

Evaluating Organizational Phishing Awareness Training on an Enterprise Scale

Doron Hillman , Yaniv Harel , Eran Toch

PII: S0167-4048(23)00274-2
DOI: <https://doi.org/10.1016/j.cose.2023.103364>
Reference: COSE 103364



To appear in: *Computers & Security*

Received date: 22 February 2023
Revised date: 28 May 2023
Accepted date: 22 June 2023

Please cite this article as: Doron Hillman , Yaniv Harel , Eran Toch , Evaluating Organizational Phishing Awareness Training on an Enterprise Scale, *Computers & Security* (2023), doi: <https://doi.org/10.1016/j.cose.2023.103364>

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Evaluating Organizational Phishing Awareness Training on an Enterprise Scale

Corresponding Author: Doron Hillman, MSc

Order of authors:

1. Doron Hillman, MSc - Tel Aviv University, Israel
2. Yaniv Harel, Dr. - Tel Aviv University, Israel
3. Eran Toch, Prof. - Tel Aviv University, Israel

Abstract

Employees are often the victims of phishing attacks, posing a threat to both themselves and their organizations. In response, organizations are dedicating resources, time, and employee effort to train staff to identify simulated phishing attacks. However, the real-world effectiveness of these training efforts in large enterprises remains largely unexplored. To address this, we carried out a controlled experiment in an Israeli financial institution with approximately 5,000 employees. The experiment included three simulated phishing emails, and we examined how different factors influence the phishing Click-Through Rate (CTR). Our findings suggest that employees are more likely to engage with phishing simulation emails that use personalized phrasing. We also found that phishing CTR varies between business units, and that the timing of training before the simulated email did not significantly affect phishing CTR. Furthermore, it became clear that training prior to phishing simulations and adopting a data-driven approach that includes process, variable and measure analysis, can enhance organizational awareness of phishing. Although advanced technologies can mitigate some phishing attacks, our research indicates that employee awareness and proactive behavior will continue to play a critical role in the foreseeable future. The paper concludes by providing guidelines to information security officers on establishing effective organizational awareness to prevent phishing attacks.

Keywords

Phishing; Phishing wave; Social engineering; Organizational cyber security; Awareness; Training;

1. Introduction

Phishing attacks present a significant threat to both enterprises and overall internet security (Jain and Gupta, 2022; Johns, 2020; Vega et al., 2022; Wash and Cooper, 2018). The FBI reported that these attacks resulted in a staggering financial loss of over \$1.8 billion in 2020 for the American population (Wang and Song, 2021). Phishing attacks wreak havoc on organizations by illicitly obtaining critical information like social security numbers, residential addresses, account details, product specifics, and intellectual property. This stolen information is then often sold for a significant sum in black markets (Jain and Gupta, 2022). In recent years, the propagation of ransomware attacks, a type of extortion-based malware threat, has been one of the more devastating consequences of phishing. In such instances, the harm isn't due to the selling of data but the attack's nature, which often involves ransom payments, system restoration efforts, and the need to address resulting compensation claims or expenses (Harel, 2021; Thomas, 2018). Aside from financial losses exceeding millions of dollars, phishing attacks damage reputations and result in the loss of customers.

The State of the Phish report by Proofpoint (2022) suggests that organizations are allocating more resources and efforts to thwart phishing attacks and educate their employees and users on how to recognize such threats. However, despite these efforts, phishing emails remain a prevalent attack method against both companies and individuals. Security reports have found that 41% of organizations face phishing attacks daily, and 77% are targeted at least once a month (De Bona and Paci, 2020). Over the past few years, numerous countermeasures have been introduced to combat phishing. These include email filtering systems, phishing website detection, analysis of phishing campaign patterns, and employee-focused interventions such as susceptibility triggers to phishing attempts, along with various educational approaches. In addition, a whole industry has sprouted up offering services and products aimed at phishing prevention. These offerings range from training and educational services to databases of known URLs and email addresses associated with phishing attacks, as well as email filters informed by threat intelligence gathered by experts and feedback from customers (Lain et al., 2021).

A primary strategy organizations use to heighten employee awareness of phishing emails involves conducting simulated phishing tests. During this process, the organization disseminates simulated phishing emails to various employees and tracks their responses, such as the number of employees who click on links within these emails. Despite the increasing implementation and potential efficacy of training and awareness programs (Longtchi et al., 2022), employees remain susceptible to phishing. This ongoing vulnerability led the UK National Cyber Security Centre to publish specific guidelines on how organizations can fortify their defenses (Williams et al., 2018).

Organizations generally operate under a centrally managed security policy, which is shaped by the insights, recommendations, and regulatory mandates provided by security practitioners. However, a challenge lies in the fact that these policies are often drafted without sufficient regard for the objectives and capabilities of the employees expected to abide by them (Beautement et al., 2016). The security measures implemented by an organization influence its ability to mitigate risk through prevention strategies (whereby employees are required to report security lapses and issues to minimize the fallout of security breaches) and detection methods (aimed at identifying security vulnerabilities before they evolve into major incidents). Despite the deployment of technical solutions, a high rate of information security breaches persists, many of which could be prevented with appropriate

attention to employee actions (Amankwa et al., 2018; Hart et al., 2020; Pac and Capstone, 2017; Safa and Maple, 2016). While organizational security systems offer an increasingly robust technical defense against certain types of attacks (such as detecting anomalies in data movement), they can prove largely ineffectual if not used correctly by users who neglect to follow security guidelines (Torten et al., 2018). At present, the metrics used to measure security behavior are fairly limited, typically focusing on technical information about attempted intrusions, virus logs, access requests, and network traffic data (Kirlappos et al., 2015).

Research indicates that cyber-attack prevention guidelines can enhance secure online behavior. This was demonstrated in a study where internet users made simulated online purchases of a digital product on an e-commerce platform (van Bavel et al., 2019; Wen et al., 2019). In a simulated secure browsing scenario on an e-commerce site, participants responded most effectively to guidance on online self-protection. Security alerts that encourage sustained behavioral modifications stand a better chance of succeeding (van Bavel et al., 2019). In another user study, educational games were shown to be an effective tool in improving learner performance. A role-playing simulation game called "What. Hack", which offers engaging anti-phishing training, was successful in improving players' ability to identify incoming threats by 36.7% (Wen et al., 2019). Cybersecurity awareness and behavior can be enhanced through warnings. However, the majority of users do not modify their behavior in response to financial incentives. This suggests that psychological elements should be taken into account when devising defense strategies (Longtchi et al., 2022).

While phishing attacks are a prevalent method of breaching an organization's information systems, the majority of studies on phishing primarily concentrate on individual susceptibility. This leaves a significant gap in research related to the organizational aspects of phishing attacks, particularly in industrial contexts. Given the significant risk that phishing attacks present to organizations, leading to detrimental consequences and data breaches, our study focused on real-world environments. Our aim was to offer pragmatic solutions that can aid organizations in reducing the harm caused by phishing attacks. Organizational effectiveness and individual improvement are two separate concepts. Individual improvement primarily focuses on personal growth and development, whereas organizational effectiveness pertains to enhancing the overall performance of the entire organization. This improvement is achieved through various means, including engaging in discussions about phishing emails and delivering training sessions (Wash and Cooper, 2018). Research carried out in real-world organizational settings points to the complexities of deploying anti-phishing education in the real world (De Bona and Paci, 2020), but the organization's effectiveness in large-scale real-world settings, particularly regarding the phrasing of the phishing attacks (personalized versus general) and how different business units have been affected over time, was still not carried out. This research aims to examine the effectiveness of organizational phishing training in relation to the aforementioned factors. Specifically, the study intends to accomplish the following objectives:

- A) Assess the impact of longitudinal Information Security Awareness (ISA)-simulated anti-phishing campaigns within the context of an organization. This evaluation includes establishing baselines to measure the success rate of phishing campaigns.
- B) Evaluate the effectiveness of different types of simulated phishing attacks, specifically personalized versus general approaches.
- C) Analyze and compare the variances in outcomes between different business units operating within the organization.

2. Background

2.1 Phishing attacks

Phishing is “a fraudulent activity that involves the creation of a replica of an existing web page to fool a user into submitting personal, financial, or password data” (Jain and Gupta, 2022). These attacks usually involve tricking the victim into taking the attacker's desired action, and they merge social psychology, technical systems, security subjects, and politics. A typical phishing attack involves a message (e.g., email) sent to a user that purports to be from a legitimate entity such as a bank; the message would attempt to convince the user to visit the phisher's website (which mimics the real website), in which the phisher requests sensitive data (e.g., username and password) from the victim (Lin et al., 2010). In this work, we look at broad phishing campaigns, which target large number of users at the same time. These broad attacks are in contrast to spear phishing, which are targeted attacks which involve carefully crafted email messages containing contextual information specific to the individuals and organizations targeted in the attack (Alkhalil et al., 2021; Halevi et al., 2015; Schuetz et al., 2016). The number of phishing attacks is on the rise, with nearly 90% of organizations experiencing targeted phishing attacks in 2019 (Alkhalil et al., 2021; Michelle et al., 2005). As people are the weakest link in any security chain (van Bavel et al., 2019; Das et al., 2020), a technical report by Cisco (2017), says that adversaries have a wide range of resources to make their attacks more effective such sophisticated attacks are characterized by social engineering tactics that manipulate the victims.

Phishing countermeasures in an organization can usually be divided into educational approaches (which are further discussed in the next section), technical approaches such as automated classification techniques that minimize human involvement (blacklist, machine learning, search engine, and visual similarity-based approaches) (Jain and Gupta, 2022), and security warnings such as a browser plug-in to warn users about suspicious webpages (Egelman et al., 2008; Jain and Gupta, 2022) or domain highlighting (Lin et al., 2010). However, if the user does not understand what phishing is, then they are less likely to give attention to these attacks (Egelman et al., 2008), which makes such techniques only effective for some types of users (Lin et al., 2010). In an organizational context, as many employees find it difficult to recognize phishing emails, training them to guard against phishing emails would be beneficial (Jevšček et al., 2019). To date, most of the research has concerned how individual internet users react to phishing messages rather than how organizations' employees react (e.g., detect, respond) to such attacks (Jevšček et al., 2019) or how these reactions can be measured and used to improve organizational aspects.

2.2 Cyber-security compliance

There is evidence that non-compliance with policy is common (Becker et al., 2017). Compliance with organizational security policies is influenced by the outcomes (costs and benefits) of compliance. Employees, however, are naturally motivated to focus on their primary task for which they are compensated rather than on secondary tasks for which they are not directly compensated (e.g., security). This makes actual security practices influenced by several factors, such as the company's security culture, employees' trust in the company, and employees' beliefs about the effects of their primary task (Petrykina et al., 2021). Security breaches, however, cause more damage to enterprises than to individuals, except in rare circumstances, which makes compliance with security policies inherently difficult (Petrykina et al., 2021). Beautelement and Sasse show that users have a

finite capacity for complying with security regulations through an economic model of compliance. A user's productivity can be harmed if this capacity is exceeded: "when the compliance budget begins to run out, an employee may see more utility in completing tasks more directly relevant to them, rather than trade these off with security tasks" (Beautement and Sasse, 2009). Additional compliance challenge is cybersecurity fatigue, which is a weariness or aversion to cybersecurity-related workplace behaviors or advice and occurs as a result from overexposure to workplace cybersecurity advice (e.g., training) or cybersecurity actions (e.g., forced password updates). There can be two types of cybersecurity fatigue: attitudinal (e.g., a belief that cybersecurity is not important) and cognitive (e.g., habituated bad behaviors) (Reeves et al., 2021).

