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, Munich, Germany

[‡]University of Luxembourg, Esch-sur-Alzette, Luxembourg

[§]Ludwig Maximilian University of Munich, Munich, Germany

[¶]Aalto University, Espoo, Finland

Abstract—Organizations depend on their employees' long-term cooperation to help protect the organization from cybersecurity threats. Phishing attacks are the entry point for harmful followup attacks. The acceptance of training measures is thus crucial. Many organizations use simulated phishing campaigns to train employees to adopt secure behaviors. We conducted a preregistered 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 of "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 extensively use email for internal and external communication, making them highly susceptible to phishing attacks. Phishing involves deceptive messages (for example, emails) designed to extract personal or confidential information from victims, often leading to harmful actions [1], [2]. Consequences include personal and financial harm to individuals and organizations [3], infrastructure disruptions such as power or internet outages [4], [5], and broader societal impacts, including the failure of public services depending on critical infrastructure (e.g., hospitals).

Phishing detection involves both human and technical components. Technical measures are critical for phishing pre-

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/ vention, such as blocking phishing emails before they reach the target's inbox. Common strategies include identifying known phishing URLs, removing landing pages [1], detecting phishing-associated website characteristics, and page similarity analysis [6]. Technical approaches are often most effective against large-scale phishing campaigns targeting multiple victims. They are less effective at protecting initial victims, particularly in highly targeted spear-phishing attacks [7], [6].

Technical measures alone cannot prevent all phishing attacks. Consequently, organizations train employees to enhance phishing awareness, encourage cybersecurity incident reporting [8], [9], and mitigate risks. Simulated phishing campaigns are widely used for training and evaluation. These campaigns involve sending realistic phishing emails to employees, often conducted by commercial vendors offering user-friendly interfaces for execution and result analysis [8], [10]. Employees who interact with phishing emails are typically redirected to a training site highlighting phishing indicators [10]. In some organizations, additional consequences for such interactions include mandatory training, discussions on cybersecurity, or, in extreme cases, disciplinary actions [11], [12].

Simulated phishing campaigns have faced significant criticism. Their effectiveness in fostering more secure email behaviors among employees remains uncertain, and they may even undermine organizational security [13], [14], [15]. Beyond their questionable impact, these campaigns can introduce subtle negative effects, such as creating mistrust toward the IT department, burdening employees already under job-related stress [16], and even fostering insider threats through deliberate manipulation of phishing scenarios [17]. Maintaining positive relationships is crucial as IT departments depend on employees' long-term cooperation to safeguard the organization and adapt to evolving security threats. Recent research highlights the importance of making training interventions useful, aligning them with employee motivations, and ensuring positive user experiences to enhance engagement with phishing-related training [18], [19].

Despite criticism and questions about their effectiveness, phishing simulation campaigns remain a common practice in organizations. These campaigns are implemented in various ways, such as informing employees in advance or keeping the simulation covert, and using pretexts that may be perceived as offensive or avoiding them. Such decisions significantly shape employees' perceptions of the campaigns. For example, a simulated phishing email offering bonuses to employees sparked public backlash [20].

This paper investigates factors contributing to the (un)acceptability of simulated phishing campaigns. We aim not to promote such campaigns but to provide a preliminary exploration of elements affecting employee acceptance. While this study does not examine the behavioral impacts of these campaigns, we refer readers to existing discussions and critiques on the topic [13], [14], [15].

We conducted a pre-registered¹ between-subjects vignette experiment (N=793) to examine factors influencing the acceptance of simulated phishing campaigns in organizations. Participants assumed the role of a caseworker reviewing a draft for a potential phishing campaign in their company and assessed its acceptability based on the provided information. Our findings indicate that prior consent to participation positively impacts acceptance of phishing campaigns, whereas including incentives in phishing emails negatively affects acceptance. Additionally, varying the consequences of interacting with the campaign had differential effects on acceptance. However, these variations did not significantly influence the likelihood of manipulation (e.g., clicking on the phishing link) despite awareness of the simulation. Based on these results, we recommend organizations prioritize obtaining informed consent from employees before implementing simulated phishing campaigns, clearly communicate potential consequences, and carefully design phishing pretexts to align with employee expectations and maintain acceptability.

Contribution Statement. This paper offers the following contributions: (1) We explore factors influencing the acceptance of simulated phishing campaigns through a vignette experiment, enabling causal insights into the determinants of acceptance and manipulation probability. (2) We discuss the implications of these findings for user-centered security research and cybersecurity practice, guiding the design of antiphishing training and informing future research directions.

II. BACKGROUND AND RELATED WORK

A. Phishing and its Consequences

Phishing is the act of eliciting sensitive information from victims by impersonating a trusted entity, often following automated patterns and typically delivered via email [21]. In today's interconnected world, phishing represents a significant threat to nearly every company and government institution [22]. The prevalence of phishing attacks continues to rise, with increasingly severe consequences [23], [22]. While email remains the primary medium, attackers also exploit other channels such as instant messaging and SMS [24].

Phishing leverages human psychology to manipulate victims into specific actions. Common tactics include threats, urgency,

¹Pre-registration link: https://osf.io/vz9f2/?view_only=3f95e86f9a4743cba32c9877bb05338f and time pressure [25], often combined with social engineering techniques (distraction, authority, deception) [26], [27].

Research on phishing spans various perspectives. Early studies focused on linking phishing susceptibility to individual characteristics, but these approaches have faced criticism due to the lack of a robust psychological foundation [28]. Recent research suggests that phishing attacks trigger peripheral information processing, reducing critical analysis and increasing susceptibility[29].

