur.*, vol. 30, no. 3, p. 435–464, jan 2022. [Online]. Available: https://doi-org.proxy.bnl.lu/10.3233/JCS-210046

[76] D. L. Paulhus, "Socially desirable responding on self-reports," in *Encyclopedia of Personality and Individual Differences*, V. Zeigler-Hill and T. K. Shackelford, Eds. Springer International Publishing, pp. 1–5. [Online]. Available: http://link.springer.com/10.1007/978-3-319-28099-8_1349-1

APPENDIX

APPENDIX A
VIGNETTES

Vignette 01

Please put yourself in the following situation and evaluate the following draft of the simulated phishing campaign:

Your employer obtains consent from employees in advance to participate in the upcoming phishing campaign. In the campaign, a simulated phishing email will be sent to all employees of the organization.

In the simulated phishing email, the boss asks the recipient to open a link with important information for an upcoming meeting. In addition, the recipient is promised a salary increase if they commit to the project and provide the requested information on the link.

If an employee falls for the phishing email, there are no consequences for them.

Vignette 02

Please put yourself in the following situation and evaluate the following draft of the simulated phishing campaign:

Your employer obtains consent from employees in advance to participate in the upcoming phishing campaign. In the campaign, a simulated phishing email will be sent to all employees of the organization.

In the simulated phishing email, the boss asks the recipient to open a link with important information for an upcoming meeting. The link leads to information about the meeting.

If an employee falls for the phishing email, there are no consequences for them.

Vignette 03

Please put yourself in the following situation and evaluate the following draft of the simulated phishing campaign:

Your employer obtains consent from employees in advance to participate in the upcoming phishing campaign. In the campaign, a simulated phishing email will be sent to all employees of the organization.

In the simulated phishing email, the boss asks the recipient to open a link with important information for an upcoming meeting. In addition, the recipient is promised a salary increase if they commit to the project and provide the requested information on the link.

If an employee falls for the phishing email, they have to see their boss for an appraisal interview.

Vignette 04

Please put yourself in the following situation and evaluate the following draft of the simulated phishing campaign:

Your employer obtains consent from employees in advance to participate in the upcoming phishing campaign. In the campaign, a simulated phishing email will be sent to all employees of the organization.

In the simulated phishing email, the boss asks the recipient to open a link with important information for an upcoming meeting. The link leads to information about the meeting.

If an employee falls for the phishing email, they have to see their boss for an appraisal interview.

Vignette 05

Please put yourself in the following situation and evaluate the following draft of the simulated phishing campaign:

Your employer obtains consent from employees in advance to participate in the upcoming phishing campaign. In the campaign, a simulated phishing email will be sent to all employees of the organization.

In the simulated phishing email, the boss asks the recipient to open a link with important information for an upcoming meeting. In addition, the recipient is promised a salary increase if they commit to the project and provide the requested information on the link.

If an employee falls for the phishing email, the employee will have to take part in phishing awareness training.

Vignette 06

Please put yourself in the following situation and evaluate the following draft of the simulated phishing campaign:

Your employer obtains consent from employees in advance to participate in the upcoming phishing campaign. In the campaign, a simulated phishing email will be sent to all employees of the organization.

In the simulated phishing email, the boss asks the recipient to open a link with important information for an upcoming meeting. The link leads to information about the meeting.

If an employee falls for the phishing email, the employee will have to take part in phishing awareness training.

Vignette 07

Please put yourself in the following situation and evaluate the following draft of the simulated phishing campaign:

Your employer obtains consent from employees in advance to participate in the upcoming phishing campaign. In the campaign, a simulated phishing email will be sent to all employees of the organization.

In the simulated phishing email, the boss asks the recipient to open a link with important information for an upcoming meeting. In addition, the recipient is promised a salary increase if they commit to the project and provide the requested information on the link.

If an employee falls for the phishing email, the employees' work contract will be terminated.

Vignette 08

Please put yourself in the following situation and evaluate the following draft of the simulated phishing campaign:

Your employer obtains consent from employees in advance to participate in the upcoming phishing campaign. In the campaign, a simulated phishing email will be sent to all employees of the organization.