We can divide the factors that influence employees' compliance with information security policies into organizational factors (e.g., awareness and training) and human factors (e.g., personality and habits) (Alotaibi et al., 2017). Overall, the primary cause of information security breaches is unintentional factors such as negligence, ignorance, and a lack of awareness (Safa and Maple, 2016). This means that despite the use of technology such as system logs, humans are most likely to recognize abnormal activities within the organization. The literature has varying conclusions, as some studies have found that security training improves a user's ability to detect deception attempts (Michelle et al., 2005), while others present contrary evidence that shows that training has no significant effect on detecting such attempts (Jevšček et al., 2019). This can be explained as not all types of users behaving the same way. For example, students who are prone to taking risks are more likely to become victims of cyberattacks (Wen et al., 2019). Given that it is practical in large organizations to use employees as a collective phishing detection mechanism (Lain et al., 2021) and since well-trained employees may become the strongest links regarding information security (Michelle et al., 2005), analyzing employees' behavior in real-world settings is vital for understanding how to motivate them to comply.

2.3 Phishing awareness training

According to a meta-analysis of 48 papers that describe field experiments, the mean phishing emails' click-through rate across all studies and measurements is 24% (Sommestad and Karlzén, 2019). Recipients are as susceptible to phishing emails that target them specifically (e.g., saluting them by name) as they are to general undirected phishing (17% for both), and individualized adaptations (e.g., saluting the recipient by name) are on average no more successful than general emails.

Several studies have examined what influences clicking on links in phishing emails (van Bavel et al., 2019; Benenson et al., 2017; Conway et al., 2017; De Bona and Paci, 2020; Downs et al., 2006; Jevšček et al., 2019; Sahni et al., 2016). Among the factors that increase the susceptibility of users to phishing, we can find principles of influence as users rely on context-specific information, such as having an account with the company that sent the mail or receiving an unexpected message (Downs et al., 2006), authority (Butavicius et al., 2015; De Bona and Paci, 2020), the low cognitive processing of emails, e.g., during busy hours (Conway et al., 2017), self-efficacy, security breach experiences at work, and perceived susceptibility (van Bavel et al., 2019; Blythe and Coventry, 2018). A university experiment (Benenson et al., 2017) found that curiosity about the content is the most common reason for clicking (34%), followed by explanations that the message met expectations (27%) and that the respondents thought that they knew the sender (16%). These results demonstrate that people's decisional heuristics are relatively easy to misuse in a targeted attack, which makes defense more challenging (e.g., a spear

phishing message that may address victims by name, mention their immediate interests, or resemble a genuine email from a trusted sender). Other common reasons to click were curiosity and a match of the message to the recipient's expectations (Jevšček et al., 2019) and adding the name of the message recipient to the email's subject line, which increased the probability of the recipient opening it by 20% (Sahni et al., 2016).

Among the factors that decreased susceptibility to phishing emails were liking, scarcity, and social proof (Butavicius et al., 2015; De Bona and Paci, 2020), and based on another study, the most prominent reason for not clicking was an unknown sender name, then a message that did not address the recipient by name or was "anonymous", followed by suspicion of fraud and the situation context (Benenson et al., 2017).

Notably, click rates vary quite substantially according to the experiment setup and treatment group, from 2% to 97% (Jevšček et al., 2019), while the average is 24% (Sommestad and Karlzén, 2019). The causes for such extreme variations in click rates may be sought in the message content, research methods, demographics of the targeted population, and timing and frequency of the messages (e.g., the number of emails sent to an individual). When comparing different studies, click rates are influenced most by persuasiveness, reputation mechanisms, and other cues that cause the recipient to recognize a fraudulent message as legitimate (e.g., credibility of the message content, recognizability of the message design, and recognizability of the sender) (Jevšček et al., 2019).

2.4 Phishing attack simulation

From an organizational perspective, taking several variables into account is essential. Information security awareness (ISA) is a practice that refers to the set of skills that help a user successfully mitigate security threats. This is a very effective method to prevent security attacks, identify vulnerable employees, and mitigate vulnerabilities to improve security (Bitton et al., 2019). As technical means of mitigating cyberattacks are becoming increasingly sophisticated, the human factor remains a "lever" that hackers use to compromise organizations' security systems (Hagen et al., 2008), which makes employees the most vulnerable point in any phishing ecosystem (Wang and Song, 2021). This makes phishing attacks the most widely used social engineering technique to achieve fruitful results (Hagen et al., 2008). Most current studies on ISA assessment depend on the subjects' responses to questionnaires or surveys, although actual phishing exercises can improve the subjects' ISA with regard to phishing (Bitton et al., 2019).

To mitigate and prepare for phishing attacks, education efforts are currently the most widely used intervention (Wang and Song, 2021). Organizations teach employees how to identify and avoid phishing emails and websites through training, workshops, and awareness programs. Using such approaches aims to improve the users' ability to identify phishing attacks, for example, by game-based training approaches (which are convenient and easy to learn as they provide teaching in a natural environment) or training following a study on user's response (which analyzes why they fell for the scam and involves training them accordingly) (Jain and Gupta, 2022). Training users to make more secure decisions means persuading them to make different choices (Wash and Cooper, 2018). Effective ways to improve cybersecurity behavior are by educating about threats (threat appraisal), including information about appropriate protective measures (van Bavel et al., 2019), sanctions that are promptly enforced (Siponen et al., 2010), adjusting security policies based on analyzing employee security behaviors (Beautement et al., 2016), enhancing employees' security knowledge and experiences (Blythe et al., 2019) and increasing the number of phishing simulations (Sutter et al., 2022). Some studies have also examined demographic data and found that these factors had little effect on vulnerability to phishing (e.g., higher education

had no effect) (Benenson et al., 2017; Rastenis et al., 2019), while some have shown that certain parameters (e.g., age and gender) affected vulnerability (De Bona and Paci, 2020; Jevšček et al., 2019).

2.5 Phishing Training Content

To enhance resilience against phishing attacks, a common practice is to utilize content training that offers direct guidance on dealing with phishing emails, often presented by individuals perceived as experts. An illustrative example of such training is the use of "manifests," which denotes educational content dispatched prior to simulated phishing attack training. A study conducted on internet users compared the effectiveness of manifests, in-class training, and the absence of training. The findings revealed that manifests were more effective than both in-class training and no training in improving resilience to phishing attacks (Carella et al., 2017). However, studies vary in their findings on the effectiveness of these training attempts. Some studies demonstrate that content training can reduce employee susceptibility to phishing, such as a study involving 191 employees receiving content (De Bona and Paci, 2020) or a study with embedded training within simulated phishing attacks targeted at university students (Kumaraguru et al., 2008).

On the other hand, a study involving 14,000 employees reveals that embedded training during simulated phishing exercises does not enhance employees' resilience to phishing (Lain et al., 2021). These mixed results align with other studies that raise doubts about managerial approaches to combatting phishing attacks, such as using threats to emphasize the consequences of unsafe behavior (van Bavel et al., 2019), suggesting tangible rewards for managerial appreciation (Longtchi et al., 2022; Siponen et al., 2010), or sharing phishing stories with non experts (Wash and Cooper, 2018).

The use of nudging techniques to steer users to pay more attention to cybersecurity risks had gained increasing attentions in the last few years (Petrykina et al., 2021). Nudges aim to guide individuals toward predefined choices using principles of persuasion (Cialdini and Goldstein, 2004). However, detecting phishing emails is a highly technical task, which may require boosting technical capabilities rather than just paying more attention. Moreover, when individuals rely heavily on heuristics and social norms, they make predictable mistakes and repeatedly make bad decisions (Mirsch et al., 2017). Boosts promote more advantageous decisions under gain frames, while disclosure nudges promote more advantageous decisions under loss frames. Additionally, boosts are typically more effective for those who initially make suboptimal choices, according to financial research (Franklin et al., 2019).

2.6 Research Questions

The identification of a research gap, stemming from the absence of real-world experiments, compelled us to adapt our research methodology to align with the technical constraints and capabilities of an actual organization, such as the organizational mailing system. Our study involved comparing a baseline wave (without any training content) to waves that included training content, with each wave employing different phrasings sent randomly. The objective was to address the following research question:

RQ1: What is the effectiveness of phishing training content (sent before simulated phishing attacks) in real-world organizations and over time on the click rate and report rate?

H1: We hypothesize that employee vulnerability is reduced through phishing training and that phishing detection improves continuously over time. Therefore, the expectation is that the click rate decreases and the report rate increases.

Additionally, as other papers have found that common reasons to click on a phishing mail were curiosity, a match of the message to the recipient's expectations (Jevšček et al., 2019) and adding the name of the message recipient (Sahni et al., 2016), we wanted to distinguish between personalized and general phrasing to answer the following question:

RQ2: What is the effect of the phrasing of the simulated phishing attack (personalized versus general)?

H2: We hypothesize that personalized phrasing makes employees more curious, so they will click on phishing links more often.

In an analysis of the correlation between employees' data and phishing attack vulnerability, an organization's real-world study found that the participants whose job type involved the frequent use of computers but in a very specialized setting (e.g., branch workers who mainly use a single dedicated program) clicked on more links in phishing emails and took more dangerous actions than those in comparable groups. A possible explanation is that a group's greater interaction with emails may make them more suspicious of incoming emails (Lain et al., 2021). It is unknown how phishing training will affect different business units (selected branches, financial departments, and all employees) over time.

RQ3: What is the impact of phishing training across different business units?

H3: Training content has the same positive impact on employees in general and in specific business units in particular, without a significant differentiation.

Another question that remains is when is the best time to invest in phishing training? In related literature, the results indicate that over very short periods of time (10 days), there is no significant difference between training and susceptibility. In contrast, over a longer period of time (63 days), training does significantly reduce susceptibility (Jevšček et al., 2019). Given organizational limitations, we explored other short time periods to see if they differed.

RQ4: Does the timing of the training (before the actual simulated attack) affect the phishing click rate?

H4: Training is more effective if it is delivered closer in time to the phishing attack rather than farther in time from it.

In addition to timing differences, we also controlled different content phrasing. We designed our study so that the informative training content includes a general description of what phishing is, how it differs from spam, and general questions that might assist in detecting phishing, while the example-based training content includes very detailed and intuitive instructions to detect phishing by sending a simulation of a computer with a phishing mail on it and highlighting all the suspicious indications in the mail. This means that the emphasis is on what the employees need to do to detect phishing (rather than sharing a general description of what phishing is).

2.7 Measuring organizational security behavior

Traditional approaches to measuring security have primarily focused on gathering technical information related to intrusion attempts, virus logs, access requests, and traffic data (Kirlappos et al., 2015). These measurements serve as indicators for assessing the performance of technical systems. However, to explore alternative methods, information security managers have examined various security measures (Hagen et al., 2008). Their evaluations have revealed that measures related to awareness (such as training, awareness programs, user participation, and top management's commitment) and technology (such as personal passwords, redundancy of critical systems, intruder detection systems, anti-virus software, and firewalls) are more effective compared to measures involving procedures and control (e.g., nondisclosure agreements), tools and methods (e.g., incident handling), and information security policies.

