Unpacking Spear Phishing Susceptibility

Zinaida Benenson[1(✉)], Freya Gassmann[2], and Robert Landwirth[1]

[1] Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany
{zinaida.benenson,robert.landwirth}@fau.de
[2] Universität des Saarlandes, Saarbrücken, Germany
f.gassmann@mx.uni-saarland.de

Abstract. We report the results of a field experiment where we sent to over 1200 university students an email or a Facebook message with a link to (non-existing) party pictures from a non-existing person, and later asked them about the reasons for their link clicking behavior. We registered a significant difference in clicking rates: 20% of email versus 42.5% of Facebook recipients clicked. The most frequently reported reason for clicking was curiosity (34%), followed by the explanations that the message fit recipient's expectations (27%). Moreover, 16% thought that they might know the sender. These results show that people's decisional heuristics are relatively easy to misuse in a targeted attack, making defense especially challenging.

Keywords: Spear phishing · Facebook · Decisional heuristics

1 Introduction

Phishing attacks that persuade users to click on malicious attachments or links have become a standard means of gaining an entry point into the systems during the APT (Advanced Persistent Threat) attacks and data breaches, and also have recently caused substantial damage in form of ransomware infections. The popularity of this attack vector has inspired numerous research efforts on susceptibility of the users to different targeting techniques and on user education [16]. Most of this research concentrated on link clicking in emails and submission of information on phishing webpages.

However, although harvesting users' login details via phishing websites and spreading malware through attachments remain important attack vectors, also just clicking on a link can result in a security incident. For example, according to two surveys published in 2016, email links leading to infected websites accounted for around 30% of malware infections in organizations [32,43].

Along with the phishing messages that address general Internet population, several variants of the so-called *spear phishing* evolved [10,17,30]. This term refers to phishing attacks targeted at specific individuals or groups, for example

Targeted Attacks Workshop at Financial Cryptography and Data Security 2017. ©IFCA.

M. Brenner et al. (Eds.): FC 2017 Workshops 2017, LNCS 10323, pp. 610–627, 2017.
https://doi.org/10.1007/978-3-319-70278-0_39

customers of a specific organization (bank, online retailer, telecommunications company) or employees in a specific department (human resources, accounting, customer support). Spear phishing messages can address victims by names, refer to their immediate interests or job tasks and appear to come from trusted senders [15,42].

Considering previous research, two areas remain relatively unexplored. Firstly, different media by which the phishing message could be received, such as email, Facebook or Twitter, could make a difference in success rates. Although phishing attacks via Facebook happen in practice [23], the first step towards direct comparisons of success rates between email and Facebook was made in our previous study in 2014 [4]. Secondly, researchers rarely directly asked users to explain the reasons behind their reactions to "suspicious" messages. Although some small-scale studies with 20 or less participants [6,13] interviewed users to find out how they would decide in a hypothetical scenario whether an email is legitimate or not, we are not aware of large-scale behavioral studies on this topic. In this work, we make the following contributions:

- We show in a between subjects field experiment with 1255 users that receiving the same message with a "suspicious" link via Facebook or via email leads to significantly different click rates. Our study partially replicates our previous study [4] and validates its results.
- We analyze the reasons for clicking and not clicking reported by the participants in a post-experiment survey and discuss how lessons learned from this experiment can be applied to a broader range of scenarios involving spear phishing attacks.

This paper is organized as follows. We present related work in the next section, and research questions and hypotheses in Sect. 3. We then elaborate on study method in Sect. 4. We present results of the behavioral field experiment in Sect. 5 and results of the post-experimental survey in Sect. 6. We discuss our findings and their implications in Sect. 7 and conclude in Sect. 8.

2 Related Work

Early works in the mid 2000's investigated the criteria according to which users categorize incoming emails as genuine messages or scam. Downs et al. [13] used interviews and role plays to analyze how users classify emails. Jakobsson and Ratkiewicz conducted so-called "context-aware" experiments in which they used publicly available data as well as the communication patterns of Ebay users to increase the plausibility of emails [20]. In another field experiment by Jagatic et al. [19], sending a phishing email that spoofed a social network friend increased the success rate from 16% (for emails from unknown senders) to 72%.

The results of these works indicate that users have difficulties in recognizing malicious emails, and that their corresponding decision criteria do not fit the problem. Five years later, a phishing study conducted by Blythe et al. [6] came to the conclusion that users still have the same difficulties, as they consider sender

address, design and language of an email as criteria for genuineness. They also cannot interpret technical details such as the composition of links.

Recognition of phishing websites also has been difficult for non-expert users. Their strategies, first uncovered in 2006 [12], still remained unsuccessful in 2015 [1]. Help provided by technical tools is also limited. Whereas passive indicators are rarely noticed by the users, active warnings are more often heeded [14,25,47]. Unfortunately, technical recognition of phishing websites, which is a precondition for effective warnings, still remains a challenging task. Most tools appear to have too high false positive and/or false negative detection rates [24,48].

Considering inability of non-expert users and of technical tools to reliably detect phishing attacks, education and training constitute alternative anti-phishing measures. Prominent academic tools for supporting anti-phishing user education and training are "Anti-Phishing Phil" [37] and "PhishGuru" [26]. In a comparative study of both systems [28], their developers found that both measures reduced the numbers of victims. The ability of non-experts to recognize (mostly non-targeted) phishing emails could be significantly increased from guessing (approx. 50% detection rates) to detection rates of 75–85%. Detection rates for users with initially higher expertise could be improved, using different education techniques, to nearly 100% [9,28,39].

