

# *Spotlight on Phishing: A Longitudinal Study on Phishing Awareness Trainings*

Florian Quinkert, Martin Degeling, and Thorsten Holz

Ruhr-Universität Bochum

{florian.quinkert,martin.degeling,thorsten.holz}@rub.de

**Abstract.** Phishing is in practice one of the most common attack vectors threatening digital assets. An attacker sends a legitimate-looking e-mail to a victim to lure her on a website with the goal of tricking the victim into revealing credentials. A phishing e-mail can use both technical (e.g., a forged link) and psychological vectors (e.g., an authoritarian tone) to persuade the victim.

In this paper, we present an analysis of more than 420,000 phishing e-mails sent over more than 1.5 years by a consulting company offering awareness trainings. Our data set contains detailed information on how users interact with the e-mails, e.g., when they click on links and what psychological vectors are used in the e-mails to convince the recipient of its legitimacy. While previous studies often used lab environments, the e-mails in our data set are sent to real users during their day-to-day work so that we can study their behavior in a genuine setting. Our results indicate a continually decreasing click rate (from 19% to 10%) with progressing awareness training. We also found some psychological vectors, including an *authoritative tone* and *curiosity*, to be more effective than others to trick a user into falling for this type of scam e-mails.

**Keywords:** Phishing, Measurement Study, Awareness Training

## 1 Introduction

For businesses [18] and private persons [21], phishing is still one of the most commonly used attack vectors. In phishing, an attacker sends a legitimate-looking message (often via e-mail) to a victim in order to lure her on a website under the attacker’s control [11]. These e-mails have both technical (e.g., a forged link or a spoofed sender address) and psychological characteristics (e.g., an authoritarian tone or a luxurious offer) to convince victims of the message’s legitimacy [14, 38]. If a victim enters sensitive information (e.g., passwords or credit card information) on the attacker’s website, the adversary can use it for malicious purposes, such as identity theft or financial fraud. The Anti-Phishing Working Group (APWG) announced the detection of about 147.000 phishing websites in the second quarter of 2020 [2], an increase of about 25,000 compared to the fourth quarter of 2018 [1]. Due to the popularity of this attack technique, it is crucial to understand *why* users fall for phishing and study methods that help us to educate them accordingly.

Previous research in this area already studied users’ susceptibility to phishing [5, 11] and proposed different education methods [7, 26, 27]. Multiple studies analyzed how psychological and technical vectors are used in phishing attacks [6, 12, 13, 34]. These studies rely on experiments with a limited number of participants (typically students), often in a (potentially artificial) lab environment. As a result, we argue that existing work is limited in its generalizability, and it remains unclear which factors influence a victim’s likelihood of falling for this type of scam e-mails in a real-world setting.

In this paper, we present a comprehensive analysis of phishing e-mails sent by a consulting company. In this paper, we refer to it as *PhishCo* as pseudonym. Clients hire PhishCo to send phishing e-mails to their employees to raise security awareness within the company. A key aspect of PhishCo’s approach is to create phishing e-mails that are as close as possible to real-world phishing e-mails. For that purpose, they analyze real-world phishing e-mails and use the same techniques in their own e-mails. In particular, they use a combination of 14 different psychological and technical characteristics and customize the phishing e-mails to fit the hiring company’s business.

PhishCo provided us with a data set consisting of 429,418 e-mails sent over almost 1.5 years to employees of 77 different clients. For each e-mail, the data set contains detailed information about both the sent e-mail (e.g., used psychological and technical vectors) and the employees’ interaction with it (e.g., when the provided link was clicked). The data set is anonymized and does not contain any personally identifiable information (PII) related to the employees’ interaction. Unlike previous studies, the e-mails were sent to actual employees of companies belonging to a variety of industrial sectors during their day-to-day routine instead of a lab environment. Furthermore, the sent e-mails are close to real-world phishing e-mails so that we firmly believe employees likely behave similarly to when they receive a real phishing e-mail. In addition, the number of analyzed e-mails in our data set is substantially larger than previous studies, further improving the generalizability of our results.

In our analysis, we discovered an improving click rate over time. That is, employees clicked less often on the provided links in phishing e-mails in later phases, demonstrating that sending phishing e-mails to employees over a long time can be helpful to counter carelessness. Moreover, we found that the psychological vector *authority* has the most significant positive influence, i.e., employees were more likely to click on a link in an e-mail if this vector was used.

In summary, we make the following key contributions in this paper:

- We analyze a data set that is magnitudes larger than ones used in previous studies and consists of (fake) phishing e-mails that use both technical and psychological vectors to deceive users, enabling us to empirically study their influence on a large scale.
- In contrast to the majority of previous publications, we do not rely on an artificial lab environment, but gain insights into user’s behavior when receiving (fake) phishing e-mails during their day-to-day routine over about one year.

- Our results show a constantly improving click rate over time, indicating that long-lasting awareness training can help users to better identify phishing e-mails.

The remaining of this paper is structured as follows: in Section 2, we introduce background information and present related work. Afterwards, we explain PhishCo’s approach and the workflow in Section 3, and then present the measurement results in Section 4. Finally, we discuss lessons learned, limitations, and ethical considerations of our work in Section 5 and conclude along with recommendations in Section 6.

## 2 Background and Related Work

In this section, we describe related work which dealt with technical and psychological vectors. In addition, we present related phishing surveys.

### 2.1 Technical Vectors

An attacker creates a phishing e-mail that imitates a legitimate message ideally so that a victim might expect such an e-mail. For example, an attacker uses actual e-mails from the targeted company as templates, generates a similar layout, and includes company logos to let the e-mail look believable [29, 35]. Moreover, an attacker can try to mimic the writing style and tone of the targeted companies’ e-mails [40]. Many recipients judge incoming e-mails based on the sender’s e-mail address so that spoofing this address is a preferred method to persuade victims [19]. PhishCo uses comparable techniques (e.g., layout similarity, writing style, and spoofed e-mail addresses) to create believable e-mails and convince clients’ employees to open the sent e-mails.