The consequences of phishing attacks are severe, concerning personal, financial, and societal domains. For instance, in 2015, spear phishing was used to disrupt Ukraine's power grid, resulting in a six-hour blackout affecting approximately 80,000 people [4], [5]. Financial damages can escalate with follow-up ransomware attacks, where victims face time-sensitive demands to restore access to encrypted data. Such attacks can disrupt production facilities or critical infrastructure, including medical equipment, with potentially life-threatening outcomes [3]. Successful phishing attacks can also damage a company's reputation and erode customer trust [30].

B. Phishing Countermeasures

Organizations and authorities employ a range of countermeasures to combat phishing, including intelligent anomaly detection powered by machine learning, two-factor authentication, and sandboxing techniques [31]. Despite combining these technical defenses, residual risk remains, particularly in organizations where employees frequently interact with external actors. Fully mitigating organizational vulnerability to phishing through technical means alone is challenging [7].

The dynamic nature of phishing attacks creates a constant race between attackers and defenders: as filtering rules improve, attackers adapt their strategies to circumvent detection. This necessitates not only robust technical defenses but also adaptable and vigilant employees. An increasing number of companies and government institutions complement technical solutions with human-focused interventions, such as general cybersecurity training and simulated phishing campaigns.

C. Phishing Simulation Campaigns

Phishing campaigns mimic real phishing attacks but are conducted by internal or external teams acting as controlled adversaries rather than genuine attackers. These campaigns aim to test an organization's defenses by sending tailored phishing emails without causing lasting harm. Volkamer et al. recommend predefining these procedures with leadership to assess organizational vulnerability [32] systematically.

In addition to assessing vulnerabilities, simulated campaigns are often used to evaluate and improve employees' security awareness [32]. Employees who fall for a simulated phishing email, such as by clicking on a phishing link, are typically redirected to training material to enhance their understanding of phishing indicators and improve future behavior [32].

However, the effectiveness of phishing simulation campaigns in fostering secure email practices remains contested. These campaigns do not consistently yield the intended outcomes and may even have unintended negative effects [13], [14], [15]. Also, the cost of simulated campaigns is often underestimated due to extensive personnel hours, as has been revealed by an analysis of Brunken et al. [33].

Simulated campaigns can also provoke adverse reactions from employees. For instance, at Tribune Publishing Company, employees were sent simulated phishing emails promising financial bonuses of \$5,000—\$10,000 following years of real layoffs and wage cuts. The campaign sparked public outrage, eroded trust, and ultimately damaged the company's reputation [20].

Research has further revealed negative behavioral outcomes linked to simulated campaigns. A study involving over 6,000 employees found that those who had fallen for a phishing simulation were more likely to engage with phishing emails in the future [11]. Distler's in-situ deception study identified unintended consequences such as shame and inaction following interaction with simulated phishing emails [17]. Volkamer et al. warn that excessive simulations may lead to employee resignation or loss of motivation. Employees might misinterpret real phishing emails as simulations or deliberately engage with phishing links as an act of protest [32]. Similarly, Mihelic et al. observed reduced vigilance in employees exposed to consecutive phishing attempts, as attackers could exploit such distractions to increase their success rates [34].

The psychological impact of phishing attacks, whether real or simulated, can also be severe. Wood highlights serious outcomes, including anxiety, depression, shame, disrupted sleep, and even increased suicide risk [35]. These findings underscore the importance of carefully considering the psychological effects of simulated campaigns.

Organizations often use click rates—the number of employees who clicked on phishing links—as a performance indicator to measure security awareness. However, this metric fails to capture the context or reasons behind employee actions [34]. Volkamer et al. caution against treating training or simulations as mere checkboxes and shifting blame onto employees who fall victim to phishing despite having undergone training [32].

D. Acceptance of Phishing Simulation Campaigns

Employee and stakeholder acceptance of phishing simulation campaigns is crucial to prevent negative outcomes, such as loss of trust in organizational leadership or IT and disengagement from future training initiatives. We define acceptance of phishing campaigns as the approval of a specific implementation. Acceptance is a multifaceted construct influenced by several factors, including the perceived importance of the campaign, its individual usefulness, personal attitudes toward the measure, the intention to modify behavior based on the intervention, and subsequent engagement [36].

Reed et al. examined public views on the ethics and efficacy of punishment, finding that participants believed punishment should be reserved for serious infractions and viewed it as less effective than positive reinforcement [37]. This raises the question of whether the type of consequence in phishing simulations impacts acceptance. Volkamer et al. emphasize that consequences in phishing campaigns should be transparently communicated and not overly punitive; otherwise, employees may avoid reporting phishing incidents out of fear of repercussions [32]. Similarly, Jampen et al. highlight the need to adapt anti-phishing campaigns to employees' needs to prevent added pressure that could lead to stress, decreased health, and reduced job performance [38]. These findings underscore the importance of employee acceptance in ensuring the success of phishing simulations.

While phishing simulation campaigns aim to educate employees, certain implementation factors can lead to adverse effects, potentially eliciting behaviors contrary to those intended by the intervention. Research in organizational security has shown that restrictive security measures can provoke non-compliance [39] and even computer abuse [40]. These outcomes can be explained by Reactance theory, which posits that individuals may respond to perceived threats to their freedom by engaging in behaviors that counteract the restriction [41]. Reactance is characterized as an active, negative response, distinct from mere avoidance behavior [42]. In some cases, employees may even adopt the very behaviors the measures are designed to prevent [43].