Among organizational measures, awareness creation is a relatively newer and less commonly implemented approach that focuses on behavioral aspects and has been found to be highly effective (Hagen et al., 2008). On the other hand, technical-administrative measures, such as policy implementation, procedures, controls, and administrative tools, are more commonly adopted but are considered to be less effective compared to awareness creation. In essence, there exists an inverse relationship between the implementation of organizational information security measures and their effectiveness. This highlights the importance for organizations to prioritize addressing employees' behavioral aspects and promoting awareness, despite the lower implementation rates, as it is a more impactful approach.

In organizations, the most common measure for evaluating phishing waves is the "click rate", the percentage of employees who clicked on the phishing link (Alkhalil et al., 2021; Caputo et al., 2014; Carella et al., 2017; Gordon et al., 2019; Greene et al., 2018; Jevšček et al., 2019), since it gives insight into how the simulated attack impacts employees, although there seems to be a significant lack of a systematic analysis beyond click rates (Jevšček et al., 2019). Another measure that organizations use is reporting the click rate - the percentage of employees who reported on the phishing wave (Greene et al., 2018; Jevšček et al., 2019). For a better understanding of what affected their employees, some organizations also use questionnaires for assessing security awareness, such as SEBIS (Egelman and Peer, 2015).

3. Methodology

This study analyzes a set of phishing campaigns targeted at the entire organization rather than individually tailored spear phishing campaigns. Insider threats are also not included. We conducted a series of controlled experiments in an organization with nearly 5,000 employees to assess the impact of cybersecurity awareness interventions on employees' security behavior in simulated phishing attacks (waves). Our target organization was an Israeli financial institution. The participants were all employees or subcontractors of the organization. Out of the employees who work for the institution, ~45% of them work in branches (our main targeted research population), ~15% in IT, and the rest in administration/headquarters (including a particular financial department that was specifically targeted in the research).

Our experiment consisted of three phishing waves. First, all employees were sent a simulated phishing email. Then, three months later, a second email was sent using "phishing training content" before it at different times. Finally, three months later, a third email was sent by using different types of "training content phrasing".

To add credibility to our simulated phishing attacks, we used addresses that our target population would normally trust and topics that seemed realistic. The emails were distributed with an automated delivery system at midday on a regular weekday to ensure rapid user discovery. The recipients of the email were all using laptops with Microsoft Outlook (not mobile devices) and were able to view the underlying URL in the mail addresses. To ensure that the emails were sent to all employees, we worked together with the IT department and representatives of the automated delivery system at the targeted organization. This was performed with the full knowledge and approval of the target institution's human subjects and with no prior knowledge of the participants.

Participants were randomly selected from the employee population. The employees had consented to the training, but they were not warned about specific phishing simulation messages. This allowed for a more realistic and controlled experiment compared to most prior studies (where volunteers were required to opt-in).

The phishing wave phrasing was designed to be different between the rounds so that employees would not recognize them. Additionally, in each round, there were different phrasings that were sent randomly to the employees so that it would make more difficult their capability of sharing the content with their coworkers.

Employees who clicked on a phishing link were automatically directed to an internal landing page, which was a website that informed them that they had clicked on a link within a phishing simulation and that offered additional training and awareness-raising tutorials (Fig. 1).

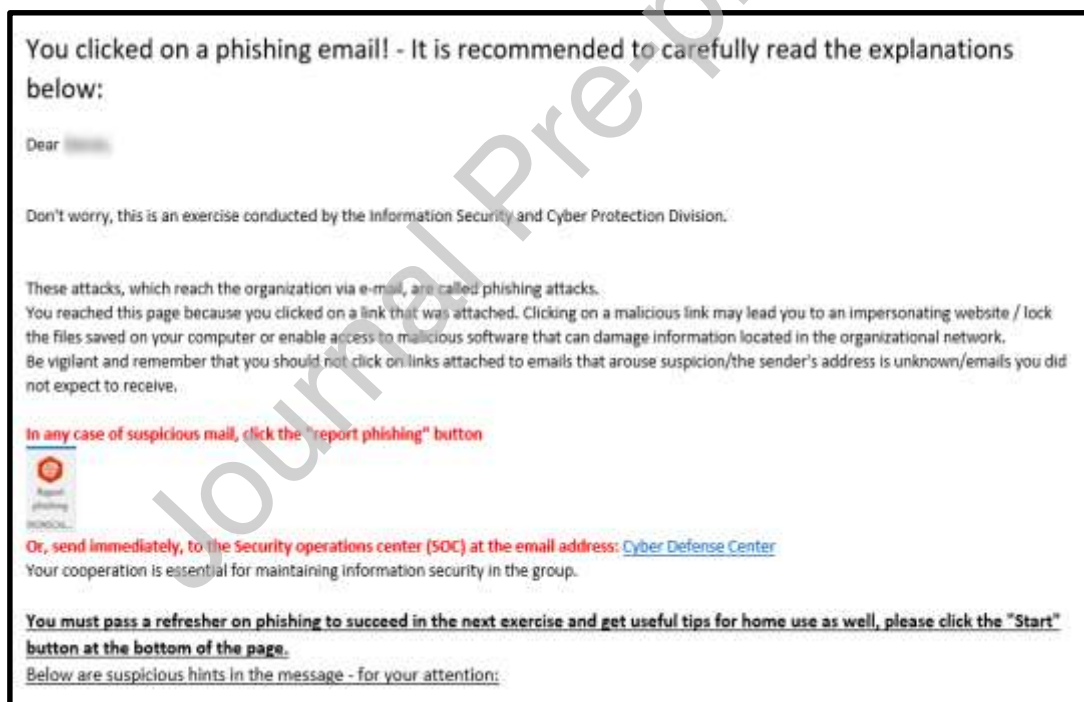


Fig. 1 - Landing page displayed after clicking on phishing link

In this paper, we consider phishing campaigns that target the entire organization rather than spear phishing campaigns that target a specific individual. Each of the simulated phishing campaigns represents a typical phishing scenario. The full details of each group and the controlled variables are described in Table 1.

3.1 Dependent variables

A controlled experiment within a single organization was conducted to examine the impact of cybersecurity awareness interventions on employees' security behavior in phishing waves. We chose a leading Israeli financial institution as our target organization. All participants worked for or were subcontractors of this organization. The dependent variables that we tested included:

(V1) Phishing Click Rate - the percentage of employees who clicked on the phishing link.

This link was combined as part of the simulated phishing attack. Among organizations, it is the most common measure for evaluating phishing waves (Alkhalil et al., 2021; Caputo et al., 2014; Carella et al., 2017; Gordon et al., 2019; Greene et al., 2018; Jevšček et al., 2019) since it gives insight into how the simulated attack impacts employees. Cybersecurity behavior and the ability to detect threats are shown by this indicator. In the literature, there is no clear percentage as a baseline for organizations, and the range varies from a few percent to dozens (Jevšček et al., 2019). Additionally, there seems to be a significant lack of a systematic analysis beyond click rates (Jevšček et al., 2019).

(V2) Reporting Click Rate - the percentage of employees who reported on the phishing wave.

This was collected by using a designated "Outlook" button or by forwarding the mail to the IT department. Usually, organizations focus only on their phishing click rate, but one number cannot tell the whole story, so the reported rate is also crucial (Greene et al., 2018). When employees know what to look for and how to report it, they can significantly contribute to the organization's security (Greene et al., 2018; Jevšček et al., 2019). Additionally, users may be the only option for catching phishing emails that receive technological defenses (Frauenstein and Von Solms, 2013; Greene et al., 2018; Jain and Gupta, 2022; Jevšček et al., 2019). Reporting buttons allow email security providers to detect attacks early and inform other potential victims before the attack spreads (Wang and Song, 2021). It has been noted in the literature that the results of the two dependent variables differ greatly. An example would be a study that recorded a 46% click rate for phishing and a 2% click rate for reporting (Jevšček et al., 2019).

Table 1 – Total number of employees who received phishing emails per wave, and a wave overview according to the variables, research groups and simulated phishing phrasing

Wave	Total employees	Independent variables	Treatment group	Simulated phishing phrasing
1st round - July 2021	N = 4650	Simulated phishing phrasing - personalized versus general	None	Personalized: Holiday gifts (Fig. 2), Parking arrangements (Fig. 12) General: COVID-19 (Fig. 3)
2nd round - October 2021	N = 4750	Timing of sending training content before the phishing wave (2/7/14 days)	3 branches (N = 76) - different training content timings, informative phrasing	Changing employees' badge photo (Fig. 13), Malicious email detection software (Fig. 14)

3rd round - January 2022	N = 4403	A) Simulated phishing phrasing - personalized versus general B) Timing of sending training content before phishing wave: 2/7/14 days C) Training content phrasing: informative, example-based	A) 3 branches (N = 76) - same timing as 2nd round, example-based phrasing B) 6 New branches (N = 142) - permutation of different training content timings (2/7/14 days), different training phrasing (informative/example-based)	Personalized: COVID-19 green pass (Fig. 15) General: ice cream coupon (Fig. 16), Financial lectures (Fig. 17)
--------------------------	----------	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------

* Treatment Group N was small compared to the total number of employees (~4,750), but we also analyzed the general results.

3.2 Phishing wave phrasing: personalized versus general

Round one of our phishing waves targeted 4,650 organizational members, which was divided into 9 branches, employees who clicked in previous rounds, and all the rest (Table 2). The purpose of this round was to have a baseline before we altered the variables in rounds two and three and to examine how different wave phrasing types affect the participants. To increase the credibility of our phishing emails, three different types of phishing emails were randomly sent to our participants.

Table 2 - Round 1: phishing wave phrasings

Phishing Mail Content	Holiday gifts (personalized) (Fig. 2)	Parking arrangements (personalized)	COVID-19 (general) (Fig. 3)
Mail Description	A routine email that employees receive before the holidays with a catalog of gifts to choose from. The sender was a fictitious "gift purchase" site, and the mail included a gift image.	An email that informed employees of changes in parking policies at different branches. The sender was a fictitious IT notifications address associated with HR.	An informative email that provided guidelines on how to protect yourself from COVID-19 in public places and a "spreading map". The sender was a fictitious address related to the Ministry of Health, with its logo attached.
Sent	1651 (35.5%)	1499 (32.25%)	1500 (32.25%)

* The sent column represents the distribution of each email phrasing to the mentioned number of employees.

Managers in the organization reviewed these emails to ensure that they were consistently deceptive and attractive.

Summer benefits for all company employees



Dear [redacted],
 Summer is here! And with it a selection of gifts and attractive treats for the [redacted] contractors and employees. Chef restaurants, pampering spas, stunning beach kits, winery tours, discounted entrances to the municipal water park, gadgets and more.
 With us, you will find all the coolest gifts and products you could think of...
In order to get these benefits, please enter the code in the coupon code field: [redacted]
 To the products catalogue website, click [here](#).