Interestingly, similar educational efforts in a corporate environment proved to be unexpectedly challenging: majority of the users who clicked on a "suspicious" link that in reality led to training materials did not read these materials [10,27]. Moreover, although training effects were evident after one week in one study [27], these effects seemed to be lost after three months in another study [10].

Numerous studies measured factors that influence users' ability to recognize phishing emails, such as age, gender and technical background of the recipients, sender's gender and familiarity to the recipient, or design, spelling and content of the message. These measurements were conducted via surveys (e.g., [6,36,45]) or in behavioral studies that simulated phishing attacks (e.g., [10,19,20,28,31]). For example, emails with logos of the corresponding companies were significantly more difficult for the users to recognize as phish [6]. Some studies did not find any correlations between demographic factors and vulnerability for malicious messages [12,37], whereas others found that younger people (between 18 and 25) are more vulnerable than the middle-aged, and that women are more susceptible than men [5,19,36]. However, older adults (especially women) seem to be more vulnerable than younger adults [31].

Susceptibility of users to phishing attacks on various social media has been investigated to a lesser degree than susceptibility via email. Some studies considered acceptance rates of friend requests from strangers or from spoofed acquaintances on Facebook and other social networks, and the amount of information that can be gained from the users via this attack [5,38]. Another interesting research direction created fake social network profiles and observed which kind of friend requests they receive [18,40]. An automated infiltration attack built a network of fake accounts that successfully befriended more than 3000 users [7].

Also leveraging social network information for crafting spear phishing emails has been investigated [8,19]. A highly sophisticated method of leveraging Twitter for spear phishing was presented at BlackHat USA 2016 [35]. To the best of our knowledge, our research group conducted the first study that directly compared phishing susceptibility between email and Facebook [4].

As mentioned previously, although some small-scale studies interviewed users to find out how they *would* decide whether an email is legitimate or not [6,13], we are not aware of large-scale *behavioral* phishing studies that directly asked participants for the reasons of their clicking behavior.

Two studies combined a social engineering field experiment with a subsequent questionnaire similarly to our study. Vidas et al. [44] distributed flyers with "suspicious" QR-codes in different locations. Users that scanned a QR-code were taken to a website with a survey that asked them to indicate the main reason for their scanning action. Tischer et al. [41] distributed "suspicious" USB sticks on a university campus in a similar fashion. Users that found a stick and inserted it into a computer were also asked for reasons of their action. We further discuss their findings in Sect. 7. In contrast to our study, these two studies could only ask for the reasons of unsafe behavior, as users who behaved securely could not be reached by their surveys.

3 Research Questions and Hypotheses

The present study is a follow-up to a similar study we conducted in 2014 [4]. We partially replicate this previous study by considering its research question and the hypotheses H1–H5 presented below.

Research Question 1: Do people react to a "suspicious" link differently depending on whether the link was received via Facebook or via email?

Hypotheses: The following factors will be correlated to the higher success rate of the attack:

- H1: Message reception via Facebook,
- H2: Friend request from the sender,
- H3: Message sent from an open Facebook profile,
- H4: Female gender of the sender,
- H5: Female gender of the recipient.

These hypotheses were grounded in the previous work on demographic characteristics of phishing victims [36] and on social network phishing [5]. In study [4], none of the hypotheses could be supported. Whereas hypotheses H2-H5 did not yield any statistically significant results, the effect of Facebook was highly significant, but reversed: 56% of email recipients, but only 38% of Facebook recipients clicked on the "suspicious" link. Therefore, we decided to conduct a follow-up study to validate the findings of the previous study.

Moreover, effect sizes in the statistical analysis in study [4] were small, indicating that some other factors, unrelated to our hypotheses, led to clicking. This assumption resulted in the second research question for the present study:

Research Question 2: How do people explain their reasons for clicking or not clicking on a link?

To answer the above research questions, we designed a field experiment and a follow-up survey presented below.

4 Method

Win the following we present design of our study. In a nutshell, we conducted a field experiment where we sent to the participants an email or a personal Facebook message with a link from a non-existing person, claiming that the link leads to the pictures from a party. When clicked, the corresponding webpage showed the "access denied" message. We registered the click rates, and later sent to the participants a questionnaire that asked about the reasons for their clicking behavior.

4.1 Ethical Considerations

Jakobsson et al. [20,21] discuss the ethical issues of phishing studies in depth and arrive at the conclusion that, under certain circumstances, it is ethically permissible to conduct phishing studies without participants' consent and without debriefing. The above position is controversial, however, as experimenting with humans without their consent can negatively influence participants. For example, one of the first phishing experiments at the Indiana University [19] resulted in a serious controversy and media outcry as the students found out that they unwittingly participated in the study [29].

Therefore, we recruited the participants for a "cover story" survey of their Internet habits in order not to prime them about phishing. To offer an incentive for participation, we drew ten online shopping vouchers with the value of 10 EUR each. We fully debriefed participants after their participation by sending to them cumulative anonymized statistics about the study results and explaining why clicking on a link might result in a security incident. We also provided a possibility for anonymous study feedback, as well as a contact person for further questions. Our study plan was approved by the data protection office of the University of Erlangen-Nuremberg that verified its compliance with the German data protection laws and ethics.

