Should Interventions be Prevented? A Research into the Right Way to Prevent People for Phishing Attacks

Floris-Jan Overink
University of Twente
P.O. Box 217, 7500AE Enschede
The Netherlands
F.Overink@student.utwente.nl

ABSTRACT
Nowadays there is a high need in protecting the human being against cyber-attacks. Cyber-attacks are not only carried out against machines but also against humans. Attackers use social engineering, phishing and spear phishing techniques for persuading the victims in giving personal identifiable information. What can be done to prevent these attacks? Is it possible to raise awareness via intervention messages? In this study two brief interventions are tested using questionnaires. The first intervention asks questions about cybercrime. The second intervention uses a warning message to create awareness. The interventions are both tested on different research groups. In the end results are compared using a control group. Results show that there are no major differences between the different kind of interventions. However the results show that approximately one third of the research population is a potential victim of a spear phishing attack based on the obtained personal information.

Keywords
Social Engineering, Phishing, Spear Phishing, Intervention, Self-Disclosure.

1. INTRODUCTION
1.1 Research aim
The increasing use of new Information and Communication Technologies in our daily life has increased our vulnerability to new cyber-attacks[4][12]. Alongside the cyber-criminal professionalization, the ICT threats have become more sophisticated and their impact can be very significant in our lives. In parallel, today’s security mechanisms are not adapted sufficiently yet, to face these new types of ICT attacks [17]. There are several ways to carry out successful ICT attacks, one of which is via personal information. A way to obtain personal information is through Social Engineering.

Social Engineering is a technique used by hackers or other attackers to gain access to information technology systems by obtaining the needed information from a person rather than breaking into the system through electronic or algorithmic hacking methods. Such attacks occur on both physical and psychological levels [14]. To gather personal information social engineers use more and more different types of techniques. They often carry out phishing attacks or even spear phishing attacks[8][9].

Phishing is a form of deception in which an attacker attempts to fraudulently acquire sensitive personal information from a victim by impersonating a trustworthy entity. Phishing attacks typically employ generic “lures”. For instance a phisher misrepresenting himself as a large banking corporation or popular online auction site will have a reasonable yield despite knowing little or nothing about the recipient[9]. Spear phishing attacks are much more targeted and involve duping particular individuals. These attacks are successful as phishers send customized, credible emails that appear to come from a trusted source. Spear-phishing attacks have success rate of 19 percent, compared to just 5 percent for standard phishing attacks [15]. So cybercrime has major consequences, including financial matters [4]. Therefore, research is needed on preventing phishing attacks.

In order to prevent these phishing attacks, various forms of interventions have been developed. Two of these interventions include training and raising awareness against cyber-attacks. One of these studies is focused on an anti-phishing approach that uses training intervention [1]. But they are not all that effective. Phish-Guru training systems teach users how to identify phishing attacks [11]. The system sends out simulated phishing emails and delivers training messages when users click on the included URLs. The effectiveness of the system was tested with 515 participants; 28 days after the first email, despite being given training more than once, 17.5 percent of participants still entered personal details onto simulated phishing websites. This was a significant improvement in comparison with the 40.1 percent a control condition revealed beforehand. However, this still leaves one in five users vulnerable after the improvement [10]. A possible reason current education and training efforts are not effective is because they assume users are keen to avoid risks, and thus likely to adopt behaviors that might protect them. For example, most online shoppers are looking for good deals. They start from a search engine and get presented links to various websites that present (often very tempting) offers. Unfortunately, training is not always effective [10].

Another method consists of a more direct intervention. Beunder et al. (2015) studied the effect on disclosure of personal identifiable information [3]. They wanted to find out whether people easily share their personal information and if they are less likely to do so with risks, especially cybercrime, in mind. The result of their research was striking. As an unexpected result awareness questions did not have a positive effect at all. Where does this apparent contradiction originate from?