Furthermore, an attacker needs plausible domains to lure victims on websites under the attackers’ control. Previous publications studied multiple technical attack techniques for that purpose, which are often referred to as *domain squatting* techniques. In *typosquatting*, an attacker creates a domain which differs from a well-known domain only by a typical typing error, e.g., `paypl.com` or `paypaal.com` [3]. A *combosquatting* domain consists of a well-known domain with added suitable terms so that the resulting domain still looks believable, e.g., `bankofamerica-security.com` or `secure-paypal.com` [23]. A domain cannot only contain Latin letters but also letters from other alphabets, such as Cyrillic. An attacker can replace one or multiple characters in well-known domain with similar looking characters from other alphabets, e.g., `bankofámerica.com`, which is referred to as *homograph domain* [33]. Registering a domain which sounds similar to a well-known domain, e.g., `guaranty-bank.com` instead of `guarantee-bank.com`, is called *soundsquatting* [30]. PhishCo registered a set of domains using such techniques to use them in their e-mails sent to clients’ employees as part of the awareness campaigns.

## 2.2 Psychological Vectors

Attackers do not only use technical vectors but also psychological ones, e.g., to create pressure or letting the victim feel important. Eventually, psychological vectors aim at convincing a victim to click on a link or open a provided attachment so that the attacker can, for example, collect personal information. Cialdini et al. introduce the six principles of persuasion and provide multiple examples of successful applications [9]. Similarly, Gragg analyzes triggers used to perform successful social engineering attacks and derives multiple defenses to counter these triggers [16]. Stajano et al. present principles based on real-life scams and conclude that security designers have to remember these principles in the development process [36]. Ferreira et al. combine the vectors of Cialdini et al., Gragg, and Stajano et al. to five principles of persuasion in social engineering attacks [12]. In a follow-up study, they extract the most effective elements of phishing e-mails and analyze them with regard to the previously introduced principles [13]. Van der Heijden et al. use the principles introduced by Cialdini et al. to build a classifier estimating how likely a human will fall for a certain phishing mail so that response teams in companies can prioritize incoming phishing e-mails [17]. Williams et al. conduct studies to analyze the effects of *urgency* and *authority* in e-mails [39]. PhishCo use a subset of these psychological vectors identified in prior publications for use in e-mails sent to clients' employees.

## 2.3 Phishing Surveys

Dhamija et al. present a user study in a lab environment with 22 participants, who classify 20 websites into legitimate and malicious [11]. Kumaraguru et al. introduce an embedded training method and a game to teach users how to identify phishing e-mails and malicious URLs [26]. In a follow-up study, Kumaraguru et al. use their training method and conduct a study with 515 faculty, staff, and students from Carnegie Mellon University, who received 10 e-mails within 28 days [25]. Caputo et al. send three phishing e-mails to 1,359 participants, show training material to a subset of them, and analyze whether it prevents participants from clicking on links in phishing e-mails [8]. Butavicius et al. ask 121 university students to classify 12 e-mails into legitimate, phishing, and spearphishing [6]. Rajivan et al. perform a two-phase experiment in which 105 participants first create phishing e-mails, which are classified in a second phase along with legitimate e-mails by 340 other participants [34]. Oliveira et al. conduct an experiment with 158 participants to understand whether younger persons or older persons have a different susceptibility for phishing attacks [31]. For that purpose, they sent spearphishing e-mails to the participants on 21 consecutive days in their day-to-day life. They concluded that older women were most prone to phishing attacks. Petelka et al. analyze the effect of different positions for suspicious URL warnings (close to suspicious URL, display on hovering the suspicious URL, deactivating the original URL and let user click it in the warning) [32]. 701 participants recruited via Mechanical Turk opened e-mails in a lab environment and answered questions about the e-mails. Wash et al. explore

the influence of training users with facts about phishing or stories of previous victims to reduce the users’ susceptibility for phishing [37]. They sent phishing e-mails to 2,000 faculty members out of which 26,8% clicked on a link in an e-mail. Afterwards, these participants got one form of the aforementioned phishing training. The authors inferred that facts about phishing is the more convincing form of phishing training.

While some of the previously mentioned publications already sent e-mails to actual users in their day-to-day business, they still often rely on artificial lab environments and a comparably small number of participants. In contrast, we base our survey on a much larger data set, not collected in artificial lab environments, but during day-to-day business. Hence, we believe that the observed user reaction for our data set is likely very similar to actual phishing attacks. Furthermore, PhishCo’s fake phishing campaigns last for one year which is a lot longer than previous studies, enabling long-term observations. In addition, our data set covers more employees from a wider variety of industry sectors and does not focus on employees/students from one company/university. We are convinced that this aspect further improves the generalizability of our results.

### 3 PhishCo’s Approach

In the following, we describe PhishCo’s approach and how our data set is generated. We first characterize how PhishCo creates e-mail templates, followed by a description of the process when a company hires PhishCo.

#### 3.1 E-Mail Generation

PhishCo offers awareness training and educational material in combination with sending (fake) phishing e-mails for one year. To avoid copyright infringements, the e-mails do not contain content from well-known companies, i.e., PhishCo does not send, for example, a fake Paypal or Amazon e-mail. Instead, PhishCo analyzes real-world phishing e-mails and creates e-mail templates which replicate typical content of phishing e-mails. PhishCo creates templates for different industrial sectors, such as *finance* or *education*, and customizes the templates for each client with respect to e-mail signatures and senders. Furthermore, each template utilizes a combination of psychological and technical vectors to convince victims of its legitimacy. PhishCo uses 14 different psychological and technical vectors in the templates, following publications on phishing and behavioral psychology discussed above. The psychological vectors include phrases drawing the victim’s interest (*curiosity*) or flattering the recipient (*praise/flattering*). The technical vectors contain, for example, the use of a spoofed e-mail address (*mail address spoofing*) or domains similar to a well-known one (*domain squatting*). Table 1 provides a full list of the used vectors, which we will analyze in detail in Section 4.4. The majority of templates use two or three of these vectors, which we will analyze in more detail in the next section.

```

+++ This message has been generated automatically +++

Dear REDACTED,

During our regular breach and data leak database scan we have identified your account
REDACTED@REDACTED as being compromised. Specifically, this could mean that your password is or has
been visible to third parties. A fraudulent usage of your data or even liability claims could be the
consequence of this data breach.

We strongly advise you to change your password via the password management section of your account
preferences:
Change password

Please make sure that the new password comprises at least one of the following features:


- At least 8 characters
- Letters and numbers
- Special characters



Kind Regards,
REDACTED Defense
-----
Please note: REDACTED Defense Ltd. is an official security provider of REDACTED

```

**Fig. 1.** Sample e-mail which has been sent to an employee of a client using the vectors *pressure/anxiety*, *trust/intimacy*, and *input mask*

Figure 1 shows an example of an e-mail that was sent to an employee as part of a campaign. The e-mail claims that the employee’s account was compromised and requests a change of the corresponding password. It provides a link along with information on how to pick a new password and uses three of the aforementioned psychological and technical vectors. First, *pressure/anxiety* by pretending the employee’s account was compromised and suggesting a possible data breach as a consequence. Second, the e-mail utilizes *trust/intimacy* by addressing the employee with his/her name (redacted in the example e-mail), appearing to be helpful, and pretending to be from an official security provider of the client. Third, the e-mail uses the technical vector *input mask* when the employee clicks on the *change password* link. That is, it will show input fields, asking the employee for his current password and a new one. PhishCo prevents data leakage by disabling input fields if a user actually tries to enter a password or, in other cases, similar private information. The URL uses the scheme *client-company.example.com*, i.e., a vigilant employee could detect that it is not the client’s website.