Regards,
 Purchasing team

Fig. 2 - A screenshot of the simulated phishing mail that contains personalized phrasing. This type of phishing contains the name of the recipient. The subject of the email is Holiday gifts

A green underline marks the independent variable “simulated phishing phrasing” - personalized phrasing (name and interest). Employee expectations are aligned with the timing (holidays are coming, so employees should select a gift accordingly). Through the phishing link, employees are offered a self-benefit (choosing a gift).

Updated guidelines - protection against COVID virus



Dear employees,
 In light of the possibility of another wave of the virus, updated guidelines on protection against corona virus in public places are attached.
 For your health - your vigilance is most important!
To view the virus spread map and the new guidelines click [here](#).

Wishing you good health,
 Minister of Health - Public Health Division



Fig. 3 - A screenshot of the simulated phishing mail that contains general phrasing. This type of phishing does not contain the name of the recipient. The subject of the email is COVID-19

A red underline marks the independent variable “simulated phishing phrasing” - general phrasing (name and interest). Through the phishing link, employees are not offered a self-benefit but only general information.

3.3 Training timing before the simulated attack

Three months after our first round of emails, we sent a second round of phishing emails to all participants (in this round, $N = 4,750$, as seen in Table 3). In this round, we targeted and affected three specific branches, which received additional “informative phishing training” at different times prior to the wave (2/7/14 days), while all other employees received it far in advance (60+ days). Informative phrasing was used in the training, as shown in Fig. 4. Our goal was to determine whether and how timing affects the chance of being phishing.

In this round, as in the previous round, we customized the phishing wave email phrasing based on the knowledge of the target organization. To increase the credibility of the simulated phishing emails, two different types of phishing emails were randomly sent to our participants (this was also useful to examine how different wave phrasing types affected the participants):

Table 3 - Round 2: phishing wave phrasings

Phishing mail content	Changing employees' badge photo	Malicious email detection software
Mail description	An instruction email was sent on updating an employee's badge photo, with a warning that employees who do not renew it will be sanctioned. The sender was the HR department.	An announcement email was sent about new cyber security software that the employees must install (including guidelines). The sender was a fictive department (which does not exist in the organization)
Sent	2371 (49.9%)	2379 (50.1%)

3.4 Phishing wave phrasing and training timing before the simulated attack

Three months after our second round of emails, we sent a third round of phishing emails to all participants (in this round, $N = 4,403$, as seen in Table 4). In addition to the different timings before the wave (2/7/14 days), we also used different training phrasings during this round. In the previous round (round two), we presented “informative phrasing” (Fig. 4), but here, we used an alternative “example-based phrasing” (Fig. 5) that provides an example with step-by-step instructions for detecting phishing mail (rather than general instructions).



Dear employees,

In order to avoid activities that could harm the financial institution or its customers, and as part of the social, value and strategic responsibility we have taken on, attached a number of information security issues that you encounter as part of your work routine in the organization.

What is phishing?

Phishing is an attempt to steal sensitive information in which the attacker impersonates a legitimate person/site and is done by sending an email with a link or a malicious file. The attacker uses impersonates email accounts to spread malicious links or attached files, that may commit identity theft, theft of account and personal information, malware injection into the corporate network and more.

What is the difference between phishing and spam?

Spam is unwanted e-mail, sometimes of a commercial nature (advertisements for example).

Phishing mail is sent fraudulently and with malicious intent to obtain sensitive information or to perform malicious actions in the organization. In any case of suspicion, this e-mail should be reported using the **"Report Button"**.

Therefore, we would like to emphasize again,

When you receive an e-mail, please pay attention and look for the following signs:

- Do you expect to receive the email?
- Check - what is the source of the email? Is it internal or external to the organization?
- Is the refer to click the link or opening of the attachment unusual, irrational, or are there feelings of pressure/fear?
- Is there a match to the writing style: spelling mistakes and wording from the recipient of the email?


The **"Report Phishing"** button is designed to report suspicious messages received in your inbox. If a suspicious message is received (and the report button has not yet been added) it should be forwarded, in accordance with the instructions, to the Cyber Protection Center.

Maintaining vigilance, a high level of awareness and continuous cooperation help protect the organization's systems.

Thank you for your cooperation,
Information Security and Cyber Defense Division

Fig. 4 - Informative training content

Red flags in social engineering



The responsibility for protecting privacy is in your hands.

In any question and concern about an information security incident You can contact the Cyber Protection Center.

Thank you for your cooperation,
Information Security and Cyber Defense Division

Pay attention!

In light of the situation, there is an increase in online fraud attempts of various kinds. Among other things, attempts to impersonate the organization's technical support personnel in order to obtain connection details for the organization's systems. Do not transfer the login and/or identification data.

If there is a problem or technical problem with the computer, you can contact:

Chat – available to you at any time
You can talk to him by sending a message

WhatsApp number - [redacted]
Scan QR Code with your smartphone's camera:

Add the chat to contact list: [QR Code] | Send a direct message to the chat: [QR Code]




Fig. 5 - Example-based training content

Our goal was to determine if and how different phrasings affect the likelihood of being phished (along with the timing in which the training content is sent). As in the previous round, we customized the phishing wave emails based on the target organization's knowledge.

To increase the credibility of our phishing emails, three different types of phishing emails were randomly sent among our participants (this was also useful to examine how different wave phrasing types affect the participants).

Table 4 - Round 3: phishing wave phrasings

Phishing mail content	COVID-19 green pass (personalized)	Ice cream coupon (general)	Financial lectures (general)
Mail description	An instruction email was sent on how to renew the COVID-19 green pass, which emphasized that employees who do not renew it may be sanctioned. The sender was a fictitious address related to the Ministry of Health, with its logo attached.	An announcement email was sent on the opening of a new ice cream shop, with a time-limited coupon for 3 kg of ice cream for the price of 1 kg. The sender was the new shop's customer service.	An email was sent that offered a free subscription to a financial lecture website for the organization's employees. The sender was the finance team.
Sent	1413 (34.5%)	1313 (32.1%)	1369 (33.4%)

4. Results

Overall, the Phishing Click Rate (V1) and Reporting Click Rate (V2) improved across the three waves in which we sent simulated phishing attacks. Thus, organizational awareness activities prior to phishing waves seem to be effective. Additionally, we noticed that mail phrasing matters.

To measure the in-branch variance (e.g., due to external or unexplained reasons), we calculated the standard deviation across all participating branches (Table 5) and found it to be 7.43%. In the relevant charts, we visualized this parameter as a vertical line.

Table 5 - Standard deviation of the branches in July 2021 (baseline wave)

Branch	1	2	3	4	5	6	7	8	9	STD
Sample size	26	18	32	15	21	11	33	21	41	
Phishing click-through rate	12%	33%	28%	27%	29%	18%	21%	19%	34%	7.43%

Due to the correlation between selected branches and all employees in the July wave, we used this standard deviation. Other waves had similar standard deviations (in October 8.61% and January 6.48%). The overall results of our measures of the “phishing click rate” and “reporting click rate” are presented in Table 6, depending on different independent variables. The results support H1 that employee vulnerability is reduced through phishing training, and phishing detection improves continuously over time.

Table 6 - Overall results: measures per wave per independent variable

	Wave 1 (July 2021)	Wave 2 (October 2021)	Wave 3 (January 2022)
--	-----------------------	--------------------------	--------------------------

Independent variables *	Simulated phishing phrasing	Training timing	Simulated phishing phrasing Training timing Training phrasing
Treatment group	0	76	218
Sent (N)	4650	4750	4403
Phishing click rate	1165 (25%)	595 (13%)	320 (7%)
Reporting click rate	371 (8%)	599 (13%)	785 (18%)

* The independent variables were not applied to all employees but only to the treatment group.

Overall, analyzing phishing click rate (measure 1), the number of employees who were phished significantly decreased from wave to wave by nearly half ($\chi^2 = 595.588$, $p < 0.0001$, $DF = 2$, $N = 13,803$). This means that fewer employees clicked on the phishing simulated attack over time, as seen in Fig. 6:

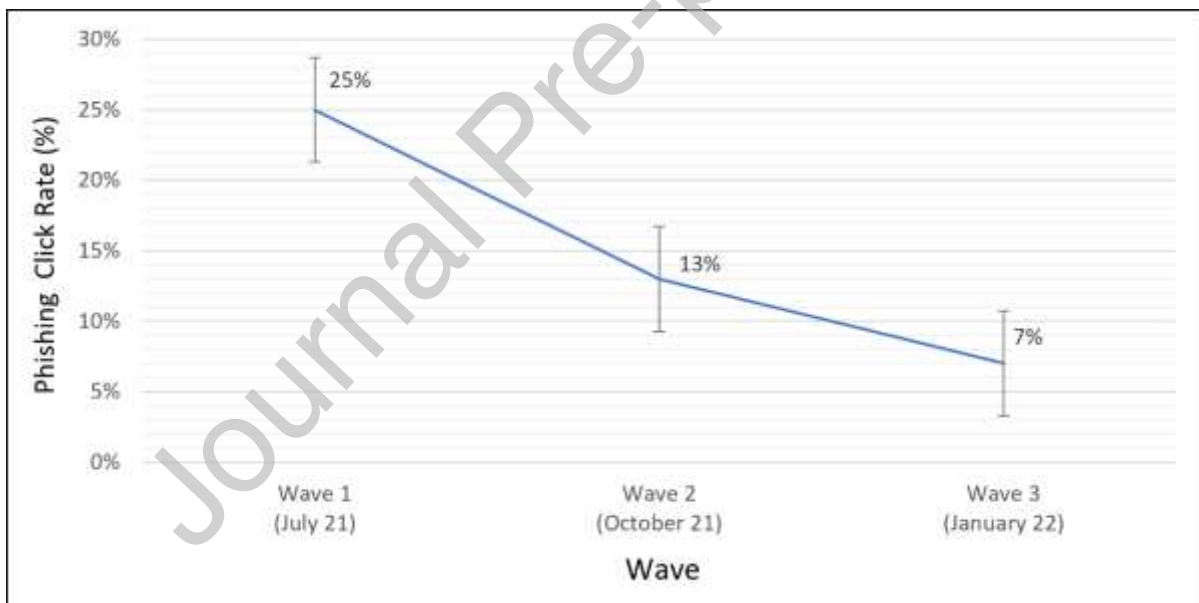


Fig. 6 - Overall phishing click rate: employees click less on the simulated phishing campaign over time

Overall, analyzing reporting click rate (measure 2), the number of employees who reported on phishing increased from wave to wave by nearly double ($\chi^2 = 197.794$, $p < 0.0001$, $DF = 2$, $N = 13,803$). That is, more employees reported the phishing simulated attack to the IT department over time, as seen in Fig. 7.