Reactance theory further suggests that not all restrictions provoke the same reactance level. Factors influencing reactive responses include poor organizational communication, such as inadequate explanations of security measures that employees perceive as threats to their autonomy [42], lack of perceived organizational justice (where employees feel unfairly treated), and general distrust toward the organization [40]. Parameters of simulated phishing campaigns, such as the absence of employee consent, may align with these factors, potentially provoking reactive behavior and undermining the intended objectives of the campaign.

E. Summary

- Phishing poses a threat to organizations and government institutions, exploiting human psychology through strategies like appeals to authority and distractions.
- Simulated phishing campaigns are contentious, as they can lead to negative outcomes such as loss of motivation, diminished trust in organizational leadership, and psychological harm, including shame and fear.
- Employee acceptance plays a crucial role in shaping the outcomes of simulated phishing campaigns, influencing trust in employers and potentially leading to adverse behaviors, such as deliberate non-compliance. However, the factors affecting acceptance remain unclear.
- *This paper* examines the factors contributing to the (un)acceptability of simulated phishing campaigns through a vignette study.

III. RESEARCH OBJECTIVES

We address two main research questions:

- **RQ1** What factors influence the acceptance of a simulated phishing campaign within an organization?
- **RQ2** What factors influence the likelihood of participants clicking on the link in a phishing email, even when they are aware it is part of a simulated phishing campaign?

In the field of usable privacy and security, most research involves obtaining informed consent, with deception studies being conducted only rarely [44], [45]. However, organizational contexts operate under different legal and ethical frameworks. Prior studies highlight the importance of psychological contracts in shaping employees' acceptance of organizational cybersecurity measures [46]. Transparent organizational communication, including obtaining employees' consent before launching a simulated phishing campaign, is a key component of maintaining such contracts.

Research has shown that transparent communication positively influences employee engagement in organizational processes [47], acceptance of organizational changes [48], and overall understanding and support for organizational decisions [49]. In the context of simulated phishing campaigns, informing employees in advance and obtaining their consent can enhance the perceived usefulness and credibility of the initiative, fostering greater acceptance and engagement [50].

In contrast, the absence of consent can violate psychological contracts, potentially triggering negative emotional and behavioral responses from employees [51]. Such violations may lead to disengagement, resistance, or even counterproductive behaviors, undermining the campaign's objectives.

We hypothesize that several factors influence employee acceptance of simulated phishing campaigns and their behavior during such campaigns.

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

The case of a newspaper company conducting a simulated phishing campaign that resulted in public backlash [20] serves as a cautionary example of how monetary promises in phishing emails can severely damage trust and have lasting negative consequences for both employees and the organization. This issue was particularly pronounced when employees were experiencing financial difficulties, leading to heightened aversion and a lack of understanding of the campaign's intent [20]. Strong incentives, such as monetary rewards, are known to motivate desired behaviors, which is why they are commonly employed in phishing attacks [52]. However, the perception of being deceived can trigger negative emotional responses, including feelings of betraval and aversion, further exacerbating distrust [53], [54], [35], [55]. Based on these observations, we hypothesize that monetary incentives included in phishing campaign messages may reduce employee 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 investigate the effects of various types of consequences in the context of simulated phishing campaigns. Previous research suggests that punishment is generally perceived as acceptable only when applied to severe misdoings [37]. Weinzimmer and Esken stressed the importance of promoting a culture that tolerates mistakes and is not based on fear of repercussion, as this improves psychological safety and positively impacts organizational learning and performance [56]. Similarly, Wang et al. [57] demonstrated that error tolerance in organizational settings is linked to improved psychological well-being, highlighting the necessity of constructive error management practices to support employees' mental health.

The organizational culture concerning errors matters, and we investigated the effects of different types of negative consequences after interacting with a simulated phishing email. Specifically, we hypothesize that any consequence requiring employees to divert time away from their primary work tasks will likely be perceived negatively, thereby reducing the acceptance of simulated phishing campaigns.

H3: Consequences for the employee resulting from clicking on the phishing link negatively affect the acceptance of the simulated phishing campaign².

Volkamer et al. [32] argued that employees might intentionally click on a phishing link as a form of protest against a simulated phishing campaign. This aligns with research on reactance theory, which has been used to explain noncompliance [58] and computer abuse behavior [40] in organizational contexts. Reactance occurs when humans perceive restrictions on their freedom of choice, leading to oppositional behaviors that aim to restore autonomy.

Studies suggest that mistakes in implementing security measures—such as a lack of transparency, which can cause employees to view organizational communication as "broken"—may exacerbate reactance and increase non-compliance [59], [60]. In contrast, clear and effective communication about security initiatives can reduce reactance. In the context of simulated phishing campaigns, explicitly communicating the purpose of the campaign and obtaining the consent of employees to participate demonstrates respect for their autonomy, which may mitigate reactive behaviors. Based on these insights, 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 long been one of the most commonly employed strategies in phishing attacks, predating the

²In 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.

advent of modern email-based communication [61]. From a user's perspective, emails mentioning money or banking alerts are among the most recognizable forms of phishing, often prompting increased caution [62]. In organizational contexts, the perceived likelihood of receiving such emails is a key factor influencing users' evaluation of their legitimacy [62].

Since monetary-incentive emails are generally uncommon in most organizational settings, they are more likely to be perceived as suspicious or fraudulent. This heightened suspicion may lead to greater user frustration, particularly when monetary incentives are included in simulated phishing emails. Such frustration can, in turn, provoke reactance-based behaviors, as previously discussed in H4, where users respond negatively to perceived manipulative or deceptive attempts.