Beunder et al. (2015) discussed the results in the discussion of their paper. They presented three possible reasons for this observed phenomenon. The first reason had to do with the awareness-generating questions. The questions created a sense of trustworthiness. When someone informs a respondents about the dangers of phishing that person will less likely take personally part in phishing practices targeted at the respondents. Beunder et al. (2015) made the equation with “Scareware”; software which is actually malicious but marketed as legitimate software, such as a
fake anti-virus [3]. A second reason had to do with creating trustworthiness. Beunder et al. (2015) told the respondents that they were students from the University of Twente. The respondents could see that Beunder et al. (2015) were from the University of Twente because of the used paper, pencils, clothing and possibility to check their college cards. This could have raised additional trustworthiness with the respondents. The last reason is that the questions that Beunder et al. (2015) tried to create awareness, did in fact not create awareness at all.

The present research aims to explain how it is possible to use intervention messages to create awareness. The findings of Beunder et al. are used to find a way to use interventions the right way as they were intended to. So research is being done into creating awareness.

1.2 Structure

This paper starts with an introduction about social engineering, phishing and previous work done by Beunder et al. (2015). After the introduction the research questions are presented. These research questions are discussed in the theoretical background. After the theoretical background, the research design is presented. The research design describes how the data was collected, what questionnaire is used, what the sample size is, what variables are constructed and how these variables are analyzed. After the research design results are presented. Once the results are presented, a short conclusion follows and followed by a discussion about the strengths and weaknesses of the study. In the end a conclusion is made based on the results and the discussion and some recommendations are stated.

1.3 Research Questions

This study answers two research questions. The first research questions of this paper is:

1. Do awareness raising questions lead to lower information disclosure, or higher information disclosure?

Beunder et al. (2015) reported that giving information is reciprocated with same type of information.

The second research question is:

2. Does an entirely different intervention, namely a much more direct warning not to give information to others, lead to similar rate of disclosure as in the control group?

1.4 Theoretical Background

There are several theories that present statements about influencing the human behavior. Perhaps one of these theories can explain why the interventions of Beunder et al. (2015) did not help.

The intervention questions should warn a person for a particular danger and if possible the questions should also persuade a person in observing the danger. Robert Cialdini describes six principles of persuasion. Each principle is examined as to its ability to produce a distinct kind of automatic, mindless compliance from people, that is, a willingness to say yes without thinking first [6].

1. Reciprocation: This principle states that we should try to repay, in kind, what another person has provided us. If a woman does us a favor, we should do her one in return; if a man sends us a birthday present, we should remember his birthday with a gift of our own. 2. Commitment and consistency: Once we have made a choice or taken a stand, we will encounter personal and interpersonal pressures to behave consistently with that commitment. Those pressures will cause us to respond in ways that justify our earlier decisions. 3. Social proofs: When people are uncertain about a course of action, they tend to look to those around them to guide their decisions and actions. They especially want to know what everyone else is doing. 4. Liking: People prefer to say ‘yes’ to those they know and like. People are also more likely to favor those who are physically attractive, similar to themselves, or who give them compliments. Even something as ‘random’ as having the same name as your prospects can increase your chances. 5. Authority: People respect authority. They want to follow the lead of real experts. Business titles, impressive clothing, and even driving an expensive automobile are proven factor in lending credibility to any individual. Giving the appearance of authority actually increases the likelihood that others will comply with requests – even if their authority is illegitimate. 6. Scarcity: In fundamental economic theory, scarcity relates to supply and demand. Basically, the less there is of something, the more valuable it is. The more rare and uncommon a thing, the more people want it. Familiar examples are frenzies over the latest holiday toy, or urban campers waiting overnight to pounce on the latest iPhone[16].

What could be another way to persuade a respondent? Bada et al. (2014) describe fear as an influencing strategy. Bada et al. (2014) explain that a meta-analysis research conducted on communication invoking fear held between 1953 and 1980 showed that increases in perceived level of fear led to increases in the acceptance of the proposed adjustment or Behavioural intention. When researchers refer to a strong condition of fear invocation, they usually mean that the message represents a big threat and the recipient perceived a big threat. Typically, the invocations of fear offer recommendations that are as efficacious in preventing the threat [2]. So if an intervention raises fear does it also raise awareness?