4.2 Experimental Design

For sending the messages with links we created three email accounts (a male, a female and an anonymous account with unidentifiable gender) at a popular German provider, and four Facebook accounts, two male and two female. We used first names that were most popular in Germany around 1990 (the estimated years of birth of our participants, university students in their twenties), and the most popular German surnames, ending up with attacker names such as Sabrina Müller and Frank Bauer.

One male and one female Facebook account were "closed", that is, they contained only names and a symbolic male or female profile picture that Facebook shows by default. Two other profiles were "open", containing a profile photo, some other pictures, postings and friends, see Fig. 1.

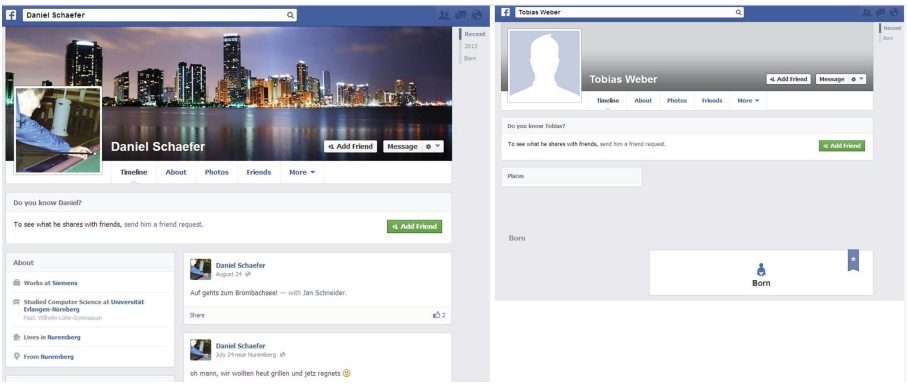

Fig. 1. Fake Facebook senders in the study: an open profile of Daniel Schäfer and a closed profile of Tobias Weber.

The field experiment started in the first week of January 2014. The participants were sent the following message with an individualized link via email or as a personal message on Facebook. The link contained an IP address from our university:

Hey!
The New Year's Eve party was awesome! Here are the pictures:
http://<*IP address*>/photocloud/page.php?h=<*participant ID*>
But please don't share them with people who have not been there!
See you next time!
<*sender's first name*>

When the users clicked on the link, their participant ID (a randomly assigned 7-digit number) was recorded in the database, and the website showed an "access denied" message.

4.3 Recruitment

The participants for the email-based study were recruited using the internal student mailing list of our university, whereas the participants for the Facebook-based study were recruited via the Facebook student groups of several German universities[1].

[1] We could not recruit enough Facebook participants at a single university and therefore used several universities.

We had a technical reason for recruiting Facebook participants via a Facebook group. At the time of the study, there were three folders in Facebook accounts into which new personal messages could be delivered: "Inbox", "Others" and "Spam". Users are only notified about new messages that are delivered into the Inbox. Furthermore, the users could choose between two settings called "Basic Filtering", which is the default setting, and "Strict Filtering" for incoming messages. We found out by experiment that users that chose strict filtering will always receive personal Facebook messages from strangers in the Others folder. However, if people use basic filtering, a message from a stranger will be delivered in the Inbox if the receiver and the sender are members of the same Facebook group. Thus, in order to make our message go to Inbox for as many participants as possible, we put our fake sender profiles into the Facebook participant group. As several potential participants in the Facebook groups explicitly asked us whether we want their email address as well, and commented that they are not willing to provide it, we recruited the email participants via email.

Participants were randomly assigned to all other experimental conditions: gender of the sender on both communication channels, friend request or no friend request from the sender on Facebook, open or closed sender profile on Facebook.

4.4 Sample Characteristics

We recruited 280 Facebook users (80 male, 200 female) and 975 email users (265 male, 710 female). Groups have a comparable gender structure with 27% and 29% of male participants, respectively.

Table 1. Key demographic facts about the participants. σ denotes standard deviation.

	All users	Email group	Facebook group
Recruited participants	1255 (28% male)	975 (27% male)	280 (29% male)
Survey response rate	57% (22% male)	56% (21% male)	62% (28% male)
Average age (survey)	23.1 ($\sigma = 4.4$)	23.2 ($\sigma = 4.1$)	22.9 ($\sigma = 5.1$)
% of students (survey)	93%	96%	86%

Other demographic characteristics of participants were not collected at the time of recruitment, but later during the survey. Therefore, these characteristics are only known for the survey participants. As presented in Table 1, response rate for the survey was 57%. The differences in response rates between the groups (56% for email and 62% for Facebook) are not statistically significant (Pearson's $\chi^2 = 2.98$, $p < 0.10$). Both groups have a comparable age structure (the differences are not statistically significant) and a strong majority of students. The number of students is significantly higher in the email group, although the effect is relatively small ($\chi^2(1) = 8.93$, $p < 0.001$, Cramer's V $\varphi_{\mathbf{c}} = 0.162$).

5 Behavioral Clicking Results: Facebook vs. Email

We extracted the behavioral clicking data from the web server logs. During this process, page requests by bots, such as Facebook or Google, were removed. We used the same statistical analysis method as in the previous study [4].