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.

Previous research emphasized the importance of transparent communication with employees about the consequences of simulated phishing campaigns, recommending that these consequences be handled with care to avoid being overly punitive. Failure to do so may discourage employees from reporting actual phishing incidents due to fear of repercussions [32].

The use of coercive power by authorities or organizational officials can trigger a psychological reaction, where people resist perceived threats to their autonomy [63], [40]. Security measures relying on fear-based tactics or emphasizing severe consequences for non-compliance (overly stringent policies) can increase non-compliant behaviors and provoke reactive responses [63], [40].

In the case of simulated phishing campaigns, awareness of severe consequences for interacting with a phishing attempt may amplify employees' intention to protest or resist these initiatives. Based on these considerations, 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. [64] demonstrated that people with greater computer experience show higher resilience to phishing attacks. We hypothesize that higher IT affinity is positively associated with greater 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 using a $2 \times 4 \times 2$ (Consent \times Consequences \times Incentive) between-subjects design. The independent variables were *Consent* (Yes vs. No), *Consequences* (No impact, Employee interview, Training, Termination after clicking on the phishing link), and *Monetary Incentive* (Yes vs. No). These factors were systematically varied across 16 unique scenarios.

Participants were instructed to assume the role of a caseworker tasked with evaluating a proposed phishing simulation campaign within their organization. Based on the vignette provided, participants assessed the acceptability of the campaign and their likelihood of engaging in manipulative behavior (e.g., intentionally clicking the phishing link despite knowing it was a simulation). Figure 1 illustrates the study procedure. The study received approval from the Ethics Committee of the University of the Bundeswehr Munich, Germany.

Participation in the study was anonymous, so that no conclusions could be drawn about individual persons. Participants first provided informed consent to participate and were introduced to phishing and the study background. They were then randomly assigned to one of the 16 vignettes, ensuring a between-subjects approach to minimize bias and prevent participants from comparing scenarios. After reviewing their assigned vignette, participants evaluated (1) the acceptability of the described scenario and (2) their likelihood of engaging in manipulation despite knowing the campaign was a simulation. Additionally, participants reported prior experience with phishing campaigns, IT affinity, and demographic information. The vignettes are included in Appendix A, and the complete questionnaire, dataset, and analysis syntax are available in the pre-registration³.

a) Pre-tests: To ensure the clarity and usability of the questionnaire, three experts in user-centered security and HCI reviewed the survey while thinking aloud as they completed it. Feedback from this process informed refinements to the question items. A subsequent pre-test with N=35 participants identified potential comprehension issues, particularly with the vignette scenarios. Based on the pre-test results, the wording of the manipulation probability item was revised to enhance clarity and improve the quality of the main study.

B. Vignettes

Each participant was presented with one vignette from the 16 possible scenarios. Figure 2 illustrates the context participants were asked to imagine. Participants were instructed to assume the role of a caseworker evaluating the design of a simulated phishing campaign planned by their company.

The vignettes systematically varied across three factors:

- *Consent:* Whether the employer obtained consent from employees about the upcoming campaign (Yes vs. No).
- *Monetary Incentive:* Whether the phishing email promised a salary increase in exchange for following the instructions in the email (Yes vs. No).
- Consequences: The outcomes for employees who clicked on the phishing link (No impact, Employee interview, Training, or Termination).

Participants evaluated the acceptability of the campaign and their likelihood of manipulation based on the scenario. The detailed vignette descriptions are provided in Appendix A.

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



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. They 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).

C. Measurements

1) Acceptance (Dependent Variable): Measuring acceptance lacks a universally recommended approach. A review in the field of driving automation identified eight major methods for measuring acceptance, varying based on study objectives. Many studies employed a single-item measure of acceptance [36]. Following this precedent, we assessed acceptance of the vignette using a 10-point scale (1 = not acceptable at all; 10 = fully acceptable). Participants were asked: "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 by asking participants: "What is the likelihood that you would click on the phishing link if you already realized it was a phishing email from your employer?" Responses were recorded on a 10-point scale (1 = very unlikely; 10 = very likely). Additionally, participants were asked to justify their responses in an open-text field to provide insights into their motivations for potentially engaging in manipulative behavior during a phishing campaign.

This measure was exploratory in nature, as no prior studies explicitly investigated similar concepts. However, we deemed an empirical examination of this construct valuable for understanding participant behavior.

3) Previous Experience of Phishing Campaigns: After completing the vignette experiment, participants were asked if they had previously participated in a simulated phishing campaign as an employee. This was a filter question with a yes/no response format. Participants with prior experience were invited to describe the campaign in more detail using an open-text field, providing information about the content, number of phishing emails, duration, and scope of the campaign.

Additionally, these participants were asked whether clicking on the phishing link had any consequences and, if so, to describe them in an open-response field. Participants were also surveyed on how they were informed about the campaign, either in advance or afterward, using the following categories: (1) Not at all, (2) Verbally by a supervisor, (3) Email, (4) Work meeting, (5) Training, (6) Note during recruitment, or (7) Other (open response). Finally, participants rated their agreement on two statements—whether the simulated phishing campaign improved their relationship with their employer and whether they viewed phishing campaigns positively—on a 6-point scale (1 = strongly disagree, 6 = strongly agree).

4) IT Affinity: Participants' technical affinity was measured using the short version of the Affinity for Technology Interaction Scale (ATI Scale) by Franke et al. [65].