Figure 2 shows a second example e-mail, which appears to be a job application. The job description is rather generic, so that most companies will likely search for applicants in this area. The e-mail uses the vector *curiosity* because the receiving employee is tempted to take a look at the documents in the attachment or dropbox (especially if the employee is working in the HR department). The vectors *attachment* and *link execution* refer to the attached documents and the link in the e-mail. The URL uses the scheme *dropbox.example.com* so that a careful employee could identify it as a link that does not lead to the actual dropbox website.

Analyzing real-world phishing e-mails and using them as basis for the sent e-mails along with using psychological and technical vectors ensures that the e-mails sent to clients’ employees are close to real-world phishing e-mails. We

Dear Mr. Doe

With great interest I have seen the [job advertisement for the office assistant on your website](#).

After eight years of work experience with a major provider of fire protection and security services, I now would like to develop professionally. I'd like to offer you numerous advantages due to my experience as an office clerk with a lot of organizational talent and a lot of experience in customer service. Please refer to my detailed application documents attached to this mail.

All documents, scans and work samples can also be downloaded via dropbox:  
<https://www.dropbox.com/sh/740kljwt8usm/AAAJrtkZheapHRH8-mnsOsEa?dl=0>

I'm looking forward to hearing back from you.

Kind regards  
 Chloe Parker

**Fig. 2.** Second sample e-mail which has been sent to an employee of a client using the vectors *curiosity*, *attachment*, and *link execution*

argue that analyzing the interaction of employees with these e-mails allows us to make conclusions about how users understand real-world phishing e-mails.

### 3.2 Workflow with Client

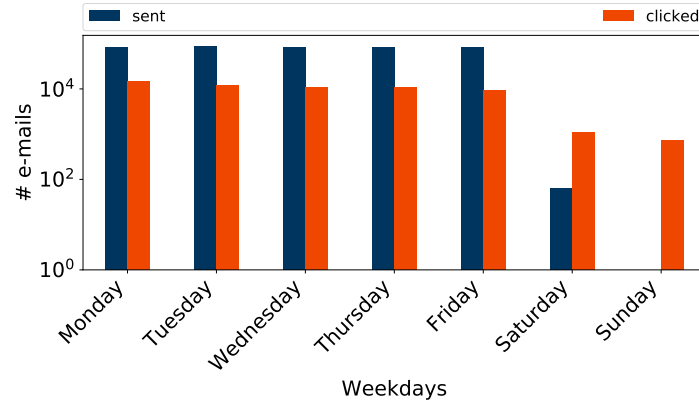
Clients hire PhishCo to train their employees on how phishing works, enabling them to identify malicious phishing e-mails in their day-to-day routine. Each training campaign consists of two phases (called P1 and P2 in the following) and usually runs for at least one year:

- In phase P1, which only lasts a couple of weeks, each employee receives less than five phishing e-mails.
- In phase P2, PhishCo sends less than 15 phishing e-mails to each employee distributed over the remainder of the year-long campaign.

The shorter phase P1 informs employees about the general problem of phishing, raises awareness, and allows teaching them not to fall for phishing. The longer phase P2 enables both reviewing and deepening employees' phishing understanding and knowledge. Before a campaign starts, the client notifies its employees that a phishing training will be conducted. The whole process is performed in accordance with data privacy laws, and both the data protection officer and the workers' council are made aware of the training. To ensure that the phishing e-mails reach the targeted employees, the client whitelists incoming e-mails from PhishCo.

Each e-mail sent to an employee contains a link to draw the employee's interest. If an employee clicks a link, he/she is notified that the e-mail was sent by PhishCo. Furthermore, PhishCo notes when an e-mail was sent and whether and when the link in the e-mail was clicked. Additionally, the e-mails contain a tracking pixel so that PhishCo can also log when an e-mail is opened by a client's employee.

For data protection, the collected data does not contain any information about the employee so that no personally identifiable information is stored.



**Fig. 3.** Number of sent and clicked e-mails as a function of the weekday. The figure uses a logarithmic scaling on the y-axis to show that at least a small number of e-mails are clicked on Saturdays and Sundays.

Hence, neither we nor the client can link multiple e-mails to the same employee and measure the employee’s performance. That is, a client cannot use collected data for disciplinary punishments or releases of employees.

## 4 Results

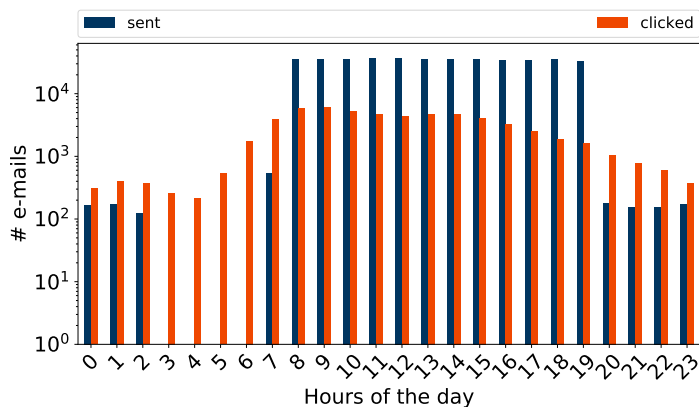
Next, we present the results of our data analysis. We start with a short description of the data set, followed by an analysis of the send and click times, the distribution among clients, the psychological and technical vectors, the click rate, and the effect of the psychological vectors.