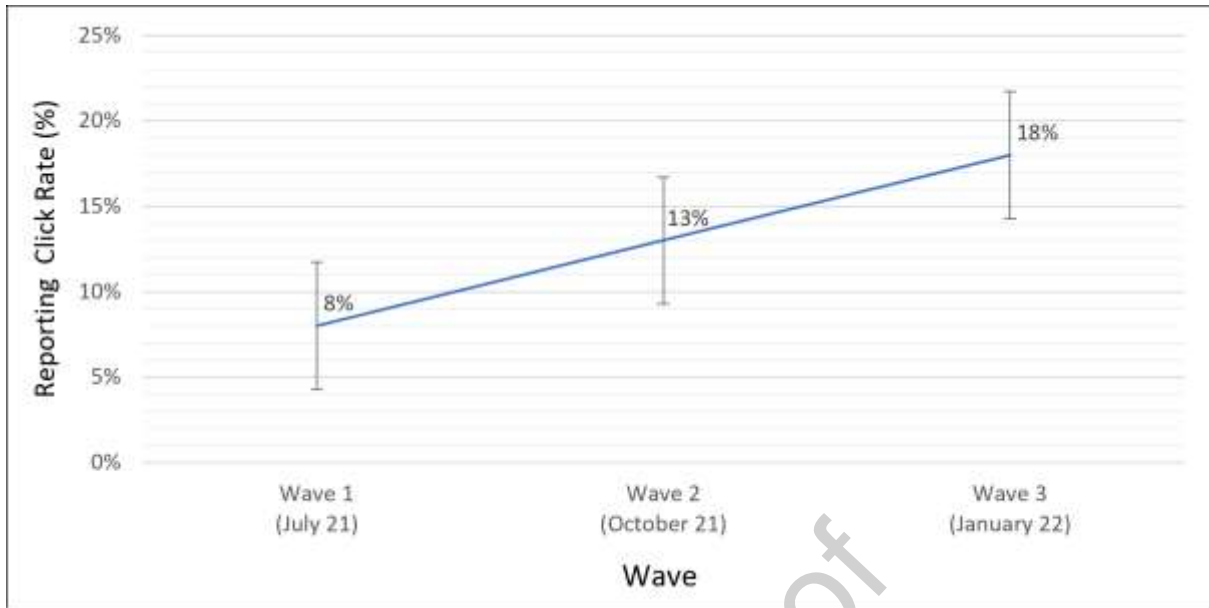


Fig. 7 - Overall reporting click rate: employees report more to IT on the simulated phishing campaign over time

Analyzing the overall results based on wave phrasing shows interesting findings (Table 7) about the measures per wave for each phishing mail phrasing (personalized versus general):

Table 7 - Results per wave phrasings (without October, as the phrasings there were mixed)

	Wave 1 (July 2021)			Wave 3 (January 2021)		
Phishing mail content	Holiday gifts	Parking arrangements	COVID-19	COVID-19 green pass	Ice cream coupon	Financial lectures
Phrasing	Personalized	Personalized	General	Personalized	General	General
Sent	1651	1499	1500	1598	1374	1431
Phished	492 (29.8%)	518 (34.5%)	155 (10.3%)	189 (11.8%)	65 (4.7%)	66 (4.6%)
Reported	148 (8.9%)	101 (6.7%)	122 (8.1%)	367 (22.9%)	234 (17%)	184 (12.8%)

In the first wave (July 2021), personalized phrasing was 210% larger than general phrasing, which indicates statistical significance ($\chi^2 = 265.001$, $p < 0.0001$, $DF = 2$, $N = 4,650$). This means that employees were phished more with personalized phrasings than with general phrasings. This significant difference was also evident in January ($\chi^2 = 77.389$, $p < 0.0001$, $DF = 2$, $N = 4,403$), in which personalized phrasing was 150% larger than general phrasing (Fig. 8 - Phishing click rate by phrasing: employees click more (Y-axis) on a personalized phrasing simulated phishing mail rather than on a general phishing mail, per wave (X-axis)). Our explanation is that employees were more likely to respond to the mail since their names were mentioned explicitly, and the content was tailored to their interests. October was filtered out as the phrasings were not aligned with these definitions but were a mix of them.

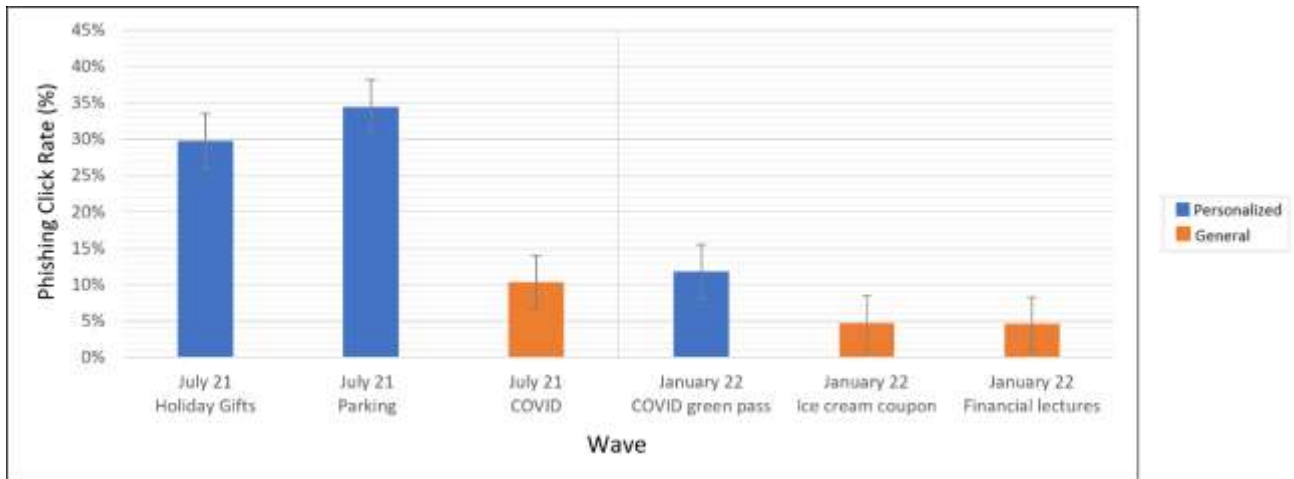


Fig. 8 - Phishing click rate by phrasing: employees click more (Y-axis) on a personalized phrasing simulated phishing mail rather than on a general phishing mail, per wave (X-axis)

This result supports H2 that personalized phrasing makes employees more curious, so they click on phishing links more often.

Taking a deeper dive into these overall results based on wave phrasing (personalized versus general) reveals interesting results. Unlike the “phishing click rate” measure, in the “reporting click rate” measure, the wave phrasings did not follow a consistent pattern (personalized versus nonpersonal). In the first wave (July 2021), there were no significant differences in the reporting rates of personalized versus general phrasing ($\chi^2 = 5.37$, $p = 0.067 > 0.05$, $DF = 2$, $N = 4,650$), and in the third wave (January 2022), personalized phrasing led to more reporting ($\chi^2 = 53.52$, $p < 0.0001$, $DF = 2$, $N = 4,403$), as shown in Fig. 9 - Reporting click rate by phrasing: no significant difference existed in the reporting rate (Y-axis) per simulated phishing phrasing among the waves (X-axis).

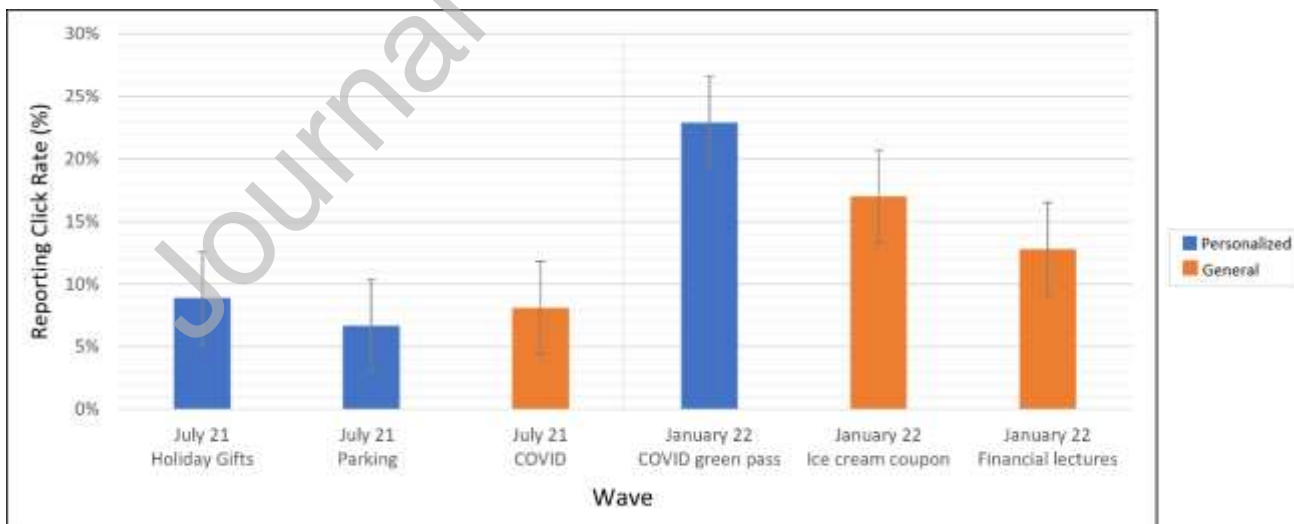


Fig. 9 - Reporting click rate by phrasing: no significant difference existed in the reporting rate (Y-axis) per simulated phishing phrasing among the waves (X-axis)

Additionally, we examined different groups of employees within the organization. Even though we were interested in selected branches (since they were closed groups) during the research, we wanted to determine how a broader population (a financial department) would react to the simulated phishing attacks. We noticed a significant difference between employee departments (selected branches, financial department, and all) on July

21 ($\chi^2 = 42.756$, $p < 0.0001$, $DF = 2$, $N = 4,650$) and October 21 ($\chi^2 = 6.611$, $p = 0.0366 < 0.05$, $DF = 2$, $N = 4,750$) but not on January 22 ($\chi^2 = 3.543$, $p = 0.17 > 0.05$, $DF = 2$, $N = 4,403$), as shown in Fig. 10.

The conclusion is that different populations might respond differently to the interventions and education that an organization offers; it is worthwhile for the organization to examine itself and see if it is moving in the right direction with the specific populations and not just on a general level. Additionally, the training content over time affected the business units differently as some of the units improved while others (financial department) did not improve, which does not support H3 that training content has the same positive impact on employees in general and per a specific business unit in particular, without a significant differentiation.