D. Recruitment and Participants

A total of N=793 participants were recruited via the Prolific platform⁴ in July 2023. Notifications were sent to eligible platform members. Participants were required to use technology at work at least once a week. Individuals who participated in the study's pre-test were excluded, as were participants who did not meet a minimum English proficiency level of B1. Recruitment targeted UK residents, and our goal was to collect data from 800 participants to ensure approximately 50 responses per vignette, as recommended for sufficient statistical power in multi-factorial designs [66].

1) Data Exclusion: Data were collected from 803 participants who completed the questionnaire. Following our preregistration, 10 participants were excluded due to self-reported English proficiency levels of A1 or A2. This ensured that all participants could fully comprehend the vignettes and study materials. The final sample included N = 793 participants.

2) Description of the Sample: The final sample was 48.7% female, 50.2% male, 0.6% non-binary, and 0.5% did not specify their gender. The mean age was M = 41 years (*SD* = 12.85, range = 18–78). Participants were relatively highly educated, with many holding bachelor's or master's degrees.

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



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.



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).

The average ATI score for technology affinity was M = 14.62 (*SD* = 4.59), with high internal consistency ($\alpha = .87$).

E. Experimental Data

Each vignette was viewed an average of 49.56 times. Gender differences on the acceptance scale were analyzed using a *t*-test. Males (M = 5.76, SD = 3.37) reported significantly higher acceptance than females (M = 5.01, SD = 3.28; t(782) = -3.20, p < .001, 95% CI [-1.22, -0.29]). No significant relationship was found between age (p = .50) or education (p = .35) and vignette acceptance. Additionally, no significant gender difference was observed on the manipulation probability scale (t(782) = -0.93, p = .35, 95% CI [-0.15, 0.43]). Correlations between manipulation probability and age (p = .46) or education (p = .31) were also not significant.

The distribution of responses on the acceptance scale was Ushaped, with the highest frequencies at the extremes and fewer responses in the middle range (Appendix B.1). In contrast, responses on the manipulation probability scale were rightskewed, with most responses clustering at low values and a smaller proportion indicating higher likelihoods (Appendix B.2). All response options were used across both scales.

F. Data Analysis

To validate the assumptions of our linear regression model, we ensured the data satisfied the criteria of linearity, independence, homoscedasticity, and normality through visual inspection. Both visual inspection and the Shapiro-Wilk test indicated that the residuals were not normally distributed. However, Schmidt and Finan's findings [67] indicate that violations of the normality assumption do not significantly affect results for large sample sizes (i.e., more than 10 observations/variable).

We conducted separate regression models to estimate the overall effects of the independent variables (*Consent, Mone-tary Incentive*, and *Consequences*) on the dependent variables: (1) acceptance of campaign and (2) manipulation probability. Subsequently, we assessed the effects of individual levels of

each independent variable on the dependent variables. For significant effects, we conducted post-hoc analyses to identify the specific variable levels contributing to the observed effects.

Additionally, we explored the relationship between individual affinity for technology and acceptance of simulated phishing campaigns by calculating and analyzing the individual sum scores for the ATI scale for each participant.

V. RESULTS

A. Bivariate Correlations between Dependent Variables

We conducted a correlation analysis to examine the relationship between the dependent variables, *acceptance* and *manipulation probability*. A significant negative correlation was identified (r = -0.08, p = .02), suggesting that higher acceptance ratings are slightly associated with lower manipulation probabilities. However, the small magnitude of the correlation coefficient indicates that this relationship is weak.

The mean acceptance rating across all vignettes was M = 5.39 (SD = 3.34), while the mean manipulation probability was M = 2.12 (SD = 2.06). These results highlight that, on average, participants rated the vignettes as moderately acceptable and reported a low likelihood of engaging in manipulation.

B. Experimental Evidence

1) Acceptance: We performed a linear regression analysis to examine the relationship between the dependent variable *acceptance* and the independent variables, using a significance level of $\alpha = 0.05$. While the primary focus was on the single effects of the independent variables, the overall effects are also reported for context (see Table T.3 and Figure B.3).

The analysis revealed a significant positive effect of obtaining consent at the beginning of a campaign on *acceptance* (r(789) = 0.28, p < 0.001). Regarding consequences, a significant negative effect on *acceptance* was observed (r(789) =-0.32, p < 0.001). No significant effect of monetary incentives on *acceptance* was found in the overall analysis (r(789) =-0.05, p = .07). A correlation table detailing the relationships between the independent variables is included in Appendix B.

For further analysis, we calculated the single effects of the independent variables' levels (Table I). A statistically significant positive effect on *acceptance* was observed (p < 0.001). Specifically, obtaining prior consent increased the acceptability rating by nearly one scale point (b = 0.90, p < 0.001).

The single effects of the consequences revealed significant impacts on *acceptance*. An employee interview as a consequence of clicking the phishing link led to a decrease in the acceptability rating by more than one and a half scale points (b = -1.64, p < 0.001). Termination of employment resulted in an even larger drop in acceptance, reducing the rating by almost three and a half scale points (b = -3.37, p < 0.001).

Monetary incentives in the phishing campaign were associated with a slight but significant decrease in *acceptance*, reducing the rating by approximately half a scale point (b =-0.47, p = 0.02). Conversely, training as a consequence of the simulated phishing campaign did not have a significant effect

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

on *acceptance* (b = 0.29, p = .32). A visual representation of the coefficients is provided in Figure 4.

2) Manipulation Probability: Overall, participants reported a very low likelihood of clicking on a phishing link if they recognized it as originating from their employer (see Appendix Figure B.2). We analyzed the overall effect of the independent variables (*Consent, Incentive,* and *Consequences*) on the dependent variable manipulation probability at a significance level of $\alpha = .05$. The analysis revealed no statistically significant relationships between the independent variables and manipulation probability (Appendix Table T.4 and Figure B.4).