### 4.1 Data Set

The data set we use for this study covers a period between November, 19th 2018, and April, 10th 2020. It consists of 429,418 e-mails sent in campaigns for 77 distinct clients. All 77 campaigns finished phase one (P1), and 14 campaigns already completed the full phase two (P2). On average, P1 lasts 20.22 days (standard deviation 9.67, minimum of 2, and a maximum of 73 days). P2 of the 14 finished campaigns lasted, on average, 285.29 days (standard deviation 115.53, minimum of 54 days, and a maximum of 372 days). The definite number of days in P1 and P2 depends on the agreement between PhishCo and the client, which explains the differences in the time periods.

### 4.2 Send and Click Times

Figure 3 depicts the number of sent and clicked e-mails per weekday. Note that the y-axis uses a logarithmic scaling. About 86.000 phishing e-mails are sent



**Fig. 4.** Number of sent and clicked e-mails as a function of the hours of the day.

on average on a weekday. Thursday is the weekday with the lowest number of sent e-mails (85,494 e-mails), while Tuesday is the weekday with the highest number of sent e-mails (86,418). A chi-square test against a uniform distribution shows that e-mails are evenly distributed among the weekdays ( $p$ -value  $< 0.01$ ). Usually, PhishCo does not send e-mails on the weekend, besides a few ones on Saturdays for testing purposes. At the beginning of the week, employees click PhishCo’s e-mails more often, with numbers decreasing from Monday to Friday. Overall, the number of clicked e-mails remains in the same magnitude. Most employees work from Monday to Friday and regularly check their e-mails so that the similar distribution among workdays is not surprising. On the weekend, only a few employees click e-mails because most employees do not work during the weekend. Furthermore, PhishCo only sends a minimal number of e-mails on the weekend.

Figure 4 shows the number of sent and clicked e-mails as a function of the hours of the day. Similar to the previous analysis, we added a table to Appendix A, which contains more detailed information on the sending and clicking times of the day. E-mails are sent during working hours from 8 AM to 8 PM. Employees interact with the e-mails especially in the morning and after lunch, with clicks decreasing during the night because most employees are likely asleep.

The observed days and times are similar to the interaction with marketing e-mails reported by multiple analyses [4, 20, 28] and show a typical diurnal pattern. Our data set suggests that sending e-mails early in the morning and after lunch as well as earlier during the week could further improve the effectiveness of the phishing e-mails because employees handle their e-mails especially at these times. In addition, clicking e-mails at the weekend or late in the evening can have multiple reasons. For example, employees can be in a different time zone than PhishCo, which sent the e-mails, or they are working long hours. Moreover, it can indicate that e-mails are opened outside the company’s infrastructure. Using pri-

vate, not properly secured devices can pose a security threat. Late in the evening and at night, we find only a few interactions because most employees are likely asleep. Interaction with e-mails starts already at 6 AM when the first employees start working. Therefore, our data suggest that starting to send e-mails already at 6 AM could reach many employees while they start their workday. Furthermore, customizing the sending times closer to times, reflecting the interaction times, could further improve the effectiveness of the phishing e-mails.

### 4.3 Distribution Among Clients

Our data set contains campaigns for 77 clients. On average, 5,576 e-mails are sent to employees of each client (standard deviation 8752.45, minimum 103, and maximum 42,268). The high standard deviation, along with the minimum and maximum of e-mails per client, indicates a large difference in e-mails per client. The number of e-mails per client highly depends on the number of employees and the point of time in the campaign, e.g., our data set contains only very few e-mails for a small client in an early stage of a campaign. In contrast, a client with many employees will account for a lot more e-mails. The top three clients are responsible for about 25% of all sent e-mails, and the top eight clients for about 50% of all sent e-mails. The results indicate that our data set contains both big and small companies, which improves the generalizability of our analysis.

### 4.4 Psychological and Technical Vectors

PhishCo uses the psychological and technical vectors in e-mails with different frequencies. In some cases, clients request the usage of specific vectors, e.g., because they have been targeted with a similar vector before. Additionally, PhishCo gained experience over time, which vectors result in higher click rates. Table 1 depicts the number of sent and clicked e-mails per psychological and technical vector. Each vector was used in at least 10,000 e-mails. The most used psychological vector is *trust/intimacy*, followed by *curiosity*, and *pressure*. In case of technical vectors, *domain squatting*, *sender spoofing*, and *attachment* are most common. We will analyze how successful and promising the different vectors are along with the best combinations of vectors in Section 4.7. Usually, PhishCo uses not only one vector in an e-mail but also combines multiple vectors to get a convincing e-mail. Table 2 shows the distribution of co-occurrences of vectors. The most common one is *pressure* and *trust/intimacy*.

### 4.5 E-Mail Timeline

We now analyze how fast employees click on a link after they opened an e-mail. We argue that a prolonged time between opening and clicking from P1 to P2 indicates that employees think longer about whether it is a legitimate e-mail or not. Even though the employee eventually clicked on the link, it can indicate a better understanding of how phishing works and lead to a correct decision for future e-mails.

**Table 1.** Overview of psychological (P) and technical (T) vectors used in e-mails sent to participants.

Vector	Description	Sent	Clicked
Pressure (P)	Urges victim to act, e.g., by giving short time to reply.	178,946	14.31%
Curiosity (P)	Appealing to the recipient’s curiosity, e.g., using a catchy subject.	203,790	13.76%
Financial appeal (P)	Pretends a fiscal advantage for the victim, e.g., by offering a discount.	35,536	7.34%
Trust/intimacy (P)	Pretends to be from a known person, e.g., by using the victim’s name.	280,773	9.21%
Praise/flattering (P)	Flattens the recipient, e.g., by addressing her as valuable resource.	148,645	9.21%
Helpfulness (P)	Asks recipient to help, e.g., by taking part in a survey.	64,288	15.15%
Authority (P)	References hierarchies e.g., pretending to be from a superior.	38,616	14.13%
Attachment (T)	Contains an attachment.	156,939	16.31%
Input mask (T)	Website behind link in e-mail contains an input field.	104,399	17.24%
Link (T)	Tries to motivate a victim to open a link.	96,538	19.40%
Bulk mailing (T)	Addressed to a larger audience, e.g., all tax consultants.	12,822	9.43%
Reply / forward (T)	Forwards another e-mail, e.g., offering discounts.	10,750	6.84%
Sender spoofing (T)	Pretends to be from different sender than it is, e.g., a co-worker.	203,353	11.90%
Domain squatting (T)	Contains domain similar to well-known one.	79,458	8.15%

As explained earlier, the opening of an e-mail is only recorded if a tracking pixel is triggered. Therefore, it is possible that employees opened e-mails but their systems blocked the tracking pixel. In such cases, we can still understand whether an e-mail was opened when a link in the e-mail was clicked. However, in the following, we focus on e-mails for which we have the opened and the clicked times to have a consistent data set. We identified 33,265 e-mails which were opened and clicked. Calculating the time between opening and clicking revealed that 753 e-mails were clicked after more than one week and 166 even after more than a month. This is noteworthy because it shows that employees sometimes click on links in phishing e-mails even after a long time has passed. As a countermeasure, companies should blacklist URLs of known phishing e-mails to prevent harm from later clicked e-mails and already handled phishing cases.