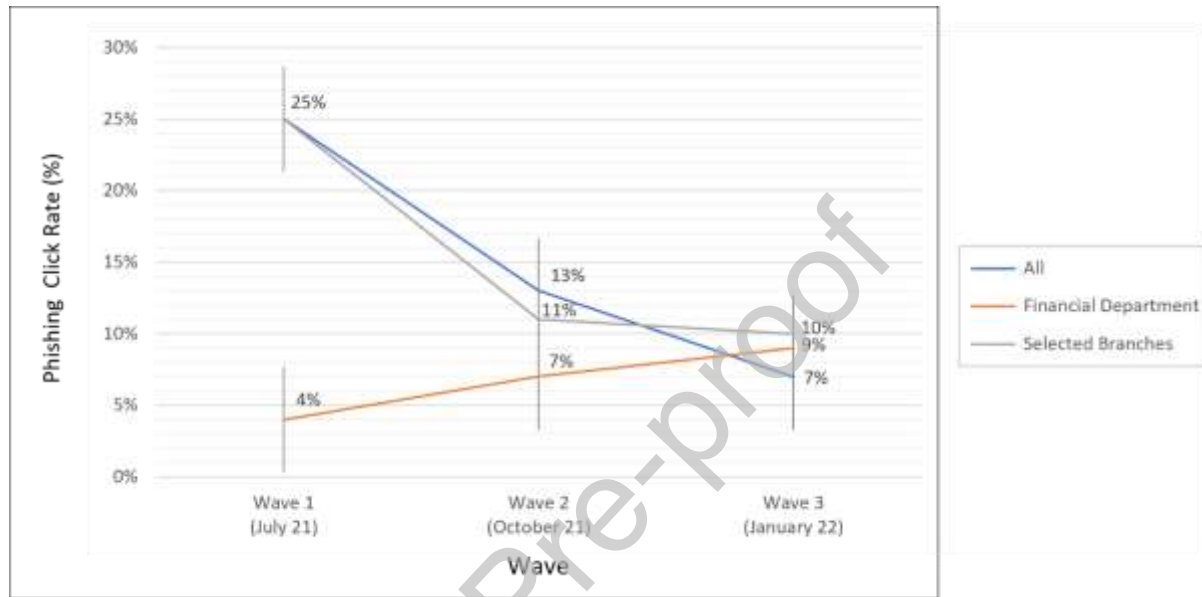


Fig. 10 - Phishing click rate by business unit: significant differences exist in the phishing click rate (Y axis) among different business units in the July 21 and October 21 waves

In analyzing the training timing, the number of days that the training was sent before the wave was not significantly affected on October 21 ($\chi^2 = 1.185$, $p = 0.756 > 0.001$, $DF = 3$, $N = 4,256$) and was barely affected on January 22 ($\chi^2 = 9.239$, $p = 0.026 < 0.05$, $DF = 3$, $N = 4,051$). We expected that when the time of the training content before the wave was shorter, fewer employees would click, but this was not the case. We suspect that the reason that this happens is that 2/7/14 days are all short terms, so the differences are not significant (possibly longer terms such as over a month would behave differently). We did notice a correlation between the short- and long-term (2 and 14 days), whereas the mid-term (7 days) behaved oppositely, as seen in Fig. 11. This result does not support H4 that training is more effective if it is delivered closer in time to the phishing attack rather than farther in time from it.

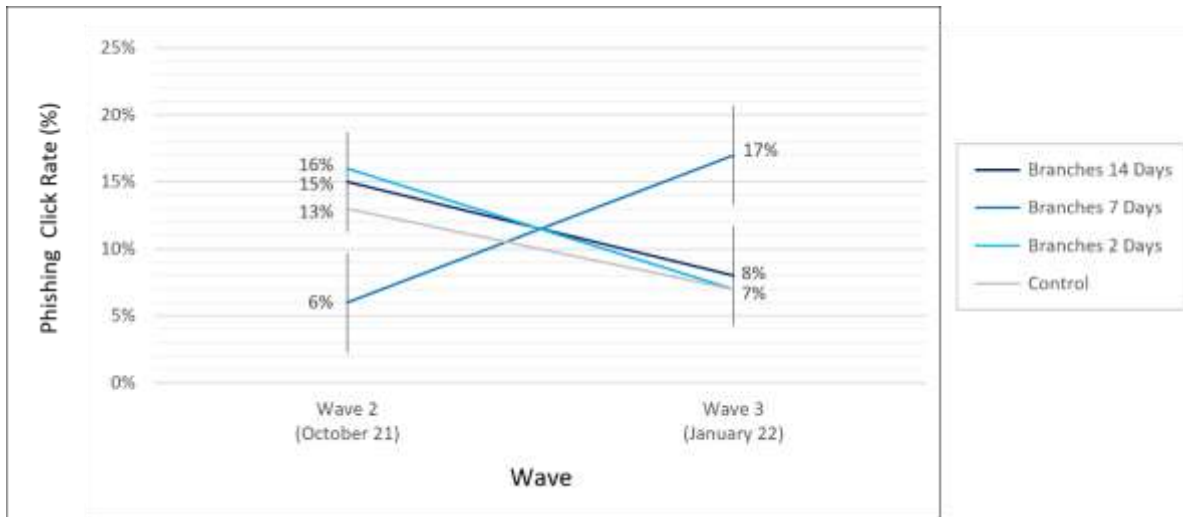


Fig. 11 - Phishing click rate by training timing: no significant differences exist in the phishing click rate (Y-axis) depending on the training timing among the waves (X-axis)

We also tested the training content phrasing and observed a non-significant increase ($\chi^2 = 4.199$, $p = 0.122 > 0.05$, $DF = 2$, $N = 4,051$) in phishing detection (lower click rate) in the example-based training (phishing mail example highlighted with detection methods) compared to the informative training (general description and information about phishing), as seen in Fig. 18. A possible theory to explain that is that this may be because we were able to conduct this comparison in only one wave (January 22) on a limited treatment group.

5. Discussion

When examining our findings, it becomes evident that various measures, including click rates and reporting, displayed improvement across the three waves of simulated phishing attacks. Notably, the phrasing used in the email (personalized versus general) had a significant impact on the click rate, with the personalized emails receiving 1.5 to 2 times more clicks compared to the general ones. However, the phrasing did not exert a noticeable influence on the reporting rate. This discrepancy might be attributed to the expectation among employees that clicking on phishing links would provide immediate benefits, such as redirecting them to pages tailored to their interests. On the other hand, factors like individual personality and habits could have a bearing on whether an employee reports risks, without a direct correlation to the email's phrasing itself.

These findings support previous literature with improved ecological validity. Lab studies have shown the effect of phrasing phishing messages (Benenson et al., 2017) or the effect of including the recipient's name in the subject line (Sahni et al., 2016) on the phishing click rate. Our study provided support to both these lab findings in a large-scale real-world organization (RQ2), showing that personalized phrasing in phishing emails, including the recipient's name and content aligned with their expectations, increases the probability of employees to click on phishing emails. Our findings also provide additional nuance: first, we show that the impact of simulated phishing email attacks have an accumulated effect beyond the first simulated attack, and that the effect is still meaningful after the second one. Moreover, we show that distinct behaviors were observed among different business units.

When evaluating the effect of content training, we show that both the timing of the training (before the real simulated phishing email) and the content of the training (informative versus example-based) did not significantly affect the click-rate. There is a chance that the differences are minor because the timing is within a short range (2 to 14 days), and this is supported by a paper that indicates there is no significant difference in phishing susceptibility based on training over a very short period of time (10 days), but over a longer time (Coronges et al., 2012). We didn't find a related literature comparing similar phrasings, but it might not be sufficiently different to make the employees behave differently. These results contradict prior literature and common industry practices that suggest providing content to users to prevent phishing attacks (Carella et al., 2017).

Several implications of the results of this study apply to organizational information security officers. Organizations can gain a better understanding of how employees respond to security threats and what should be improved by measuring security behavior in practice. First, it is beneficial to train employees about phishing-type attacks and measure whether the training reduces click rates. Although it is unrealistic to expect employees to detect all phishing emails, lowering click rates is the most effective way to improve resilience to phishing attacks. Therefore, we suggest investing in a data-driven approach to focus on what works, adjust the training per level or department within the organization, and in addition to the key performance indicators (KPI) phishing Click Rate (V1), we also recommend analyzing the other KPI Reporting Click Rate (V2). We noticed that these KPIs are not correlated, thus one effort may not necessarily improve both, as each may be affected by different parameters.

Our methodology holds significant implications for organizational information security officers who are responsible for evaluating anti-phishing tools. By employing clearly defined measures, we address the challenge of assessing the relationship between security measures and their impact on business outcomes (Ashenden and Lawrence, 2016). In particular, we emphasize the importance of how outcomes are evaluated and measured. Drawing from our partnership with an organization, we have observed that information security officers should acquire knowledge in data science and experimental design to effectively design interventions and measure their effectiveness. Specifically, assessing the efficacy of anti-phishing campaigns proves challenging without establishing a baseline for phishing click rates and other relevant measures, as well as understanding the variability inherent in these measures.

Personalized phishing emails are more likely to cause employees into clicking on their links than general ones, as it is expected from people to trust messages that are addressed to them specifically. While analyzing the results, we suggest considering the effect of the email phrasing on the KPIs, adjusting the difficulty accordingly, and educating employees about common suspicious phrasings. Incorporating technical solutions alongside awareness efforts can be useful. For example, phishing emails that include employee interests and that use personalized phrasing were found to cause more clicks on the phishing links, so it might be helpful to mark suspicious indicators in the mail itself. A feasible example to implement can be if the sender's email address is within or outside the organization, as "knowing the sender" influences click-through rates (Benenson et al., 2017).

Our findings also point the way to several organizational policies that may benefit the organization's security. IT departments should give attention to every phishing email reported, since an active attack spreads quickly to more employees, and additional attacks may follow. In addition, promoting security champions in the

organization was found useful in the literature (Becker et al., 2017), and reporting on phishing mails can be one of the options used to do this. We also suggest that organizations can lower their efforts in scheduling the best time to send the training before the simulated phishing attack, since it does not significantly affect the training's effectiveness. It is an unexpected finding since the literature indicates that people are generally more capable of detecting phishing emails when they are expecting them (Briggs et al., 2017). Hence we expected to receive better detection for shorter timeframes, though it is important to remember that different sending frequencies and timeframes longer than 14 days were not investigated.

Our findings stress out the importance of diverse engagement in phishing campaigns, using additional creative phrasings, such as social nudging. As phishing awareness training requires careful tailoring to the organization, and since reporting back to employees with statistics (e.g., the number of employees that identified a new phishing threat) can be useful (Greene et al., 2018), such engagement can be solicited, for example, through positive competition within different organizational departments, depending on previously simulated phishing attack results (e.g., "X% of your department colleagues detected the previous phishing simulated attack, and department Y reported the most on it"). This can make the training experience interesting and improve the motivation to mitigate cybersecurity threats.

In terms of ecological validity, organizations should ask what the click rates and reporting rates represent. A high click rate, for instance, may indicate more than a phishing vulnerability. For example, a study found a participant who knew it was an exercise and clicked on a suspect link or attachment out of curiosity. It means that if the reporting process is onerous or unknown to the participants, reporting rates would not measure the number of users who "caught" the phish but rather the number who were willing to navigate the organization's reporting process (Greene et al., 2018). Ultimately, it matters not how employees perform training exercises but whether they know the latest scams and how they respond to phishing emails. The training awareness programs are not the goal but the means. A data-driven approach in training and simulated phishing attacks is essential for identifying potential risks and allocating resources accordingly so that in actual phishing attacks, employees won't click on suspicious links but report them.

5.1 Limitations and future directions

This research has some limitations that the reader should be aware of. First, the case study was conducted in a financial institution in Israel. The results may be different in other settings or cultural backgrounds. Second, conducting a real-world experiment has constraints. Specifically, the constraints in our case are limited phrasing options - the training content phrasings that we used were restricted due to internal limitations (e.g., some of our suggested phrasings such as social nudging was not approved); Limited technology - the system that generates the attacks could not mix training or phrasings within a group but only between different groups (e.g., for the same group, we were not able to send to a few employees a training 3 days before a wave and for others in the same group, 14 days before); Limited resources led to small treatment groups (2nd wave $N = 76$, 3rd wave $N = 218$, out of nearly ~5,000 employees) and a limited number of targeted groups; Limited data - we did not analyze employee-level data because of privacy limitations. As we did not have access to raw data, we had to use aggregative data, excluding demographic data (e.g., male sex and gender) and employee data (e.g., profession and seniority). Having raw data collected could be used to create machine learning models and to analyze invariable

correlations more effectively (e.g., by avoiding situations where external factors hinder internal factors, as data are noisy in the real world).