The descriptive results and the Pearson chi-squared (χ^2) test results with the effect size reported using Cramer's V (φ_c) are presented in Table 2. Just as in the previous study, hypotheses H2–H5 were not supported. However, H1 was supported. Thus, in both studies, the only significant clicking factor is the communication channel. In our study, 20% of email versus 42.5% of Facebook users clicked on the link. However, the channel effect in [4] was reversed: 56% of email users versus 38% of Facebook users clicked. We discuss this difference further in Sect. 7.

Table 2. Statistics for clicking rates. The only significant factor ($p < 0.001$) in the present study is the communication channel (Facebook versus email).

Factor	Clicked	χ^2	df	p	φ_c
Communication channel	Email: 194/975 (20%) FB: 119/280 (42.5%)	59.365	1	**0.000**	0.218
Sender's gender (email)	Female: 72/325 (22.1%) Male: 59/326 (18.1%) Undefined: 63/324 (19.4%)	1.742	2	0.419	0.042
Sender's gender (Facebook)	Female: 64/140 (45.7%) Male: 55/140 (39.3%)	1.184	1	0.277	0.065
Receiver's gender (email)	Female: 152/710 (21.4%) Male: 42/265 (15.8%)	3.742	1	0.053	0.062
Receiver's gender (Facebook)	Female: 86/200 (43.0%) Male: 33/80 (41.2%)	0.144	1	0.704	0.023
Friend request (FR) from sender (Facebook)	With FR: 58/120 (48.3%) no FR: 61/160 (38.1%)	2.924	1	0.087	0.102
Profile information of the sender (Facebook)	Closed: 64/140 (45.7%) Open: 55/140 (39.3%)	1.184	1	0.277	0.065

6 Reported Reasons for Clicking Behavior

In the survey, 117 out of 720 participants reported that they clicked, and 502 participants reported that they did not click. These participants were asked in a subsequent open-ended question to explain in their own words why they clicked or did not click. The rest of the participants reported that they either could not remember whether they clicked, or that they did not receive the message.

We analyzed participants' explanations of their clicking behavior according to principles of qualitative content analysis [34]. First, two researchers independently worked through the responses, identifying relevant topics and labeling them. These topics and labels were discussed and an initial coding frame was

designed. This initial coding frame was used in a first trial coding, spanning over the first one hundred responses, coded independently by both researchers.[2] During this process, each researcher took note of occurring coding problems. Post coding, these problems were discussed and the coding frame and its categories were revised accordingly. The refined coding frame was used to recode the initially coded replies and to also code the next hundred replies. This process was repeated until no more coding frame related problems arose during trial coding. After that, all data was coded by two independent raters using the final coding frame. To assess inter-rater reliability, Cohen's Kappa κ was calculated [11], and afterwards the cases with conflicting codes were discussed to produce agreement. During this discussion, full inter-rater agreement could be reached.

Replies of clickers were coded with seven categories. Cohen's κ for four categories indicated excellent agreement (over 0.75), while the remaining three showed good agreement (over 0.6).[3] Answers of non-clickers were coded with 20 categories. 19 categories had excellent Cohen's κ (over 0.75), and the remaining category had a good one (0.62). For interpretation purposes, we clustered some of these categories into more general categories.

6.1 Reasons for Clicking

The reported reasons for clicking were similar for the email and the Facebook groups (Table 3). By far the most frequent reason was *Curiosity*. These participants explained that they knew that the pictures cannot be for them, but were interested in the supposedly funny or private content.

Table 3. Categories for the clicking reasons (117 answers). Cohen's Kappa $\kappa > 0.75$ indicates excellent inter-rater agreement, $\kappa > 0.6$ means good agreement. Some participants reported more than one category.

Category	N	%	κ	Explanation
Curiosity	40	34.2	0.91	Curios about the pictures, interested to see their content
Context	32	27.4	0.82	Reception of the message fits the situation of the New Year's Eve celebration
Investigation	21	17.9	0.84	Wish to find out more about the situation that caused this message
Known sender	19	16.2	0.62	Certainty or assumption that one knows the sender
Technical context	13	11.1	0.9	Technical features (operating system, browser, antivirus, university's network) will thwart threats
Fear	8	6.8	0.92	Fear that a stranger may have pictures of the receiver
Automatic	4	3.4	0.71	Clicked without thinking, impulsively

[2] As the first question elicited only 117 responses, all these responses were processed during each coding step.

[3] We follow the interpretation of Cohen's Kappa by Banerjee et al. [3].

The second place was taken by the explanations that the message fits the *Context* of the New Year's Eve celebration, for example P151: *"I thought these were the pictures from the company's celebration, and all of us have been waiting for them."* P483 explained: *"I did not know many people from the New Year [...] and I thought it was one of them"*.

Some participants clicked in the course of an *Investigation*, as they wanted to find out more about the situation, and maybe to correct the "mistake": *"I wanted to see to whom the message was actually addressed and forward it if possible"* (P16). Users also thought that the message is from a *Known sender*, so P8 explained: *"I thought the message was from a friend whose name is also Sabrina by chance"*. This indicates that choosing most popular German names was a good strategy for targeting. Interestingly, two users explained that they thought it was some friend who used a pseudonymous account.