17,161 e-mails belong to P1 and 13,674 e-mails to P2. In the following, we focus on the first five minutes after opening an e-mail because we consider it to be most likely that employees did not interrupt handling the particular e-mail when the link is clicked in this time frame. In P1, 12,243 e-mails (71.34%) and in

**Table 2.** Co-occurrences of vectors in the dataset

	Pressure	Curiosity	Financial	Trust	Flattering	Help	Authority
Pressure		65,581	14,399	169,990	189	50,883	31,357
Curiosity			10,969	131,613	4174	15,353	18,890
Financial				8379	504	773	419
Trust					3979	43,283	16,960
Flattering						0	0
Help							13,043

**Table 3.** Summary of sent and clicked e-mails along with click rate in relation to total, phase 1 (P1), and phase 2 (P2) numbers.

	Total	Phase 1 (P1)	Phase 2 (P2)
Sent	429,418	168,859	260,559
Clicked	59,689	32,188	27,501
Click rate	13.90%	19.06%	10.55%

P2, 9,810 e-mails (71.74%) were clicked within the first five minutes. On average, it took 1.26 minutes in P1 and 1.34 minutes in P2 from opening to clicking. That is, even though the employees eventually took a wrong decision, they spent, on average, more time on assessing the e-mails. A t-test of the average processing times in P1 and P2 led to a test statistic of -6.42 and a p-value below 0.01. Hence, the difference is significant.

In summary, our results show that links in phishing e-mails are opened even after a long time. Furthermore, constantly sending phishing e-mails leads to more time spent on a single e-mail.

#### 4.6 Click Rate

The click rate is the percentage of e-mails in which the receiving employee clicked on the provided link. In particular, we focus on the differences between phase one (P1) and phase two (P2). We are interested in how the click rate changes between P1 and P2 because a lower click rate in P2 indicates that the employees of a client gained a better understanding of how phishing e-mails look like. Note that we cannot make a statement about the performance of single employees, as this data is not collected. Table 3 summarizes the number of sent and clicked e-mails along with the click rate for both total numbers and split between P1 and P2. In summary, it shows a decreasing click rate. In this section, we use all sent e-mails, regardless of whether the tracking pixel worked and indicated an opening of a particular e-mail because we focus on clicked e-mails to calculate the click rate. Therefore, the numbers for clicked e-mails in P1 and P2 differ from the numbers in the previous section, in which we used only e-mails which have been opened, indicated by the tracking pixel.

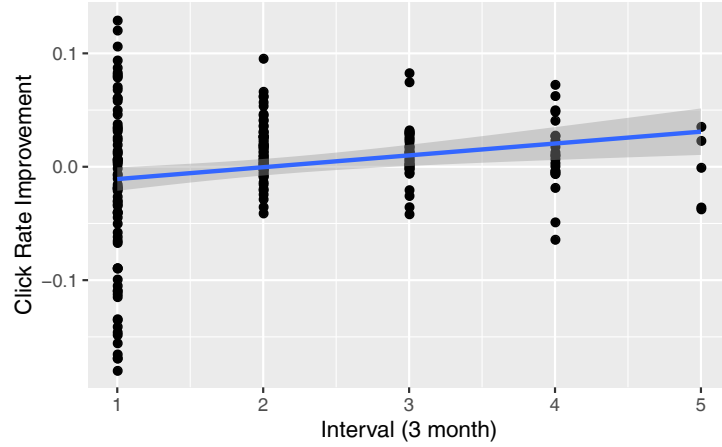
In total, PhishCo sent 429,418 e-mails to clients' employees, out of which 59,689 e-mails were clicked. Hence, the click rate is 13.90%. Due to the absence of large-scale academic studies analyzing such numbers for real-world phishing, it is difficult to compare the click rate with other publications. The security company KnowBe4 reported a click rate of 27 % for an initial phishing test, even though it remains unclear how the click rate is calculated in that particular case [24]. The Canadian government presented an infographic saying that in 10 % of phishing e-mails a link is clicked [15]. When we take into consideration that PhishCo sends highly targeted phishing e-mails where the sender is whitelisted, an overall click rate close to real-world phishing e-mails is what we expected.

Out of the 429,418 e-mails in the data set, 168,859 were sent as part of P1 and 260,559 in P2. The number of sent e-mails in P2 is higher because it lasts up to 49 weeks compared to a couple of weeks in P1. In P1, links were clicked in 32,188 (19.06 %) e-mails. The click rate in P1 is (as would be expected) higher as it is an initial test of the employees' phishing awareness.

In P2, the links in 27,501 e-mails were clicked, which leads to a click rate of 10.55 %. Hence, the click rate improves by about eight percentage points or 42 % for e-mails sent in P2 compared with P1. A possible explanation for this drop is a familiarization with the concept of phishing and this type of scam e-mails.

In addition to the improvement from P1 to P2, we expected an ongoing improvement during P2, based on the assumption that additional e-mails increase employees' understanding of phishing. Figure 5 shows the click rate improvement as a function of three months long intervals in P2. For each e-mail, we calculated the interval in which the e-mail was sent to an employee based on the start date of P2 for the corresponding client. Afterwards, we calculated the click rate as described previously and the click rate improvement compared to P1 (for the first interval) or the previous interval (for intervals two to n). Since not all clients have already finished P2, the number of clients decreases from interval to interval. Over time the click rate not only decreases, but this positive trend intensifies over time. This effect emphasizes the importance of long term training, showing that raising awareness over a long-time is helpful. There is a positive ( $r=0.21$ ) significant correlation ( $p=0,002$ ) between the improvement and the interval.