Accordingly, a real-world experimentation is the most reliable way to conduct research because it utilizes authentic processes, useful collections, predefined measurements, and independent variables aimed at improving the organization's KPIs and real needs. Further research should combine additional raw data (including demographics) that can be used to analyze and build a robust cybersecurity model for organizations. An improved monitoring environment is necessary to obtain additional information about the campaigns that can be used to maximize an organization's objectives.

5.2 Conclusions

In this study, we have conducted an experiment at a large Israeli financial institution to test its resilience to simulated phishing attacks. This study contributes knowledge on how organizations respond to phishing attacks in real-world settings. As not all employees can identify phishing emails and since the potential damage in such cases is significant, employee detection is a key challenge. Training to prevent phishing attacks is therefore beneficial, although some parts (e.g., phishing simulations) are more effective than others. The novelty of our study lies in its empirical analysis of real-world organizational settings. While the two decades of phishing training had produced effective interventions, evaluating them in the real world had led to new conclusions, which go beyond what evaluation in the lab or with a student population can teach. Our conclusions are particularly relevant to organizations that have a large and diverse workforce. real-world organizational study provides, on the one hand, validity for the variables and a more realistic perspective on organizational behavior and processes, but on the other hand, it takes time to reach significant numbers and to clean up the additional side-effects, so one can accurately describe the study phenomenon. Considering the worldwide need for increased prevention of phishing attacks, it will be necessary to continue the effort of this research, to refine characteristics and factors that can reduce the click-rate of phishing attacks.

6. References

- Alkhalil Z, Hewage C, Nawaf L, Khan I. Phishing Attacks: A Recent Comprehensive Study and a New Anatomy. *Frontiers in Computer Science* 2021;3. <https://doi.org/10.3389/fcomp.2021.563060>.
- Alotaibi M, Furnell S, Clarke N. Information security policies: A review of challenges and influencing factors. 2016 11th International Conference for Internet Technology and Secured Transactions, ICI-TST 2016 2017:352–8. <https://doi.org/10.1109/ICITST.2016.7856729>.
- Amankwa E, Looock M, Kritzing E. Establishing information security policy compliance culture in organizations. *Information and Computer Security* 2018;26:420–36. <https://doi.org/10.1108/ICS-09-2017-0063>.
- Ashenden D, Lawrence D. Security Dialogues: Building Better Relationships between Security and Business. *IEEE Security and Privacy* 2016;14:82–7. <https://doi.org/10.1109/MSP.2016.57>.
- van Bavel R, Rodríguez-Priego N, Vila J, Briggs P. Using protection motivation theory in the design of nudges to improve online security behavior. *International Journal of Human Computer Studies* 2019;123:29–39. <https://doi.org/10.1016/j.ijhcs.2018.11.003>.
- Beautement A, Becker I, Parkin S, Krol K, Sasse A. Productive Security : A Scalable Methodology for Analysing Employee Security Behaviours This paper is included in the Proceedings of the Productive Security : A scalable methodology for analysing employee security behaviours 2016:253–70.
- Beautement A, Sasse MA. The Compliance Budget: The Economics of User Effort in Information Security. 2009.
- Becker I, Parkin S, Sasse MA. Finding Security Champions in Blends of Organisational Culture. *Proc USEC 11* 2017. <https://doi.org/10.14722/eurosec.2017.23007>.
- Benenson Z, Gassmann F, Landwirth R. Unpacking Spear Phishing Susceptibility. Springer International Publishing: 2017.
- Bitton R, Boymgold K, Puzis R, Shabtai A. Evaluating the Information Security Awareness of Smartphone Users. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Pp 1-13) 2019.
- Blythe JM, Coventry L. Costly but effective: Comparing the factors that influence employee anti-malware behaviours. *Computers in Human Behavior* 2018;87:87–97. <https://doi.org/10.1016/j.chb.2018.05.023>.
- Blythe JM, Coventry L, Little L. Unpacking security policy compliance: The motivators and barriers of employees' security behaviors. *SOUPS 2015 - Proceedings of the 11th Symposium on Usable Privacy and Security* 2019:103–22.
- Briggs P, Jeske D, Coventry L. Behavior Change Interventions for Cybersecurity. Elsevier Inc.; 2017. <https://doi.org/10.1016/B978-0-12-802690-8.00004-9>.
- Butavicius M, Parsons K, Pattinson M, McCormac A. Breaching the Human Firewall: Social engineering in Phishing and Spear-Phishing Emails. *ArXiv Preprint ArXiv:160600887* 2015.
- Caputo DD, Pfleegeer SL, Freeman JD, Johnson ME. Going spear phishing: Exploring embedded training and awareness. *IEEE Security and Privacy* 2014;12:28–38. <https://doi.org/10.1109/MSP.2013.106>.
- Carella A, Kotsoev M, Truta TM. Impact of security awareness training on phishing click-through rates. 2017 IEEE International Conference on Big Data (Big Data), Boston, MA: IEEE; 2017, p. 4458–66. <https://doi.org/10.1109/BigData.2017.8258485>.
- Cialdini RB, Goldstein NJ. Social Influence: Compliance and Conformity. *Annu Rev Psychol* 2004;55:591–621. <https://doi.org/10.1146/annurev.psych.55.090902.142015>.
- Cisco, 2017. Annual Cyber Security Report 2017. Cisco 110. <https://learningnetwork.cisco.com/s/article/cisco-2017-annual-cybersecurity-report-pdf>
- Conway D, Taib R, Harris M, Berkovsky S, Yu K, Chen F. A Qualitative Investigation of Bank Employee Experiences of Information Security and Phishing. *SOUPS(Vol 2017, Pp 115-129)* 2017.
- Coronges K, Dodge R, Mukina C, Radwick Z, Shevchik J, Rovira E. The Influences of Social Networks on Phishing Vulnerability. 2012 45th Hawaii International Conference on System Sciences, Maui, HI, USA: IEEE; 2012, p. 2366–73. <https://doi.org/10.1109/HICSS.2012.657>.
- Das A, Baki S, El Aassal A, Verma R, Dunbar A. SoK: A Comprehensive Reexamination of Phishing Research from the Security Perspective. *IEEE Communications Surveys and Tutorials* 2020;22:671–708. <https://doi.org/10.1109/COMST.2019.2957750>.

- De Bona M, Paci F. A real world study on employees' susceptibility to phishing attacks. *ACM International Conference Proceeding Series, Association for Computing Machinery*; 2020. <https://doi.org/10.1145/3407023.3409179>.
- Downs JS, Holbrook MB, Cranor LF. Decision Strategies and Susceptibility to Phishing. *Proceedings of the Second Symposium on Usable Privacy and Security* (Pp. 79-90): 2006.
- Egelman S, Cranor LF, Hong J. You've Been Warned: An Empirical Study of the Effectiveness of Web Browser Phishing Warnings. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Pp. 1065-1074): 2008.
- Egelman S, Peer E. Scaling the security wall: Developing a security behavior intentions scale (SeBIS). *Conference on Human Factors in Computing Systems - Proceedings 2015;2015-April:2873–82*. <https://doi.org/10.1145/2702123.2702249>.
- Franklin M, Folke T, Ruggeri K. Optimising nudges and boosts for financial decisions under uncertainty. *Palgrave Communications* 2019;5:1–13. <https://doi.org/10.1057/s41599-019-0321-y>.
- Frauenstein ED, Von Solms R. IFIP AICT 406 - An Enterprise Anti-phishing Framework. vol. 6. Springer Berlin Heidelberg: 2013.
- Gordon WJ, Wright A, Aiyagari R, Corbo L, Glynn RJ, Kadakia J, et al. Assessment of Employee Susceptibility to Phishing Attacks at US Health Care Institutions. *JAMA Network Open* 2019;2. <https://doi.org/10.1001/jamanetworkopen.2019.0393>.
- Greene K, Steves M, Theofanos M. No phishing beyond this point. *Computer* 2018;51:86–9. <https://doi.org/10.1109/MC.2018.2701632>.
- Hagen JM, Albrechtsen E, Hovden J. Implementation and effectiveness of organizational information security measures. *Information Management and Computer Security* 2008;16:377–97. <https://doi.org/10.1108/09685220810908796>.
- Halevi T, Memon N, Nov O. Spear-Phishing in the Wild: A Real-World Study of Personality, Phishing Self-Efficacy and Vulnerability to Spear-Phishing Attacks. *SSRN Journal* 2015. <https://doi.org/10.2139/ssrn.2544742>.
- Harel Y. Ransomware incidents aren't personal attacks against an organization's management. *Israel Defense*; 2021. <https://www.israeldefense.co.il/en/node/51754>
- Hart S, Margheri A, Paci F, Sassone V. Riskio: A Serious Game for Cyber Security Awareness and Education. *Computers and Security* 2020;95. <https://doi.org/10.1016/j.cose.2020.101827>.
- Jain AK, Gupta BB. A survey of phishing attack techniques, defence mechanisms and open research challenges. *Enterprise Information Systems* 2022;16:527–65. <https://doi.org/10.1080/17517575.2021.1896786>.
- Jevšček M, Vrhovec S, Bernik I. Testing the Human Backdoor: Organizational Response to a Phishing Campaign Faculty of Organisation Studies in Novo Mesto. *Article in JOURNAL OF UNIVERSAL COMPUTER SCIENCE* 2019. <https://doi.org/10.3217/jucs-025-11-1458>.
- Johns, E. (2020). *Cyber security breaches survey 2020*. London: Department for Digital, Culture, Media & Sport.
- Kirlappos I, Parkin S, Sasse MA. "Shadow security" as a tool for the learning organization. *ACM SIGCAS Computers and Society* 2015;45:29–37. <https://doi.org/10.1145/2738210.2738216>.
- Kumaraguru P, Sheng S, Acquisti A, Cranor LF, Hong J. Lessons from a real world evaluation of anti-phishing training. 2008 *eCrime Researchers Summit, Atlanta, GA, USA: IEEE*; 2008, p. 1–12. <https://doi.org/10.1109/ECRIME.2008.4696970>.
- Lain D, Kostiaainen K, Capkun S. Phishing in Organizations: Findings from a Large-Scale and Long-Term Study. *IEEE Symposium on Security and Privacy (SP)* (Pp 842-859) 2021.
- Lin E., Greenberg S., Trotter E., Ma D., Ayccock J, Lin E, et al. Does Domain Highlighting Help People Identify Phishing Sites Does Domain Highlighting Help People Identify Phishing Sites? *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Pp. 2075-2084): 2010.
- Longtchi T, Rodriguez RM, Al-Shawaf L, Atyabi A, Xu S. Internet-based Social Engineering Attacks, Defenses and Psychology: A Survey 2022.
- Michelle K, Kowalski E, Cappeli D, Moore A, Shimeall T, Rogers S. Insider Threat Study: Computer System Sabotage in Critical Infrastructure Sectors. *National Threat Assessment Ctr Washington Dc* 2005.
- Mirsch T, Lehrer C, Jung R. Digital Nudging: Altering User Behavior in Digital Environments. *Proceedings Der 13 Internationalen Tagung Wirtschaftsinformatik (WI 2017)* 2017:634–48.