Based on one of the principles of Cialdini it is possible to give an explanation on the result of the null hypothesis. People who have been set to thinking about privacy issues will provide less personally identifiable information, of Beunder et al. (2015)[3]. Principle 5 states that people respect authority. Beunder et al. (2015) told and showed their respondents that they were students from the University of Twente. This could ensure that people want to participate in the questionnaire and answer all the questions for the students of the University. Principle 5 also provides an explanation for the second reason Beunder et al. (2015) gave for their result. Because Beunder et al. (2015) stated that they were from the University of Twente and the respondents could also see it. This could emanate authority to the respondent. What causes the respondent to carry out the tasks put forth by the questionnaire. This could cause the adverse effect of the intervention questions.

Another factor can be tailoring. Tailoring is used in the social engineering and is thus used combined with phishing attacks. According to Groves et al. (2014) tailoring can be used to increase the relevance to the sample person of information given about the interview request. This is beneficial to both the interviewer and the person. It may also be used to heighten the saliency of some minor but attractive feature of the request while distracting respondent attention from strongly unattractive components[7].

Based on principle 1, reciprocation, it is possible that people want to answer the questions in the questionnaire as good as possible after they have read the intervention questions, because they want to reciprocate something for the questioner. Or the people want to help the student with completing his thesis, so they don’t care about the intervention.

At last principle 4, liking, could be used to explain the effect of the intervention messages. In the questionnaire were some awareness-generating questions. These questions could raise a sense of trustworthiness. Which ensures that the respondent will answer the questions without doubting about what those questions meant.
So according to the principles of Cialdini and the research on fear done by Bada et al. (2014) it is possible to provide an explanation for the result of the hypothesis of Beunder et al. (2015). But to get a better reference the questionnaires of Beunder et al. will be re-done.

There is also a different way of using interventions. Beunder et al. (2015) used awareness raising questions. Bullée et al. (2015) used a different intervention method based on a poster. [5]. Bullée et al. (2015) created a poster with warnings signs. These are based on Kumaraguru et al. (2009) landing pages [11]. They did research into landing pages in order to prevent people of becoming a phishing victim. They used one single page as intervention page. This page exists of four questions how the phishing target can help himself and others. To use these questions in this research the questions need to be short and brief. For example two of these questions could be transformed into questions for a new intervention poster or leaflet. “How were you tricked”, could be transformed into “How does a phisher attack?” and “How phishers trick you into giving out personal information” could be transformed into “What does a phisher want?” [11]. The reason why these question need to be transformed lies in the fact that they needed to be short and effective. Wolgalter (1961) states that if you create a warning design with wording, the use of text should be as little as necessary to clearly convey the message. Besides this he states that you should be explicit. Tell exactly what the hazard is and what the consequences are [18]. Finally Leonard (1999) did research into signal words. Leonard (1999) wrote that if you turn a signal word into the red color, the dominant response of red is associated with risks at all levels [13]. So if signal words are used, it could be more useful to turn them into the color red.

Based on the research questions and the information above, the following two hypotheses H1: “The findings of Beunder et al. (2015) can be replicated.”
H2: “The questionnaire with the new intervention will perform as intended, so there will be a different rate on disclosure.”

2. RESEARCH DESIGN
The data is collected by means of three questionnaires. Data is collected at several places in the city center of Enschede in the main shopping area. The researcher approached potential respondents with a small introduction of himself and the main focus of the questionnaire.

2.1 Data Collection
The data is collected by approaching people at the van Heeksquare in Enschede and subsequently asking whether they would like to participate in the questionnaire [3]. Because of a larger research group minors are also allowed to fill in the questionnaire.

The research was carried out on five different days in May 2015. Beunder et al. (2015) did the questionnaires on a Saturday, Tuesday and Friday. The aim is to create a comparable setting as used by Beunder et al. (2015) to exclude distorting factors. Subsequently data collection is done on at least one Saturday and a Tuesday because there is a big and small market on those days at the van Heeksquare in Enschede. In Table 1 is the planning shown for the data collection.