Participants also expressed trust into some technical measures, or in the ability of the university to protect them, so P711: *"I have never received spam at the university email address before"*, or P461: *"I knew that my Kaspersky will protect me"*. P490 considered the combination of Mac OS and Firefox "safe enough" for clicking. Four participants stated that as the IP address belonged to the university, they considered the link to be safe.

Eight participants said that they were anxious that a stranger might actually have pictures of them (*Fear* category), so P32: *"Although I felt unsafe, my fear that a stranger might have my pictures was very strong. There are so many possibilities nowadays to make photos that one never knows who might have made them, and under which circumstances"*.

Automatic reaction was also reported: *"I first clicked on the link and then it came to me that no person with this name was actually present"* (P33).

6.2 Reasons for Not Clicking

The most prominent reason for not clicking was the *Unknown sender* name (Table 4). Although unknown sender name is an important indicator of scam messages, only three users explicitly commented that one cannot fully rely on it, as dangerous messages can also arrive from known senders.

Many participants indicated that they suspected the link to contain malware or be fraudulent without explaining how they arrived at this conclusion (*Suspicion of Fraud* category). It seems that they relied on their intuition: *"I thought is was a virus"* (P137), *"Might have been a 'spy' link"* (P196), *"I knew immediately that this was spam"* (P385).

Some people reasoned that the context of the message reception did not fit. For example, *Situation context* was an important indicator, where users explained that no pictures were made at their party, or that they spent the New Year's Eve alone. Unfitting *Life context* means that there are no people or circumstances in the person's life that would cause such a message to be sent: *"My friends would not contact me in this way"* (P36), *"I do not receive this kind of mails"* (P238). Some people also remarked that they never share pictures via email (or via Facebook), or that they do not use this particular email address

Table 4. Categories for the reasons not to click (502 answers).

Category	N	%	κ	Explanation
Unknown sender	254	50.6	0.90	Sender of the message is unknown
Suspicion of Fraud[a]	250	49.8	0.93	Assumption that the message is fraudulent, phishing, might contain a virus
Situation context[a]	195	38.8	0.96	Reception of the message does not fit the situation of the New Year's Eve celebration
Life context[a]	58	11.6	0.75	There are no circumstances in the life of the recipient that would cause such a message
Rule of conduct	47	9.4	0.91	A behavioral rule prohibits clicking on links in such messages
Privacy	28	5.6	0.93	Private message sent to a wrong person
Message context[a]	27	5.4	0.87	Wrong communication channel or email address for a message like this
Message form[a]	25	5.0	0.91	Anonymous message, not addressed by name
Link form	20	4	0.93	Link looks suspicious
Bad experience	11	2.2	0.8	Unpleasant experience in a similar situation

[a]Indicates a merged category. Some participants reported more than one category.

for communication with their friends (category *Message context*), or that the message did not address them by name, or was "anonymous" (*Message form*).

Almost 10% of users said that they acted according to a specific *Rule of conduct*, for example they never open emails from unknown senders, or never click on "such" links. Some users mentioned the "strange" link (it contained a bare IP address, *Link form* category), or that they already had an unpleasant experience with clicking on a link (e.g., the link led to a porn site), or caught a virus after clicking on a link in a similar situation (*Bad experience*).

Respecting *Privacy* of other people was stated as a reason by 5.6% of users, for example P708 said: *"I do not look up a private message that was obviously not addressed to me"*. This reason can be considered as an antipode to the most frequently stated reason for clicking (curiosity).

7 Discussion

Although this study has some limitations, we think that useful preliminary conclusions can be drawn from our study and from its comparison to study [4]. Especially the highly significant difference between the communication channels and the reasons for clicking provide important insight into targeting strategies, as we discuss in the following.

7.1 Limitations

Findings of this study have several limitations. Thus, we did not assign the communication channel (Facebook or email) randomly to participants, and moreover,

email and Facebook groups were recruited at different universities. We also had different sample sizes for email (975) and for Facebook (280). Both user groups are skewed towards female participants. However, this bias might not be important as recipients' sex did not play any role in our and in the previous study [4].

Furthermore, reported reasons for actions do not always correspond to the real reasons, as people make many decisions based on intuition or subconsciously [22,46]. Thus, although we now know more about how people reason about targeted attacks, we might still not be able to predict their behavior. This should be verified in future studies.

7.2 Facebook versus Email

In the present study, 42.5% of Facebook users, but 20% of email users clicked on the link. We hoped to find the reasons for this statistically highly significant difference in the reasons for clicking and not clicking provided by the users. Surprisingly, reasons did not differ statistically across the platforms, although a small amount of non-clickers commented that they did not expect this kind of message to arrive via email, and a small amount of clickers commented that receiving pictures via Facebook seemed plausible to them.

Several factors might be responsible for susceptibility of Facebook users. Firstly, social networks such as Facebook or LinkedIn might be considered trustworthy by users, as Kirlappos and Sasse indicate [25]. Secondly, the special characteristics of the Facebook platform, such as informal communication and easy ways to find the profile of a recent acquaintance, might have made our message especially plausible there. Thirdly, handling the messages on Facebook might be different from handling the emails, such that the users scan through their many notifications very quickly, without paying attention to what they are actually doing.