Pac R, Capstone A. PHISHING THREATS, ATTACK VECTORS, AND MITIGATION. Doctoral Dissertation, Utica College: 2017.

Petrykina Y, Schwartz-Chassidim H, Toch E. Nudging users towards online safety using gamified environments. *Computers and Security* 2021;108. <https://doi.org/10.1016/j.cose.2021.102270>.

Proofpoint, 2022. State of the Phish: An In-Depth Exploration of User Awareness, Vulnerability and Resilience. <https://www.proofpoint.com/au/resources/webinars/state-phish-2022>

Rastenis J, Ramanauskaite S, Janulevicius J, Cenys A. Credulity to Phishing Attacks: A Real-World Study of Personnel with Higher Education. 2019 Open Conference of Electrical, Electronic and Information Sciences (eStream), Vilnius, Lithuania: IEEE; 2019, p. 1–5. <https://doi.org/10.1109/eStream.2019.8732169>.

Reeves A, Delfabbro P, Calic D. Encouraging Employee Engagement With Cybersecurity: How to Tackle Cyber Fatigue. *SAGE Open* 2021;11:215824402110000. <https://doi.org/10.1177/21582440211000049>.

Safa NS, Maple C. Human errors in the information security realm – and how to fix them. *Computer Fraud and Security* 2016;2016:17–20. [https://doi.org/10.1016/S1361-3723\(16\)30073-2](https://doi.org/10.1016/S1361-3723(16)30073-2).

Sahni NS, Christian S, Chintagunta WP, Gentzkow M, Goldfarb A, Sudhir K, et al. Personalization in Email Marketing: The Role of Non-Informative Advertising Content *. *Marketing Science*, 37(2), 236-258: 2016.

Schuetz SW, Lowry PB, Thatcher JB. DEFENDING AGAINST SPEAR PHISHING: MOTIVATING USERS THROUGH FEAR APPEAL MANIPULATIONS. 20th Pacific Asia Conference on Information Systems (PACIS 2016), Chiayi, Taiwan 2016.

Siponen M, Pahlila S, Mahmood MA. Compliance with Information Security Policies : IEE Computer Society 2010;43:64–71.

Sommestad T, Karlzén H. A meta-analysis of field experiments on phishing susceptibility. IEEE: 2019.

Sutter T, Bozkir AS, Gehring B, Berlich P. Avoiding the Hook: Influential Factors of Phishing Awareness Training on Click-Rates and a Data-Driven Approach to Predict Email Difficulty Perception. *IEEE Access* 2022;10:100540–65. <https://doi.org/10.1109/ACCESS.2022.3207272>.

Thomas JE. Individual Cyber Security: Empowering Employees to Resist Spear Phishing to Prevent Identity Theft and Ransomware Attacks. *IJBM* 2018;13:1. <https://doi.org/10.5539/ijbm.v13n6p1>.

Torten R, Reaiche C, Boyle S. The impact of security awareness on information technology professionals' behavior. *Computers and Security* 2018;79:68–79. <https://doi.org/10.1016/j.cose.2018.08.007>.

Vega J, Shevchyk D, Cheng Y. A Literature Survey of Phishing and Its Countermeasures. 2022 Computer Science Conference for CSU Undergraduates: 2022.

Wang M, Song L. An Incentive Mechanism for Reporting Phishing E-Mails Based on the Tripartite Evolutionary Game Model. *Security and Communication Networks* 2021;2021. <https://doi.org/10.1155/2021/3394325>.

Wash R, Cooper MM. Who provides phishing training? Facts, stories, and people like me. *Conference on Human Factors in Computing Systems - Proceedings*, vol. 2018- April, Association for Computing Machinery; 2018. <https://doi.org/10.1145/3173574.3174066>.

Wen ZA, Lin Z, Chen R, Andersen E. What.Hack: Engaging Anti-Phishing Training through a Role-playing Phishing Simulation Game. *Conference on Human Factors in Computing Systems - Proceedings* 2019:1–12. <https://doi.org/10.1145/3290605.3300338>.

Williams EJ, Hinds J, Joinson AN. Exploring susceptibility to phishing in the workplace. *International Journal of Human Computer Studies* 2018;120:1–13. <https://doi.org/10.1016/j.ijhcs.2018.06.004>.

7. Appendix

Changes in parking arrangements



Reply
 Reply All
 Forward
 ...

Dear [redacted],

In light of our recent re-evaluations and the large number of inquiries we have received on the subject, we are making changes to the parking arrangements at the various facilities of the organization, including branches. To view the new policy and understand the significance of the change for you, click [here](#).

Regards,
Human Resources Department

Fig. 12 - Parking arrangements (translated)

Changing employees' badge photo



Reply
 Reply All
 Forward
 ...

Dear employee,

We are happy to update you that we have added an option to update the photo of the employee badge. The photo will be updated no later than **4.11.21**.

Few steps to update the image:

- Go to the attached link - [image](#)
- Enter the employee number and email address
- Attach the requested photo
- It's possible to get the tag with the updated photo at the reception desk at 7/11/21

Regards,
Human Resources Department

Fig. 13 - Changing employees' badge photo (translated)

Malicious email detection software



Reply
 Reply All
 Forward
 ...

Dear [redacted],

In honor of Information Security and Cyber Protection Awareness Week starting this month, And further the cyber-attack on the [redacted] hospital and cyber threats that renew every day It was decided that the employees **must** install by the date: **4/11/2021** a smart system for detecting malicious messages.

The purpose of the system is to identify email messages sent by malicious parties and to filter these messages so that they are not received as incoming mail at all in the recipient's email account.

Operating steps - malicious email detection system:

- Click on the attached link: [mails detection system](#)
- Type the employee number and email address
- On the desktop will appear a shortcut with the icon for the mails detection system

Thank you for your cooperation.
Regards,
System updates and upgrades department

Fig. 14 - Malicious email detection software (translated)

COVID green pass



[Reply](#)
[Reply All](#)
[Forward](#)
[...](#)

Dear [redacted],

According to our records, your COVID green card is expired. **Pay attention - failure to extend the COVID green card will result in restrictions in accordance with the government's updated instructions!**

In order to ensure automatic renewal of the COVID green card, you must go to the website of the Ministry of Health in the [link](#) and fill out the online form.

Alternatively, if a mistake has been made and your COVID green card is valid, click [here](#) and report it to us.

As well, we launched a new service that lets you watch queue status cameras for COVID tests according to regions and health insurance - to view by region [click here](#)

After filling in the details, the renewed COVID green card will be sent to you.

Making sure you stay healthy,

Ministry of Health

Fig. 15 - COVID-19 green pass (translated)

Ice cream coupon



[Reply](#)
[Reply All](#)
[Forward](#)
[...](#)

Dear clients,

We opened the new branch at [redacted], and as employees of the [redacted] you are entitled to a variety of great benefits at the new branch!!!

And this time a special January benefit in light of the opening - all the flavors on the menu are on sale: 3 kg for the price of 1!!

To redeem the benefit, go to the [link](#) and type the coupon code [redacted] in the payment line.

The benefit is valid for one week only on the dates January 17-23, 2022 - hurry to redeem!

We will be happy to send you the most delicious ice cream in the country - the delivery is on the house 🍦

See you at the next benefit

Customer service department

Fig. 16 - ice cream coupon (translated)

Financial lectures



[Reply](#)
[Reply All](#)
[Forward](#)
[...](#)

Dear employee,

As an employee of the bank, you are entitled to a free subscription to the FINANCE TALKS website.

On this website, you can enjoy a variety of interesting lectures in the financial field.

Will Bitcoin become an official currency in other South American countries?

How did COVID affect the global economy - two years into the epidemic

The shadow economy of the darknet

These are just a small part of the lectures you will be exposed to...

To redeem the benefit, [click here](#).

If you want us to remove you from the mailing list [click here](#).

Regards,

Finance team

Fig. 17 - Financial lectures (translated)

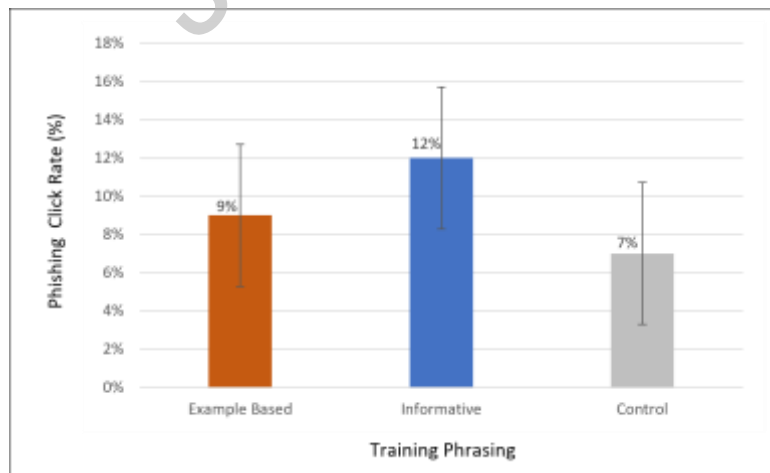


Fig. 18 - Phishing click rate per training phrasing in Wave 3 (January 22) - informative versus example-based

Doron Hillman

Doron Hillman is a Master's student at Tel Aviv University, in the Department of Industrial Engineering. In parallel, he works as a technical product manager in IBM's Trusteer Labs.

Eran Toch

Eran Toch is an associate professor at the Department of Industrial Engineering at Tel-Aviv University. He is also the co-director of the IWIT (Interacting with Technology Lab). Eran's research group works on usable privacy and security, large-scale analysis of interactive behavior, and mobile computing. The research group is currently running several projects funded by agencies such as the Israeli Science Foundation (ISF), DARPA, European Union Horizon 2020 Program, and the Israel Ministry of Science.

Yaniv Harel

Yaniv Harel acts as the Chief Strategy Officer of the Interdisciplinary Cyber Research Center at Tel Aviv University. Dr. Harel is involved in Cyber Research and serves as the Chairman of the Cyber Week Academic Conference. In his industrial capacity, Yaniv is the former GM of the Dell Technologies Cyber Solutions Group and was also the SVP Cyber Defense in Sygnia. Previously, he served years in governmental positions, leading technological divisions and units. Dr. Harel is one of the National Security Award winners of 2002. He consults Start-ups and large companies, and serves as a member in Governmental committees.

CRedit author statement

Doron Hillman: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data Curation, Writing - Original Draft, Writing - Review & Editing, Visualization, Project administration

Prof. Eran Toch: Conceptualization, Methodology, Validation, Resources, Writing - Original Draft, Writing - Review & Editing, Supervision, Project administration, Funding acquisition

Dr. Yaniv Harel: Conceptualization, Methodology, Validation, Resources, Writing - Original Draft, Writing - Review & Editing, Project administration, Funding acquisition

Declaration of interests

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐ The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Journal Pre-proof