2.2 Questionnaire Design
The questionnaire is based on Beunder et al. (2015)[3]. It consists of two parts. One is a common part that all respondents answered, and there was a series of questions on cybercrime meant to raise the security awareness of the respondents that was answered only in the ‘awareness’ condition. There were three conditions:

1. The awareness condition. In this condition respondents had to answer 4 questions about cybercrime, namely about if they were familiar with phishing, if they knew how much personal information is stated on the internet, how their Facebook privacy settings are and finally if the respondents has ever been scammed via the internet. These questions are identical to Beunder et al. (2015) and can be found in appendix A [3]. They were placed in the middle of the questionnaire.

2. The warning condition. In this condition, prior to answering any question, respondents were handed over a leaflet of A4 format, the leaflet is shown in Figure 1. In addition, a small piece of this leaflet was placed at the top of each page, as a reminder. This small piece can be found in Figure 2. This leaflet was inspired by Bullée et al. (2015) Bullée et al. (2015) created a poster for intervention in a research into reducing the success of social engineering attacks [5].

3. A control group answered only the common choice of the questionnaire.

Table 1 Block diagram for questionnaire

<table>
<thead>
<tr>
<th></th>
<th>Mon 18-5</th>
<th>Tue 19-5</th>
<th>Fri 22-5</th>
<th>Sat 23-5</th>
<th>Tue 26-5</th>
</tr>
</thead>
<tbody>
<tr>
<td>10:00</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>Q3</td>
<td>Q1</td>
</tr>
<tr>
<td>11:15</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>Q2</td>
<td>Q3</td>
</tr>
<tr>
<td>12:30</td>
<td>Break</td>
<td>Break</td>
<td>Break</td>
<td>Break</td>
<td>Break</td>
</tr>
<tr>
<td>13:00</td>
<td>Q1</td>
<td>Q2</td>
<td>Q3</td>
<td>Q1</td>
<td>Q2</td>
</tr>
<tr>
<td>14:15</td>
<td>Q2</td>
<td>Q1</td>
<td>Q2</td>
<td>Q3</td>
<td>Q1</td>
</tr>
<tr>
<td>15:30</td>
<td>X</td>
<td>X</td>
<td>Q1</td>
<td>Q2</td>
<td>Q3</td>
</tr>
<tr>
<td>16:45</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

Figure 1: Warning Leaflet
2.3 Sample size
The target was to interview 100 respondents for each experimental condition. This number is chosen after considering a balance between on the one hand obtaining results out of which statistically significant conclusions can be drawn and on the other hand the feasibility given the time frame and available resources. For statistical and methodological purposes data is also collected on the population specific properties such as gender and age, as mentioned before.[3]

2.4 Variables
There are four outcome variables, namely writing down an email address, giving information on one’s bank account and giving information about online shopping history.

1) E-mail address was asked at the beginning of the questionnaire. Respondents were first asked if they wanted to receive a copy of the results, and then if they could fill in their email address, this was also meant to prevent double respondents. As this was done in the beginning of the questionnaire the awareness and control condition are identical at this aspect, as the awareness raising question have not been asked at this point. However, the warning condition had already received the warning leaflet. Answer categories were: email address filled in (1) or not filled in (0).

2) The bank account information was asked towards the end of the questionnaire. The question asked for a part of the respondent bank account number, because of privacy issues. Respondents needed to show if they knew their IBAN number out of their minds. An IBAN number is your bank account number plus your national letters, a control number and an abbreviation of your banks name. The way the respondents could show if they knew their IBAN number was done via a figure. This figure is shown in Figure 3.

Figure 3 IBAN number fill in field.
The answer categories were up to the respondents. They had their own choice whether the wanted to fill it in or not.

3) The respondents were also asked for their online shopping history. Only if respondents have ever bought something online. The answer categories were film/music, clothes, books, furniture and additional.

4) Besides what the respondents bought online, there was a question about where they bought their products. The answer categories were some major Dutch online shops and the category additional.