7.3 How Powerful is Personalization?

Our previous study [4] provided inspiration for the present study, although we did not strictly replicate it. As mentioned above, 56% of email participants clicked in study [4], whereas only 20% of email participants clicked in our study. Clicking rates on Facebook were comparable: 38% in [4] and 42.5% in the present study. Due to differences in experimental setup, direct statistical comparison of the two studies is problematic, and therefore we consider mainly qualitative arguments in the following.

According to Table 5, Facebook groups in both studies have comparable sizes, but the email group in our study has significantly more participants. The participants in both studies have comparable age and occupation demographics, but study [4] has significantly more male participants. However, in both studies, participants' sex did not correlate to their clicking probability, and therefore, gender differences of the samples are unlikely to have influenced the differences in results. Messages sent on both studies were similar, but not identical. Especially, participants in study [4] were addressed by first name.

Table 5. Comparison of key features between study [4] and our study.

	Study [4]	Our study
Time frame	Summer 2013	Winter 2013/14
Participants	398 (61% male) 240 Facebook/158 email	1255 (28 % male) 280 Facebook/975 email
Average age	22 ($\sigma = 4.5$)	23 ($\sigma = 4.4$)
% of students	96%	93%
Message	Pictures from party last week	Pictures from New Year's Eve party (sent on January 7th)
Addressing	Hey <*receiver's first name*>	Hey!
Clicking rates	38% Facebook/56% email	42.5% Facebook/20% email

We hypothesize that addressing by first name plays the most important role in the differences between two studies in the email clicking rates. Indeed, for many years, the traditional security advice to consumers had been that legitimate emails would address them by names, but the scams would not. Recently, this advice has changed. For example, at the time of writing, the Anti-Phishing Working Group (APWG) states: *"Typically, phisher emails are not personalized, but they can be"* [2]. The 56% clicking rate in study [4] as opposed to 20% clicking rate in the present study, although the messages were fairly similar, indicates that personalization is especially important for targeted email-based attacks. On Facebook, however, addressing by first name does not seem to play an important role. This could be connected to the difference in user interface, as names of recipients are clearly visible for the senders, and therefore are not perceived by the recipients as something that a stranger cannot find out. Moreover, receiving an informally addressed message via Facebook might be more common than receiving such a message via email. We note, however, that these assumptions are not supported by evidence so far and need further investigation.

7.4 Lessons About Targeting and Spear Phishing Susceptibility

Curiosity seems to be a very powerful driver of risky Internet behavior. This was also noticed in the previous studies: 64% of survey respondents in study by Vidas et al. [44] scanned "suspicious" QR codes out of curiosity, and 18% of survey respondents in the study by Tischer et al. [41] plugged in a "suspicious" USB stick for this reason.[4] At the same time, a small amount of participants in our study was protected from the would-be danger by their lack of curiosity, or the wish to respect the privacy of the others.

Also the fitting the content and the context of the message to the current life situation of a person plays an important role. Many people did not click because

[4] Both of these studies could not reach participants that behaved in a safe manner, as they did not have any opportunity to provide them with a survey.

they learned to avoid messages from unknown senders, or with an unexpected content, as it might give them an unpleasant experience, such as a virus. For some participants, however, the same heuristic ("does this message fit my current situation?") led to the clicks, as they thought that the message might be from a person from their New Year's Eve party, or that they might know the sender.

7.5 Defense Against Spear Phishing

Defense against spear phishing and other targeted attacks seems to be especially challenging because of the ambiguity of the situations that they create, making the context and content of the message look plausible and legitimate. Because of this ambiguity, asking people to be permanently vigilant when they process their messages might have unintended negative consequences.

For example, if a person's job requires processing a lot of invoices sent via email, they might click on a ransomware-infected file called "invoice", as this fits their job expectations. And if they are taught to be "careful" with invoices, they might start missing or delaying the real ones, which stands in a direct conflict with the requirements of their job. Under these circumstances, the employees are likely to disregard this kind of user education attempts, because the only way for them to get their job done *in time* is to process their emails as quickly as possible, without "wasting" time with extra security checks.

In general, being suspicious of every message that was maybe sent in a hurry with typos from a mobile device, or is otherwise a bit strange, will deprive people from (usually reliable) decision heuristics such as "this message fits my current expectations" or "I know the sender", making them less efficient in their jobs, especially if these jobs require processing of a high number of messages.

In security practice, sending fake phishing emails to employees has become a popular method of assessing their security awareness, with numerous commercial tools designed for this purpose. However, trying to involve users into perimeter defense by means of catching them on dangerous actions, such as link clicking in fake phishing emails, might have unintended negative consequences. For example, employees of an organization may become disgruntled and unmotivated if they find out that they are being attacked by their own security staff [33], or start blaming themselves for inability to make a correct decision in an ambiguous and difficult situation [10]. Moreover, sending employees messages from spoofed colleagues, friends and bosses, although might raise their security awareness, may also seriously hamper their work effectiveness, and also social relationships within the organization, promoting the atmosphere of distrust.

We note, however, that although our study led us to hypothesize about negative consequences of the above human-centered defenses against spear phishing attacks, we do not have enough evidence to support these hypotheses. Thus, one of the most important directions for future research is development of study designs and measurement procedures for assessing not only effectiveness of anti-phishing measures, but also their impact on the work and life environment of people, and on their psychological well-being.