Besides analyzing the overall click rate, it is interesting to see how the click rate differs per client, which we further analyze in the following. Our data set contains 77 unique clients, which have an average click rate of 12.49 % with a standard deviation of 4.81. In P1, the 77 clients reach an average click rate of 15.65 % (standard deviation 7.18) and in P2 an average click rate of 10.37 % (standard deviation 4.97). The mean values for P1 and P2 again indicate an improvement. Comparing the average click rates of both phases, we found 61 clients who improved their click rate from P1 to P2 (minimum improvement -0.1 %, and maximum improvement -26.22 %). In contrast, 16 clients did not improve or decrease their click rate (minimum deterioration 0.59% and maximum deterioration 7.12 %). Figure 6 shows a scatter plot with each dot representing a client. The clients' position is defined by their click rate in P1 (x-axis) and P2



**Fig. 5.** Improvement compared between intervals

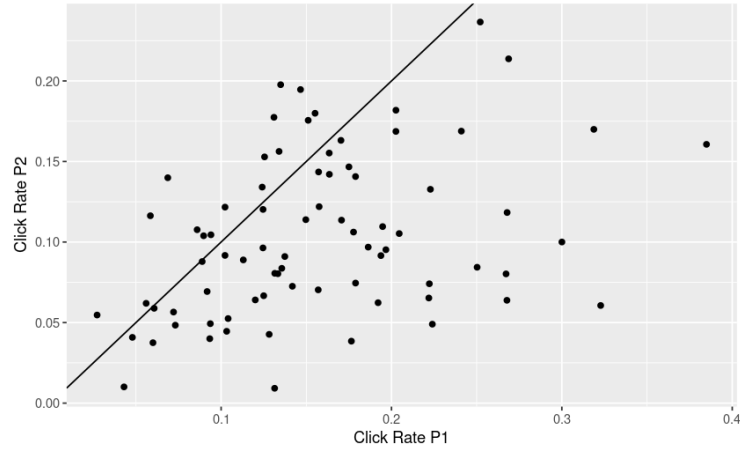
(y-axis). Clients below the line improved their click rate from P1 to P2. Besides the already discussed result that most clients improved their click rate, it shows that clients who did not improve their click rate are close to the line, i.e., they did not decrease their click rate much.

Our results reveal a decreasing click rate from P1 to P2. In addition, the click rate further decreases within P2, which emphasizes the importance of long-term awareness training. Finally, the sending of fake phishing e-mails proved to be useful for the majority of clients, as 61 out of 77 improved their click rate.

#### 4.7 Effect of Psychological Vectors

Figure 7 shows the click rate for templates with specific vector combinations. A chi-square test shows that the clickrate and the vector combinations are statistically independent. The template with the highest click rates was comprised of text, including the psychological vectors pressure, curiosity, trust, and authority (34 % click rate) as well as curiosity, trust, and authority. These e-mails all pretend to contain company internal information like updated emergency plans or training for IT security incidents as well as *CEO Fraud* [10] style e-mails. Those e-mails with the combination of curiosity, financial, trust, flattering (3 %), or curiosity and help (6 %) had low click rates. They mostly came from external contacts and either contained flattering invitations, e.g., for TV interviews or pretended to be customer e-mails.

Our results show a trend that confirms previous work that also found e-mails that claim to come from an “authority” to increase the click rate [39]. At the same time, our data shows that financial incentives – offering money or lucrative deals – often lead to lower click rates.



**Fig. 6.** Click rate in phase one (P1) (x-axis) and phase two (P2) (y-axis). Clients below the line improved their click rate.

## 5 Discussion

Compared to other publications, our data set is magnitudes larger than ones used in previous studies, covers longer time periods, and characterizes how employees react to phishing in their day-to-day business. In this section, we discuss our main results and compare them with previous publications.

We can confirm the results of Butavicius et al. that an authoritative tone increases the susceptibility of users to fall for phishing e-mails [6]. However, Butavicius et al. tested only authority, scarcity (similar to the vector “trust” in our data set), and social proof, which we do not have in our data set. The fact that “trust” also has a positive impact on the click rate can explain why *CEO Fraud* [10] has become so “successful” in practice. Coaxing users into giving away their information or offering financial gain, which is common in certain types of scams, is less successful in comparison. In contrast to e-mails using authority and trust, those which appear to offer financial benefits are less common in typical office situations and might, therefore, be easier to identify for many users. Furthermore, users often associate scam with a financial gain, e.g., the so-called Nigerian scam [22], so that they are more cautious. Even though some vectors led to higher click rates, we consider the variety of used vectors beneficial. Users are exposed to a wide range of phishing e-mails they might face in real world, which improves the chances that they do not click on an actual phishing e-mail when they receive it.

Additionally, our results show an improved click rate from P1 to P2 and further within P2. We did not observe an increasing click rate in later phases of P2, which could be an indicator of declining awareness. Hence, PhishCo’s approach has a long-lasting effect on an employee’s ability to identify phishing



Fig. 7. Clickrates for various vector combinations

e-mails. Therefore, raising awareness by regularly sending phishing e-mails helps to educate employees. In contrast to our work, Oliveira et al. could not determine a connection between the day in their 21 days long study and a click on a link [31]. The time period of only 21 days might be too short to already see an increased understanding of how phishing works.

## 5.1 Limitations

Our data set has several constraints that limit our analysis. First, we cannot evaluate the performance on a per-employee basis because PhishCo does not provide information to connect e-mails sent to the same employee. While this could potentially lead to interesting insights, data privacy concerns outweigh the benefits. In our study, we can still examine the changing performance on a per-client basis and deduce the overall improving performance. Second, our data set contains only phishing e-mails so that we cannot infer whether employees evaluate legitimate e-mails differently, such as spending more time to decide whether it is legitimate. Third, our data set contains only e-mails sent by PhishCo, i.e., no real-world phishing e-mails. Therefore, we cannot assess the influence of PhishCo’s e-mails on employees’ ability to identify actual phishing e-mails. However, PhishCo replicates real-world phishing e-mails they identify in the wild so that the e-mails in our data set are as close to real phishing e-mails as possible without using actual phishing e-mails.