5) Each group has a control number. The control group has control number one. The group with the awareness condition has control number two. The group with the warning condition has control number three. These control numbers are used to put the data into SPSS.

Risk variable
A risk variable can be calculated with the help of the variables stated above. When someone uses the results of the four variables named above, a spear-phishing attack can be organized.

Firstly, variable 2, 3 and 4 need to be coded.

2) The results of the bank account information could be coded into the following: If nothing was filled in answering category (0), zero points. This means that it is not usable for phishing attacks. If the first six empty squares were filled in (1), one point. In this case it is semi useful for a phishing attack. If everything was filled in (2), two points, it could be used for a phishing attack.

3 and 4) Variable 3 and 4 can be coded in the same way. If nothing was filled in (0), zero points. This means also that it is not usable for a phishing attack. If they filled in a product or web shop the code is (1), one point. So it is usable for phishing attacks.

So the results of these questions need to be summed up. So if someone filled in his e-mail, bank account information and shopping history he scores 4 points. 1 point from the email question, 2 points from the bank account information question, 1 point from the product information question and 1 point from the shop information. So the respondents with score 4 can easily be attacked via a spear-phishing attack.

Socio-demographic variables
Information on age, sex and educational level was recorded.

2.5 Data Analysis
The data analysis is done by using SPSS 22.0.

The analyses are done using frequency tables and Chi-Square tests.

The social demographic analysis is based on gender and age.

3. RESULTS
In total 290 persons filled in the questionnaire. Five respondents were online two days or less on a weekly basis, which means that the data of these respondents could not be used. The questionnaires of these respondents were subsequently not used in the analysis. Below the results are presented. In Table 2 the numbers and percentages of the respondents in each research group can be found. In total 290 questionnaires are collected. The control group existed out of 98 respondents. The awareness group out of 97 respondents. At last, the warning group existed out of 95 respondents.

<table>
<thead>
<tr>
<th>Table 2 Number of respondents in each group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>----------------------</td>
</tr>
<tr>
<td>1 Control group</td>
</tr>
<tr>
<td>2 Awareness Group</td>
</tr>
<tr>
<td>3 Warning Group</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>

Based on the analysis of the data a decision is made about the result of the question about the average hours online. People who answered that they were less than 3 hours online in a week are not taken into account in the analysis. Leaving out these respondents, the total number of respondents numbers equals 281. An overview of the social demographic analysis can be found in Table 3. The male/female ratio is based on a Chi-Square test. The average number of hours and days a week online was obtained via frequency analyses and taking the mean of the data. At a first glance the groups do not differ a lot from each other. The male/female-ratio is almost the same in the three groups, spanning roughly between 55% and 60%, with a p-value of 0,891. The other variables are also not that different.
Hypothesis 1: “The findings of Beunder et al. (2015) can be replicated.”

Beunder et al. (2015) concluded that awareness group had more disclosure on personal information than the control group. When you compare the results of the Chi-Square test, done on the results of the data, in Table 4 there are no major differences between the research groups. The biggest difference is whether the respondents fill in their email or not. But this is between the control group and awareness group versus the warning group. To check the validity of H1 only the data from the control group and awareness group is used, because the warning group was not taken into account by Beunder et al. (2015). So when you compare the results of the control group and awareness group there are no big differences. The differences between whether the respondents filled in the kind of products they bought online or not are the smallest. The Chi-Square test showed that this p-value is 0.652. The differences on the other variables are not more than 10 percent. The p-value for whether the respondent filled in the web shop they used is 0.374. At last, the p-value for whether the IBAN number is usable for phishing is 0.598. Concluded the differences are that small that it is not statistical significant.

Hypothesis 2: “The questionnaire with the new intervention will perform as intended, so there will be a different rate on disclosure.”

The results, needed to test hypothesis 2, can be found in Table 4 and 5. When you compare the results of a Chi-Square test, done on the results of the data, in Table 4 and 5 over the research groups then there are no major differences between the research groups. The biggest difference is again if the respondent fill in their email or not. The Chi-Square test showed that the p-value for this variable is 0.01. This means that the difference is statistical significant. But the other three variables are not that different. The Chi-Square test showed that the p-value for Table 5 is 0.220. So the results of Table 5 are also not statistical significant. Concluded the overall differences are that small that they are not statistical significant.