8 Conclusion

We conducted a study consisting of a link clicking field experiment on Facebook and via email, and a follow-up survey that investigated the reasons for clicking behavior. An important future work question is whether awareness of danger ("links can lead to infected sites") helps, and to what extent can people be expected to act rationally when they feel curiosity, or any other strong emotion. We think that expecting the full impulse control from the users is unrealistic.

This particular study revealed susceptibilities to scam in some people, and the reasons behind their susceptibility, but we think that the lesson learned is broader. By a careful design and timing of a message, it should be possible to make virtually any person click on a link, as any person will be curious about something, or interested in some topic, or find themselves in a life situation that fits the message's content and context. For example, the message might come from a known sender, or refer to a previous experience in a plausible way. In the long run, relying on technical in-depth defense may be a better solution, and more research and evidence is needed to determine which level of defense non-expert users are able to achieve through security education and training.

Acknowledgments. We thank Nadina Hintz, Andreas Luder and Gaston Pugliese for their invaluable help in data gathering and analysis. Zinaida Benenson and Robert Landwirth were supported by the Bavarian State Ministry of Education, Science and the Arts within the scope of research association FORSEC (www.bayforsec.de).

References

1. Alsharnouby, M., Alaca, F., Chiasson, S.: Why phishing still works: user strategies for combating phishing attacks. Int. J. Hum. Comput. Stud. **82**, 69–82 (2015)
2. Anti-Phishing Working Group (APWG): How to avoid phishing scams. http://www.apwg.org/resources/overview/avoid-phishing-scams
3. Banerjee, M., Capozzoli, M., McSweeney, L., Sinha, D.: Beyond kappa: a review of interrater agreement measures. Can. J. Stat. **27**(1), 3–23 (1999)
4. Benenson, Z., Girard, A., Hintz, N., Luder, A.: Susceptibility to URL-based Internet attacks: Facebook vs. email. In: 6th IEEE International Workshop on SEcurity and SOCial Networking (SESOC), pp. 604–609. IEEE (2014)
5. Bilge, L., Strufe, T., Balzarotti, D., Kirda, E.: All your contacts are belong to us: automated identity theft attacks an online social networks. In: 18th International Conference on World Wide Web (2009)
6. Blythe, M., Petrie, H., Clark, J.A.: F for fake: four studies on how we fall for Phish. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI 2011, pp. 3469–3478 (2011)
7. Boshmaf, Y., Muslukhov, I., Beznosov, K., Ripeanu, M.: The socialbot network: when bots socialize for fame and money. In: Proceedings of the 27th Annual Computer Security Applications Conference, pp. 93–102. ACM (2011)
8. Brown, G., Howe, T., Ihbe, M., Prakash, A., Borders, K.: Social networks and context-aware spam. In: Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work, pp. 403–412. ACM (2008)

9. Canova, G., Volkamer, M., Bergmann, C., Borza, R., Reinheimer, B., Stockhardt, S., Tenberg, R.: Learn to spot phishing URLs with the Android NoPhish App. In: Bishop, M., Miloslavskaya, N., Theocharidou, M. (eds.) WISE 2015. IAICT, vol. 453, pp. 87–100. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-18500-2_8
10. Caputo, D.D., Pfleeger, S.L., Freeman, J.D., Johnson, M.E.: Going spear phishing: exploring embedded training and awareness. IEEE Secur. Priv. **12**(1), 28–38 (2014)
11. Cohen, J.: A coefficient of agreement for nominal scales. Educ. Psychol. Measur. **20**(1), 36–47 (1960)
12. Dhamija, R., Tygar, J.D., Hearst, M.: Why phishing works. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI 2006, pp. 581–590 (2006)
13. Downs, J.S., Holbrook, M.B., Cranor, L.F.: Decision strategies and susceptibility to phishing. In: Proceedings of the Second Symposium on Usable Privacy and Security, SOUPS 2006, pp. 79–90 (2006)
14. Egelman, S., Cranor, L.F., Hong, J.: You've been warned: an empirical study of the effectiveness of web browser phishing warnings. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI 2008, pp. 1065–1074 (2008)
15. Goodin, D.: Crypto ransomware targets called by name in spear-phishing blast. Ars Technica, 4 April 2016
16. Hong, J.: The state of phishing attacks. Commun. ACM **55**(1), 74–81 (2012)
17. Infosec Institute: Spear Phishing: Real Life Examples. http://resources.infosecinstitute.com/spear-phishing-real-life-examples. Accessed Mar 2017
18. Irani, D., Balduzzi, M., Balzarotti, D., Kirda, E., Pu, C.: Reverse social engineering attacks in online social networks. In: Holz, T., Bos, H. (eds.) DIMVA 2011. LNCS, vol. 6739, pp. 55–74. Springer, Heidelberg (2011). https://doi.org/10.1007/978-3-642-22424-9_4
19. Jagatic, T.N., Johnson, N.A., Jakobsson, M., Menczer, F.: Social phishing. Commun. ACM **50**(10), 94–100 (2007)
20. Jakobsson, M., Ratkiewicz, J.: Designing ethical phishing experiments: a study of (ROT13) rOnl query features. In: 15th International Conference on World Wide Web (2006)
21. Jakobsson, M., Johnson, N., Finn, P.: Why and how to perform fraud experiments. IEEE Secur. Priv. **6**(2), 66–68 (2008)
22. Kahneman, D.: Thinking, Fast and Slow. Macmillan, Basingstoke (2011)
23. Kaspersky Lab Exposes Facebook Phishing Attacks: 10,000 Victims in Two Days June 2016. http://www.kaspersky.com/about/news/virus/2016/10000-Victims-in-Two-Days
24. Khonji, M., Iraqi, Y., Jones, A.: Phishing detection: a literature survey. IEEE Commun. Surv. Tutor. **15**(4), 2091–2121 (2013)
25. Kirlappos, I., Sasse, M.A.: Security education against phishing: a modest proposal for a major rethink. IEEE Secur. Priv. Mag. **10**(2), 24–32 (2012)
26. Kumaraguru, P., Cranshaw, J., Acquisti, A., Cranor, L.F., Hong, J., Blair, M.A., Pham, T.: School of Phish: a real-world evaluation of anti-phishing training. In: Symposium On Usable Privacy and Security (SOUPS) (2009)
27. Kumaraguru, P., Sheng, S., Acquisti, A., Cranor, L., Hong, J.: Lessons from a real world evaluation of anti-phishing training. Anti-Phishing Working Group (2008)
28. Kumaraguru, P., Sheng, S., Acquisti, A., Cranor, L.F., Hong, J.: Teaching Johnny not to fall for phish. ACM Trans. Internet Technol. (TOIT) **10**(2), 7 (2010)