## 5.2 Ethical Considerations

Our research institution does not have an IRB for computer science so that we could not get an IRB review to perform this study. However, as noted above, the data set we received from PhishCo does not contain any personal information about single employees or clients, but only technical information about the sent e-mails and meta data about when an e-mail was opened or a link in an e-mail clicked. Additionally, PhishCo prevented the accidental collection of private information by blocking employees from entering information in input fields on landing pages. Furthermore, due to the absence of personal information, it is not possible to identify an employee receiving an e-mail or identify multiple e-mails received by the same employee. That is, neither a client nor we can assess the performance of single employees based on the collected data. While we do not have the employees' consent for our analysis, the clients in advance notify their employees that a phishing awareness training will take place. Furthermore, PhishCo informs the employee that it was a phishing e-mail sent by them immediately after the employee clicked on a link. In summary, ethical concerns were considered during our study as the data set provided by PhishCo because is fully pseudonymized and it is not possible to make statements about employees or clients. Employees who received fake phishing e-mails had no disadvantage but got an opportunity to better understand how phishing works.

## 6 Conclusion and Recommendations

In this paper, we presented a detailed analysis of more than 420,000 phishing e-mails sent during more than 1.5 years as part of phishing awareness trainings performed in 77 companies. Compared to other publications, our data set is magnitudes larger than those used in other studies. Furthermore, the sent e-mails use 14 different technical and psychological vectors to create believable e-mails as close as possible to actual phishing e-mails. In contrast to multiple other studies, the e-mails were sent during employees day-to-day business instead of a lab situation. This leads to several unique insights into how people interact with phishing e-mails.

Employees continuously improve their click rate according to our results so that we recommend long-lasting awareness training instead of short-term ones. Similar to other publications, we identified the concept of *authority* as being the most successful and hence empirically confirmed this insight. Sending a variety of phishing e-mails that use different psychological and technical vectors proved to be useful so that employees are aware of different possible phishing schemes. Hence, we suggest using a diverse set of phishing e-mails, combined with current phishing trends, and phishing e-mails observed at the own institution to conduct successful phishing awareness trainings. An analysis of the interaction times revealed that employees especially interact with the sent e-mails at the beginning of the week, early in the morning, and after lunch. Therefore, we recommend focusing on these days and times to perform awareness trainings.

## References

1. Aaron, G.: Phishing Activity Trends Report - 4th Quarter 2019. [https://docs.apwg.org/reports/apwg\\_trends\\_report\\_q4\\_2019.pdf](https://docs.apwg.org/reports/apwg_trends_report_q4_2019.pdf) (2019)
2. Aaron, G.: Phishing Activity Trends Report - 2nd Quarter 2020. [https://docs.apwg.org/reports/apwg\\_trends\\_report\\_q2\\_2020.pdf](https://docs.apwg.org/reports/apwg_trends_report_q2_2020.pdf) (2020)
3. Agten, P., Joosen, W., Piessens, F., Nikiforakis, N.: Seven months' worth of mistakes: A longitudinal study of typosquatting abuse. In: Network and Distributed System Security Symposium (NDSS) (2015)
4. Bauer, E.: The 2017 Email Marketing Field Guide: The Best Times and Days to Send Your Message and Get It Read. <https://www.propellercrm.com/blog/2017-email-marketing-field-guide>
5. Blythe, M., Petrie, H.L., Clark, J.A.: F for fake: four studies on how we fall for phish. In: Conference on Human Factors in Computing Systems (CHI) (2011)
6. Butavicius, M., Parsons, K., Pattinson, M., McCormac, A.: Breaching the Human Firewall: Social Engineering in Phishing and Spear-Phishing Emails. In: Australian Conference of Information System (2015)
7. Canova, G., Volkamer, M., Bergmann, C., Reinheimer, B.: NoPhish App Evaluation: Lab and Retention Study. In: Workshop on Usable Security and Privacy (USEC) (2015)
8. Caputo, D.D., Pfleeger, S.L., Freeman, J.D., Johnson, M.E.: Going spear phishing: Exploring embedded training and awareness. *IEEE Security & Privacy* (2013)
9. Cialdini, R.B., Goldstein, N.J.: The Science and Practice of Persuasion. Cornell Hotel and Restaurant Administration Quarterly (2002)
10. Cidon, A., Gavish, L., Bleier, I., Korshun, N., Schweighauser, M., Tsitkin, A.: High Precision Detection of Business Email Compromise. In: Usenix Security Symposium (2019)
11. Dhamija, R., Tygar, J.D., Hearst, M.: Why Phishing Works. In: Conference on Human Factors in Computing Systems (CHI) (2006)
12. Ferreira, A., Coventry, L., Lenzini, G.: Principles of Persuasion in Social Engineering and Their Use in Phishing. In: International Conference on Human Aspects of Information Security, Privacy, and Trust (2015)
13. Ferreira, A., Lenzini, G.: An Analysis of Social Engineering Principles in Effective Phishing. In: Workshop on Socio-Technical Aspects in Security and Trust (STAST) (2015)
14. Fette, I., Sadeh, N., Tomasic, A.: Learning to Detect Phishing Emails. In: World Wide Web Conference (WWW) (2007)
15. Government, C.: Phishing: How many take the bait? <https://www.getcybersafe.gc.ca/cnt/rsracs/nfgrphcs/nfgrphcs-2012-10-11-en.aspx>
16. Gragg, D.: A Multi-Level Defense Against Social Engineering. SANS Institute - Information Security Reading Room (2003)
17. van der Heijden, A., Allodi, L.: Cognitive Triaging of Phishing Attacks. In: Usenix Security Symposium (2019)
18. Ho, G., Cidon, A., Gavish, L., Schweighauser, M., Paxson, V., Savage, S., Voelker, G.M., Wagner, D.: Detecting and Characterizing Lateral Phishing at Scale. In: 28th USENIX Security Symposium (USENIX Security 19) (2019)
19. Ho, G., Sharma, A., Javed, M., Paxson, V., Wagner, D.: Detecting Credential Spearphishing Attacks in Enterprise Settings. In: Usenix Security Symposium (2017)