We then examined the single effects of the variable levels (Figure 5). Neither obtaining consent nor including a monetary incentive in the phishing email had statistically significant effects on manipulation probability (p > .05).

To analyze the open-ended responses, we categorized participants' answers, ensuring each response was assigned to at least one category. Participants were asked to explain their reasons for potentially clicking on a phishing link, even when they knew it was part of a simulated phishing campaign. The analysis identified three primary reasons for intentionally clicking a suspected link from an employer: false trust in the email (n = 44), protest (n = 11), and curiosity (n = 9).

Participants who expressed *false trust in the email* often believed it to be legitimate, particularly when it appeared to come from a trusted authority, such as their boss. For example, responses included, "The boss has specifically asked me to open it, so I would think it is OK" (P399) and "I think I'm

 TABLE II

 Single effects of the independent variables on the manipulation probability. Manipulation probability was measured on a scale of 1=very unlikely to 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

quite trusting and do as I'm told" (P47).

Some participants cited *protest* as their motivation for clicking on phishing links, using it as a form of defiance against the campaign. For instance, they 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).

Finally, *curiosity* was another common reason for interacting with phishing emails. Responses such as "Just to read it" (P134) and "Just out of curiosity I guess.[...]" (P665) illustrate how phishing messages can pique recipients' curiosity, leading to engagement with the simulated phishing email.

3) ATI and Acceptance: Participants had an average ATI value of M = 14.62 (SD = 4.59). To examine whether technology affinity influences the acceptance of phishing campaigns, we calculated a Spearman correlation due to the violation of the normality assumption. The analysis revealed a significant positive correlation (r(791) = 0.12, p < 0.001), indicating that individuals with a higher affinity for technology tend to rate the acceptance of phishing campaigns more favorably. Based on these findings, Hypothesis 7 is supported.

C. Prior Experience with Simulated Phishing Campaigns

Of the 793 participants, n = 179 (23%) reported prior experience with simulated phishing and completed additional questions describing their experiences. This section focuses on the responses from this subset of participants and serves a descriptive purpose, independent of experimental treatments. Among participants with prior experience, 64% indicated that they were not informed about the phishing campaign in advance. When asked how they were informed beforehand, 17% reported being notified via email, 10% as part of training, 7% by their supervisor, and just under 3% each during a work meeting or the application process.

Regarding post-campaign communication, 77% stated they were informed about the phishing campaign via email, 10% each in the context of training and/or during a work meeting, 8% verbally by their supervisor, and 5% reported receiving no notification at all. A few participants mentioned alternative channels, such as an announcement on the company's intranet or informal conversations with colleagues.

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

D. Summary of the Results

We summarize the results of this study in Table III.

E. Limitations

Our study has several limitations that should be considered when interpreting the results. First, we were unable to test every possible scenario or combination of variables that might occur in real-world contexts. Consequently, not all potential expressions of the independent variables could be represented. Nevertheless, our methodological approach allowed us to isolate the effects of the variables included in the study and draw conclusions about their specific impacts.

Second, the consequence "training" may be perceived differently by different employees, depending on factors such as content, delivery method, and personal preferences. Future research could benefit from providing more detailed descriptions of various consequences to better understand their relative acceptance and effectiveness.

Third, our sample consisted exclusively of UK residents. This limits the generalizability of our findings to other cultural contexts, as cross-cultural differences may influence the dependent variables of acceptance and manipulation probability. Investigating these variables across diverse cultural settings would be an important avenue for future work.

Finally, the vignettes varied in length, with some being longer than others. As each participant was exposed to only one vignette, the average study duration was relatively short (M = 04:54 minutes), minimizing the likelihood of fatigue effects. We assume this had no significant impact on the results, but future studies could standardize vignette length.

VI. DISCUSSION

A. Acceptance

a) Consent and Simulated Phishing Campaigns: Consent and simulated phishing campaigns are a highly debated topic. In our study, obtaining consent significantly improved the

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 em- ployee resulting from clicking on the phishing link have a negative effect on the acceptance of the simu- lated phishing campaign.	Partially confirmed	Training had statistically non-significant effect on ac- ceptance. 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 neg- ative 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 de- spite knowledge of the simulated phishing campaign.	Not confirmed	No statistically significant effect of stronger conse- quences 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.

acceptance rating of simulated phishing campaigns. In practice, these campaigns often involve some degree of deception [32]. For instance, some organizations never disclose that a simulated campaign has been conducted, simply redirecting victims to a legitimate website. Others inform victims only after they have fallen for a simulated phishing attempt [32].

Securing consent prior to initiating simulated phishing training can be examined through the lens of the *psychological contract*, which refers to the implicit expectations between employees and employers regarding their mutual responsibilities [68]. Prior research highlights the critical role of psychological contracts in shaping employees' acceptance of organizational cybersecurity policies [46]. Justice and fairness are perceived as core components of these contracts, with employees expecting organizations to act transparently and declare their intentions openly [69]. Failure to secure clear consent for simulated phishing emails risks breaching this contract, which can trigger negative emotional and behavioral responses [51]. Such breaches may undermine employees' trust and commitment to security measures [70].