29. Lenz, R.: In Indiana phishing study, students take the bait. USA Today, 23 July 2007. http://usatoday30.usatoday.com/tech/news/computersecurity/2007-07-23-phishing-study_N.htm
30. Northcutt, S.: Spear Phishing (Methods of Attack Series). https://www.sans.edu/cyber-research/security-laboratory/article/spear-phish. Accessed Mar 2017
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: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI 2017 (2017)
32. Osterman Research Survey: Understanding the Depth of the Global Ransomware Problem (2016)
33. Sasse, A.: Scaring and bullying people into security won't work. IEEE Secur. Priv. **13**(3), 80–83 (2015)
34. Schreier, M.: Qualitative Content Analysis in Practice. Sage Publications, Thousand Oaks (2012)
35. Seymour, J., Tully, P.: Weaponizing Data Science for Social Engineering: Automated E2E Spear Phishing on Twitter. Black Hat USA (2016)
36. Sheng, S., Holbrook, M., Kumaraguru, P., Cranor, L.F., Downs, J.: Who falls for phish?: a demographic analysis of phishing susceptibility and effectiveness of interventions. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 373–382. ACM (2010)
37. Sheng, S., Magnien, B., Kumaraguru, P., Acquisti, A., Cranor, L.F., Hong, J., Nunge, E.: Anti-phishing phil: the design and evaluation of a game that teaches people not to fall for phish. In: Proceedings of the 3rd Symposium on Usable Privacy and Security, SOUPS 2007, pp. 88–99 (2007)
38. Sophos: Facebook users at risk of "rubber duck" identity attack. https://www.sophos.com/en-us/press-office/press-releases/2009/12/facebook.aspx
39. Stockhardt, S., Reinheimer, B., Volkamer, M., Mayer, P., Kunz, A., Rack, P., Lehmann, D.: Teaching phishing-security: which way is best? In: Hoepman, J.-H., Katzenbeisser, S. (eds.) SEC 2016. IAICT, vol. 471, pp. 135–149. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-33630-5_10
40. Stringhini, G., Kruegel, C., Vigna, G.: Detecting spammers on social networks. In: 26th Annual Computer Security Applications Conference (2010)
41. Tischer, M., Durumeric, Z., Foster, S., Duan, S., Mori, A., Bursztein, E., Bailey, M.: Users really do plug in USB drives they find. In: 2016 IEEE Symposium on Security and Privacy (SP), pp. 306–319. IEEE (2016)
42. Vaas, L.: Beware the latest tax-season spear-phishing scam. https://nakedsecurity.sophos.com/2017/02/08/beware-the-latest-tax-season-spear-phishing-scam. Accessed Mar 2017
43. Verizon 2016 Data Breach Investigations Report (2016)
44. Vidas, T., Owusu, E., Wang, S., Zeng, C., Cranor, L.F., Christin, N.: QRishing: the susceptibility of smartphone users to QR code phishing attacks. In: Adams, A.A., Brenner, M., Smith, M. (eds.) FC 2013. LNCS, vol. 7862, pp. 52–69. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-41320-9_4
45. Vishwanath, A., Herath, T., Chen, R., Wang, J., Rao, H.R.: Why do people get phished? Testing individual differences in phishing vulnerability within an integrated, information processing model. Decis. Support Syst. **51**(3), 576–586 (2011)
46. Wilson, T.D.: Strangers to Ourselves. Harvard University Press, Cambridge (2004)

47. Wu, M., Miller, R.C., Garfinkel, S.L.: Do security toolbars actually prevent phishing attacks? In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 601–610. ACM (2006)
48. Zhang, Y., Egelman, S., Cranor, L., Hong, J.: Phinding phish: evaluating anti-phishing tools. In: Proceedings of the 14th Annual Network and Distributed System Security Symposium (NDSS) (2007)