20. Hodgekiss, R.: What Our Data Told Us about the Best Time to Send Email Campaigns. <https://www.campaignmonitor.com/blog/email-marketing/2019/01/best-time-to-send-email-campaigns-by-device/>
21. Hong, J.: The State of Phishing Attacks. *Commun. ACM* **55**(1) (2012)
22. Isacenkova, J., Thonnard, O., Costin, A., Francillon, A., Balzarotti, D.: Inside the scam jungle: a closer look at 419 scam email operations. In: *EURASIP Journal on Information Security* (2014)
23. Kintis, P., Miramirkhani, N., Lever, C., Chen, Y., Romero-Gómez, R., Pitropakis, N., Nikiforakis, N., Antonakakis, M.: Hiding in Plain Sight: A Longitudinal Study of Combosquatting Abuse. In: *Conference on Computer and Communications Security (CCS)* (2017)
24. KnowBe4: Report: 2018 Phishing By Industry Benchmarking Report. <https://www.ciosummits.com/KnowBe4-Phishing-By-Industry-Benchmarking-Report.pdf> (2018)
25. Kumaraguru, P., Cranshaw, J., Acquisti, A., Cranor, L., Hong, J., Ann Blair, M., Pham, T.: School of Phish: A Real-World Evaluation of Anti-Phishing Training. In: *Symposium on Usable Privacy and Security (SOUPS)* (2009)
26. Kumaraguru, P., Sheng, S., Acquisti, A., Cranor, L.F., Hong, J.: Teaching Johnny Not to Fall for Phish. In: *ACM Transactions on Internet Technology (TOIT)* (2010)
27. Lin, E., Greenberg, S., Trotter, E., Ma, D., Aycock, J.: Does Domain Highlighting Help People Identify Phishing Sites? In: *Conference on Human Factors in Computing Systems (CHI)* (2011)
28. Mailchimp: Insights from Mailchimp's Send Time Optimization System. <https://mailchimp.com/resources/insights-from-mailchimps-send-time-optimization-system/>
29. Mao, J., Li, P., Li, K., Wei, T., Liang, Z.: BaitAlarm: Detecting Phishing Sites Using Similarity in Fundamental Visual Features. In: *International Conference on Intelligent Networking and Collaborative Systems (INCoS)* (2013)
30. Nikiforakis, N., Balduzzi, M., Desmet, L., Piessens, F., Joosen, W.: Soundsquatting: Uncovering the use of homophones in domain squatting. In: *International Conference on Information Security (ISC)* (2014)
31. Oliveira, D., Rocha, H., Yang, H., Ellis, D., Dommaraju, S., Muradoglu, M., Weir, D., Soliman, A., Lin, T., Ebner, N.: Dissecting Spear Phishing Emails for Older vs Young Adults: On the Interplay of Weapons of Influence and Life Domains in Predicting Susceptibility to Phishing. In: *Conference on Human Factors in Computing Systems (CHI)* (2017)
32. Petelka, J., Zou, Y., Schaub, F.: Put Your Warning Where Your Link Is: Improving and Evaluating Email Phishing Warnings. In: *Conference on Human Factors in Computing Systems (CHI)* (2019)
33. Quinkert, F., Lauinger, T., Robertson, W., Kirida, E., Holz, T.: It's Not What It Looks Like: Measuring Attacks and Defensive Registrations of Homograph Domains. In: *Conference on Communications and Network Security (CNS)* (2019)
34. Rajivan, P., Gonzalez, C.: Creative Persuasion: A Study on Adversarial Behaviors and Strategies in Phishing Attacks. *Frontiers in Psychology* (2018)
35. Rosiello, A., Kirida, E., Kruegel, C., Ferrandi, F.: A layout-similarity-based approach for detecting phishing pages (2007)
36. Stajano, F., Wilson, P.: Understanding scam victims: seven principles for systems security (2011)
37. Wash, R., Cooper, M.M.: Who Provides Phishing Training? Facts, Stories, and People Like Me. In: *Conference on Human Factors in Computing Systems (CHI)* (2018)

38. Whittaker, C., Ryner, B., Nazif, M.: Large-Scale Automatic Classification of Phishing Pages. In: Network and Distributed System Security Symposium (NDSS) (2010)
39. Williams, E.J., Hinds, J., Joinson, A.N.: Exploring susceptibility to phishing in the workplace. International Journal of Human-Computer Studies (2018)
40. Wright, R., Jensen, M., Thatcher, J., Dinger, M., Marett, K.: Influence Techniques in Phishing Attacks: An Examination of Vulnerability and Resistance. In: Information Systems Research (2014)

## A Detailed Information

Table 4 provides more detailed numbers on the e-mails per hour with reference to the categories *sent* and *clicked*.

**Table 4.** Total number of sent and clicked e-mails along with the numbers and percentages per hour of the day.

	Sent	Clicked
<b>Total # of e-mails</b>	429,418	59,689
<b>Hour 0</b>	164 (0.04%)	307 ( 0.59%)
<b>Hour 1</b>	173 (0.04%)	406 ( 0.68%)
<b>Hour 2</b>	123 (0.03%)	366 ( 0.61%)
<b>Hour 3</b>	0 (0.00%)	257 ( 0.43%)
<b>Hour 4</b>	0 (0.00%)	212 ( 0.36%)
<b>Hour 5</b>	0 (0.00%)	533 ( 0.89%)
<b>Hour 6</b>	0 (0.00%)	1744 ( 2.92%)
<b>Hour 7</b>	540 (0.13%)	3926 ( 6.58%)
<b>Hour 8</b>	36246 (8.44%)	5955 ( 9.98%)
<b>Hour 9</b>	36026 (8.39%)	6058 (10.15%)
<b>Hour 10</b>	35684 (8.31%)	5191 ( 8.70%)
<b>Hour 11</b>	36515 (8.50%)	4673 ( 7.83%)
<b>Hour 12</b>	37511 (8.74%)	4429 ( 7.42%)
<b>Hour 13</b>	35836 (8.35%)	4708 ( 7.89%)
<b>Hour 14</b>	36040 (8.39%)	4694 ( 7.86%)
<b>Hour 15</b>	36217 (8.43%)	4091 ( 6.85%)
<b>Hour 16</b>	34557 (8.05%)	3317 ( 5.56%)
<b>Hour 17</b>	34642 (8.07%)	2520 ( 4.22%)
<b>Hour 18</b>	35266 (8.21%)	1880 ( 3.15%)
<b>Hour 19</b>	33208 (7.73%)	1629 ( 2.73%)
<b>Hour 20</b>	182 (0.04%)	1051 ( 1.76%)
<b>Hour 21</b>	158 (0.04%)	765 ( 1.28%)
<b>Hour 22</b>	156 (0.04%)	602 ( 1.01%)
<b>Hour 23</b>	174 (0.04%)	375 ( 0.63%)