Reactance theory offers another perspective on the importance of consent in simulated phishing campaign acceptance. Reactance describes a negative emotional reaction triggered by perceived threats to or restrictions on an individual's behavioural freedom, which in turn leads to actions meant to restore freedom[41]. In organizational contexts, this response frequently arises in reaction to security measures perceived as controlling or invasive [71]. Our findings support this framework: when simulated phishing campaigns are conducted without employee consent, they may be perceived as restricting employees' freedom to participate voluntarily. Such perceptions can provoke negative responses, such as deliberate noncompliance, including intentionally clicking on phishing links, thereby countering the organization's intended objectives.

Informed consent is a fundamental ethical safeguard in most empirical studies on usable privacy and security [44]. However, the precise amount and type of information required to qualify as informed consent remain unclear [72]. Lengthy and overly complex consent documents may fail to effectively inform research participants or employees. Studies have shown that prospective participants often do not fully comprehend the information disclosed during the consent process [73], and similar challenges may apply in workplace settings.

Further research is needed to determine how employees can best be informed about simulated phishing campaigns in ways that balance respect for their time with the provision of all necessary information. Simulated phishing campaigns inherently generate personal and potentially sensitive data about employees. An informed consent process should address key aspects, including who has access to the data, how long it will be stored, how it will be secured, and how employees can revoke their consent. For further guidance, see [32]. Importantly, consent must be both informed and freely given. Employees who opt out of simulated phishing campaigns should be provided with alternative learning opportunities to enhance their phishing countermeasure skills.

To foster acceptance and effectiveness, organizations could consider involving employees in the design and refinement of security measures through co-design sessions. By incorporating employees' perspectives and experiences from their daily work, organizations can create security initiatives that are more aligned with employees' needs, thus enhancing trust and engagement.

b) False Promises of Monetary Incentives: Our findings indicate that the inclusion of a monetary incentive in the email content had a small but statistically significant negative effect on acceptance. In real-life phishing campaigns, the impact of a promised incentive likely depends on the organizational context. It is also important to consider that other pretexts may similarly have negative effects on acceptance. Sensitive topics such as vacation days, sick leave, organizational restructuring, or politics could evoke comparable adverse reactions, highlighting the need for careful consideration of email content in both simulated and real-world scenarios.

c) Consequences of Interacting with a Phishing Email: Acceptance ratings varied based on the consequences described in the vignette scenarios (see Figure 4). Contract termination, consistent with prior findings [37], and an employee interview both had a statistically significant negative effect on acceptance. Conversely, training as a consequence had a negative effect, though it was not statistically significant.

We hypothesize that the impact of training on acceptance likely depends on multiple factors, such as the duration of the training, its perceived relevance, and whether it is viewed as "embarrassing" (e.g., if supervisors are informed of participation) or as a constructive and helpful measure. These considerations suggest that the design and delivery of training programs play a critical role in shaping employees' acceptance of such consequences.

B. Manipulation Probability

Previous research has noted the possibility of employees intentionally clicking on simulated phishing links as a form of protest, driven by feelings that it is unreasonable for their organization to "trick" them, or out of curiosity [32]. In our study, the majority of participants indicated that they would not knowingly click on a simulated phishing email from their employer, and none of the vignette factors had a statistically significant effect on manipulation probability. However, open-ended responses provided some evidence of intentional clicking, motivated by curiosity or the perception that there would be no real consequences for doing so.

Several potential follow-up hypotheses could explain these findings. First, the intention to manipulate a simulated phishing campaign may be generally uncommon, which could account for its infrequent occurrence in our sample. Second, the intention to manipulate may be highly context-dependent and tied to real-life organizational settings, making it difficult to replicate using vignette-based scenarios. For example, [74] highlight the tensions that arise in organizational contexts when time, resource, and cognitive constraints intersect with incomplete information and conflicting security demands. These tensions often lead employees to make "good enough" decisions. Time and resource pressures, which are integral to workplace environments, cannot easily be replicated in a vignette study, suggesting that intentional clicking on a suspected phishing link may emerge only under such real-world conditions.

Finally, a social desirability bias may influence participants' responses, as individuals might refrain from admitting behaviors they perceive as socially undesirable, such as deliberately clicking on phishing links, in an effort to present themselves in a more positive light [75].

C. IT Affinity

This study also examined whether higher IT affinity is associated with greater acceptance of phishing campaigns. Consistent with our findings, Flores et al. [64] demonstrated that individuals with greater computer experience tend to exhibit higher resilience to phishing attempts.

VII. RECOMMENDATIONS

We provide practical recommendations for designing future simulated phishing campaigns. It is important to note that this study does not evaluate whether these campaigns improve overall security outcomes. Instead, our focus is on identifying the factors that influence their acceptance.

A. Obtain consent from participants before including them in simulated phishing training

Our study highlights the positive effect of obtaining employee consent on the acceptance of phishing campaigns. While obtaining consent may influence employees' behavior in the short term, it is crucial to balance this with the long-term objectives of maintaining engagement with security measures and fostering trust in an organization's security professionals.

Conducting simulated phishing campaigns with prior employee consent appears to be a worthwhile approach, as it aligns with ethical considerations and builds transparency. Additionally, organizations should carefully assess the effects of prior consent on both acceptance and behavior. Employees who choose not to provide informed consent should be offered alternative forms of security training to ensure inclusivity and equal access to cybersecurity education.

B. 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

Our findings indicate that the consequences of a simulated phishing campaign significantly influence its acceptance. To enhance transparency and trust, organizations should clearly communicate the intentions of the campaign, the potential consequences for employees (both positive and negative), and how employee data will be used (e.g., who has access, whether it will be used for performance evaluations). Combining this approach with prior informed consent ensures that employees understand the scope and purpose of the campaign. Additionally, we found that consequences such as employee interviews resulted in lower acceptance of the simulated phishing campaign. This insight is particularly relevant for smaller organizations, where such measures may be more feasible but could still negatively impact employee perceptions. Furthermore, there is no evidence suggesting that employee interviews positively influence phishing-related behaviors. Therefore, organizations should carefully consider alternative approaches that maintain employee trust and engagement while achieving the desired training objectives.