In the simulated phishing email, the boss asks the recipient to open a link with important information for an upcoming meeting. The link leads to information about the meeting.

If an employee falls for the phishing email, the employees' work contract will be terminated.

Vignette 09

Please put yourself in the following situation and evaluate the following draft of the simulated phishing campaign:

Your employer obtains no consent from employees in advance to participate in the upcoming phishing campaign. In the campaign, a simulated phishing email will be sent to all employees of the organization.

In the simulated phishing email, the boss asks the recipient to open a link with important information for an upcoming meeting. In addition, the recipient is promised a salary increase if they commit to the project and provide the requested information on the link.

If an employee falls for the phishing email, there are no consequences for them.

Vignette 10

Please put yourself in the following situation and evaluate the following draft of the simulated phishing campaign:

Your employer obtains no consent from employees in advance to participate in the upcoming phishing campaign. In the campaign, a simulated phishing email will be sent to all employees of the organization.

In the simulated phishing email, the boss asks the recipient to open a link with important information for an upcoming meeting. The link leads to information about the meeting.

If an employee falls for the phishing email, there are no consequences for them.

Vignette 11

Please put yourself in the following situation and evaluate the following draft of the simulated phishing campaign:

Your employer obtains no consent from employees in advance to participate in the upcoming phishing campaign. In the campaign, a simulated phishing email will be sent to all employees of the organization.

In the simulated phishing email, the boss asks the recipient to open a link with important information for an upcoming meeting. In addition, the recipient is promised a salary increase if they commit to the project and provide the requested information on the link.

If an employee falls for the phishing email, they have to see their boss for an appraisal interview.

Vignette 12

Please put yourself in the following situation and evaluate the following draft of the simulated phishing campaign:

Your employer obtains no consent from employees in advance to participate in the upcoming phishing campaign. In the campaign, a simulated phishing email will be sent to all employees of the organization.

In the simulated phishing email, the boss asks the recipient to open a link with important information for an upcoming meeting. The link leads to information about the meeting.

If an employee falls for the phishing email, they have to see their boss for an appraisal interview.

Vignette 13

Please put yourself in the following situation and evaluate the following draft of the simulated phishing campaign:

Your employer obtains no consent from employees in advance to participate in the upcoming phishing campaign. In the campaign, a simulated phishing email will be sent to all employees of the organization.

In the simulated phishing email, the boss asks the recipient to open a link with important information for an upcoming meeting. In addition, the recipient is promised a salary increase if they commit to the project and provide the requested information on the link.

If an employee falls for the phishing email, the employee will have to take part in phishing awareness training.

Vignette 14

Please put yourself in the following situation and evaluate the following draft of the simulated phishing campaign:

Your employer obtains no consent from employees in advance to participate in the upcoming phishing campaign. In the campaign, a simulated phishing email will be sent to all employees of the organization.

In the simulated phishing email, the boss asks the recipient to open a link with important information for an upcoming meeting. The link leads to information about the meeting.

If an employee falls for the phishing email, the employee will have to take part in phishing awareness training.

Vignette 15

Please put yourself in the following situation and evaluate the following draft of the simulated phishing campaign:

Your employer obtains no consent from employees in advance to participate in the upcoming phishing campaign. In the campaign, a simulated phishing email will be sent to all employees of the organization.

In the simulated phishing email, the boss asks the recipient to open a link with important information for an upcoming meeting. In addition, the recipient is promised a salary increase if they commit to the project and provide the requested information on the link.

If an employee falls for the phishing email, the employees' work contract will be terminated.

Vignette 16

Please put yourself in the following situation and evaluate the following draft of the simulated phishing campaign:

Your employer obtains no consent from employees in advance to participate in the upcoming phishing campaign. In the campaign, a simulated phishing email will be sent to all employees of the organization.

In the simulated phishing email, the boss asks the recipient to open a link with important information for an upcoming meeting. The link leads to information about the meeting.

If an employee falls for the phishing email, the employees' work contract will be terminated.

APPENDIX B
ADDITIONAL FIGURES AND DATA