4. CONCLUSION

The aim of the present study is to find out whether it is possible to prevent people of disclosing personal information. Beunder et al. (2015) did research into this study field. They discovered that giving information is reciprocated with the same type of information. However, the N in their study was low, which is could possibly be a threat to the internal validity of the study. In this thesis the number of respondents is higher than the Beunder et al. (2015) study, henceforth increasing reliability.

The results show us that there are no clear differences between the research groups. All groups have around 100 respondents and there are no big differences between the groups based on the social demographic results. A chi-square test showed that there was no significant difference between the groups.

In order to answer question 1 it is checked if hypothesis 1 can be confirmed or rejected.

H1: “The findings of Beunder et al. (2015) can be replicated.”

Because there are no major differences found in the results. The subject who received an awareness intervention did not differ from the subjects who received a warning or from the subjects from the control group. In all groups 40% to 50% give enough information on their bank account to be usable for a phishing attack. The results are not statistically significant. Therefore hypothesis one will be rejected. This means research question can be answered.

Question 1: Do awareness raising questions lead to lower information disclosure, or higher information disclosure?

It can be concluded, based on the results testing hypothesis 1, that the results from Beunder et al. (2015) could not be replicated. There were no major differences in the results to assume this. So as an answer to the first research question. Awareness raising questions do not lead to lower information or higher information disclosure.

In order to answer question 2 there will be checked if hypothesis 2 can be accepted.

H2. “Does an entirely different intervention, namely a much more direct warning not to give information to others, lead to similar rate of disclosure as in the control group?”

<table>
<thead>
<tr>
<th>Table 3 Social demographic information and hours online a day or week</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Females</td>
</tr>
<tr>
<td>Mean age (years and months)</td>
</tr>
<tr>
<td>Mean number of hours online</td>
</tr>
<tr>
<td>Mean number of days a week online</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 4 Results questions suitable for phishing attack</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Email filled in</td>
</tr>
<tr>
<td>Filled in the kind of product the respondent bought online</td>
</tr>
<tr>
<td>Filled in the web shop</td>
</tr>
<tr>
<td>IBAN Usable for phishing</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 5 Results of the risk group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>0, 1 or 2 points, no risk for phishing</td>
</tr>
<tr>
<td>3 points semi risk for phishing</td>
</tr>
<tr>
<td>4 points, full risk for phishing</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>
The results show also that there are no statistically significant differences between the control group, awareness group and the warning group. Accordingly, hypothesis 2 can also be rejected. So now question two can be answered.

Question 2. Does an entirely different intervention, namely a much more direct warning not to give information to others, lead to similar rate of disclosure as in the control group?

Now we know that H2 also is rejected. This means that the new intervention, with warning condition, wasn’t any better in its results comparing to the intervention with awareness conditions. As an answer to the second research questions; it looks like that is does not matter if you use an intervention based on a warning condition or an awareness condition. An intervention, based on warning conditions, does not lead to less personal disclosure from the respondents than a warning based on awareness conditions. The reason for this is not clear. It can be a general problem that it does not matter what intervention method you use. But it can also be a problem with the respondents. That the respondents were not that aware over the underlying idea of the questionnaire.

Beun der et al. had a relative small research population. This could lead to some interesting results because of the small research population they used. The results of this research were not the same as that of Beun der et al. So awareness-raising questions do not lead to lower or higher information disclosure. Even if you take the same numbers Beun der et al. used. The results will not be the same. So if this results was the benefit of the questionnaires or respondents cannot be said.

In the end, it is possible to prevent people, by an intervention, in disclosing personal information. But results did not say which intervention method is better. What can be said is that still too many people disclose their personal information to strangers. With the results of this research a number of 95 respondents can be attacked via a spear phishing attack. This is almost one third of the research population! So there is still a lot research to do into intervention and people disclosing personal information.