C. Organizations should carefully evaluate the appropriateness of pretexts used in simulated phishing campaigns

Our findings indicate that promising an incentive in the phishing email had a negative effect on the campaign's acceptance. While real attackers may employ any means necessary to deceive their targets, organizations should prioritize fostering employees' long-term commitment to security over employing deceptive or controversial tactics.

Simulated phishing campaigns should be designed with pretexts that align with organizational values and respect employees' trust. Using inappropriate or manipulative pretexts risks undermining employee engagement and could erode trust in the organization's security measures. By carefully considering the acceptability of pretexts, organizations can maintain transparency, build trust, and enhance the overall effectiveness of their security training initiatives.

VIII. FUTURE RESEARCH

Future research should explore the real-life acceptance of employees who have participated in simulated phishing campaigns within their organizations. In particular, interviewing employees who expressed disagreement or dissatisfaction with these campaigns could provide valuable insights into potential areas for improvement. Understanding their concerns and perspectives would help refine the design and implementation of such initiatives.

Additionally, future studies should investigate how employees perceive the impact of these campaigns on their behavior. Behavior change in organizational contexts is inherently complex and influenced by numerous factors, including the nature of the intervention and the organizational culture. Therefore, further research should examine the specific characteristics of simulated phishing campaigns—such as their frequency, transparency, and consequences—to gain a deeper understanding of their effectiveness and how they are perceived by employees. This detailed analysis could help identify best practices for fostering meaningful and lasting behavior change in cybersecurity.

IX. CONCLUSION

Our findings highlight the impact of varying key factors—consent, monetary incentives, and consequences—when designing simulated phishing campaigns. The results reveal that these factors can significantly influence employee acceptance, underscoring the importance of thoughtful campaign design. However, it is important to note that this study does not assess or advocate for the overall effectiveness of such campaigns in improving organizational security, a topic that has been questioned by multiple studies.

We encourage future research to explore ways to enhance the effectiveness of anti-phishing training, whether simulated or otherwise, while simultaneously ensuring these measures are accepted by employees. Long-term organizational security relies on the sustained collaboration, trust, and motivation of employees. Therefore, any security measure should be rigorously evaluated not only for its behavioral impact but also for how it is perceived by employees.

We hope to see more research focused on understanding employee engagement, motivation, and acceptance of security measures, as these factors are critical to creating effective, ethical, and sustainable organizational cybersecurity practices.

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

- 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 hospitals and medical care facilities." on 2021 https://www.cyber.forum.yale.edu/blog/2021/7/20/ attributing-deaths-to-ransomware-attacks-on-hospitals-and-medicalcare-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 Selfefficacy, 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," *Humancentric 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.
- "Fool [55] A. Castel, Why me once: scams leave people foolish." 2021. feeling [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] 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.
- [65] 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.
- [66] K. Auspurg and T. Hinz, "Multifactorial experiments in surveys," in *Experimente in den Sozialwissenschaften*. Nomos Verlagsgesellschaft mbH & Co. KG, 2015, pp. 294–320.
- [67] A. F. Schmidt and C. Finan, "Linear regression and the normality assumption," *Journal of clinical epidemiology*, vol. 98, pp. 146–151, 2018.
- [68] D. M. Rousseau, "Psychological and implied contracts in organizations," *Employee responsibilities and rights journal*, vol. 2, pp. 121–139, 1989.
- [69] 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.
- [70] 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.
- [71] 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.
- [72] 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.
- [73] 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

- [74] 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
- [75] D. L. Paulhus, "Socially desirable responding on selfreports," 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



Fig. B.1. Distribution of responses on the acceptance scale of all vignettes



Fig. B.3. Overall effects of the independent variables on acceptance



Fig. B.2. Distribution of responses manipulation probability

 TABLE T.1

 Correlation Table of independent variables on acceptance

Variable	М	SD	1	2	3	4
1 Acceptance	5.39	3.34	-	.28**	32**	05
2 Consent				-	34**	01
3 Consequences	2.44	1.12			-	001
4 Incentive						-
N = 702 *= < 05	**	01				

N = 793, *p < .05, **p < .01

TABLE T.2 Correlation Table of independent variables on manipulation probability

Variable	М	SD	1	2	3	4
1 Manipulation Prob.	2.12	2.06	-	.03	.01	.04
2 Consent				-	34**	01
3 Consequences	2.44	1.12			-	001
4 Incentive						-
N = 793, *p <.05, **p <.01						

 TABLE T.3

 Regression table overall effects on acceptance. Acceptance

 was measured on a scale of 1 to 10.

Term	Estimate	SE	p-value
(Intercept) Consent Incentive	5.9972 1.7731*** -0.3464	0.2665 0.2334 0.2235	< .001
Consequences (dummy)	-1.4830***	0.2255	< .001



Fig. B.4. Overall effects of the independent variables on manipulation probability

TABLE T.4Regression table overall effects on manipulation probability.Manipulation probability was measured on a scale of 1 to 10.

Term	Estimate	SE	p-value
(Intercept)	1.8898	0.1748	
Consent	0.1587	0.1531	0.300
Incentive	0.1757	0.1466	0.231
Consequences (dummy)	0.1269	0.1678	0.450