5. DISCUSSION

In this chapter the results of this thesis are discussed. The focus lies with the reliability and validity of the study.

In Table 4 and 5 are the results stated for answering the research questions. The biggest difference in these results can be found in Table 4. You can see that the warning groups percentage, about if the respondents fill in their email, is lower than the other groups. There is an average difference of almost 20%. The reason for this lies in the fact that in the questionnaire for the warning condition the email question were state as last. Instead as the first question. So it could be that the respondents did not want to spend more time into the questionnaire and just finish it as fast as possible. If the difference could be the benefit of the intervention, the other results should also be more different. The differences in the results between the research of Beun der et al. and this research could be explained by the number of respondents. Beun der et al. (2015) had 47 respondents. For this research almost 300 respondents cooperated. 200 of the respondents took part in the research into question one. That is more than 300%. So with these facts in mind, there can be said that the results of Beun der et al. are less statistically substantiated and therefore are the results of this research more likely to be true.

As shown in Table 1 a block diagram was used for the data collection. Unfortunately the weather was an important factor in gathering the data. The first day there were no problems with the weather. It was cloudy but it was not a nuisance. The collection went well. The second collection day gave the first problems. There were several rain showers, so it became impossible to gather enough data that day. So on the second day there was already a deviation from the schedule. The same happened on the Saturday morning. On this Saturday were also a lot of rain showers. The first two hours were also canceled. In the end the most data was gathered on Saturday afternoon and the last Tuesday. So there can be said that the data was not collected on equal moments, this could have lead to inconsistent results.

I made also several observations during the surveying process. When I approached some couples of groups to take part in my questionnaire, their first reaction was: “We got to help that student, he needs to get his numbers”. So people wanted to help me with my study. This could mean that the participants were not aware of any suspicious activities and that they shared all the information they could share based on authority of the questioner or they wanted to reciprocate. Sometimes the respondent reacted at first: “How long does it take?” I always responded all the time, that it takes several minutes and it does not cost much time. When they heard this, they were willing to take part in my questionnaire. But this group of participants did not take, in my opinion, enough time to check the intervention or thought enough about the underlying idea. This could be the benefit of tailoring. When I told the respondents that it does not take much time, it gave them a reaction they wanted to hear. So people filled in the questionnaire and sometimes they even hurried in filling it in.

In the end I also noticed something odd. When a couple or group was filling in the questionnaire, they did it separately until the question about the IBAN number. When they were arrived at that question, it sometimes became a small battle. A battle of ‘who knows the most numbers’. The reason why this happened can be described via social proof. Maybe it was uncertainty from the respondents and were they checking upon their friend. This could also be seen as a signal that the respondents where not aware of the underlying idea.

Based on the observations the results could definitely be influenced by the respondents. This is shown by that some of the observations could be substantiated with the help of the Cialdini principles. Some of the respondents were not aware of the underlying idea and at the same time lazy or not motivated enough to fill in the questionnaire safely. Because of the lack of motivation of some of the respondents it is now possible to try to obtain money from them via a spear phishing attack.

This discussion shows that there is more than only the obtained data. Based on the results of this research and the discussion recommendations can be done.

6. RECOMMENDATIONS

In this last chapter some recommendations are presented. First there research can be done into the results of the questionnaire with the awareness condition. The results of the intervention question were not analyzed because there was not enough time. Maybe there is a pattern between the disclosing rate on personal information and the answers on the intervention questions. Secondly the research could be redone with a clearer research population. Sometimes the respondents did not take enough time to look at the intervention or think about the meaning of the survey. The respondents wanted to help the student and go on with their pursuits. So if you use a clearer research population, you can ensure that they take enough time for the questionnaire. At last there can be done some more research into the interventions according to a specific time period. What if you ask the respondents a week or two later the same question? Do they still know something about the last intervention, or is their mind reset and is it all new to them?
7. REFERENCES


APPENDIX

A. Questionnaire Including Awareness (Question 8, 9, 10 and 11)

Ik ben Floris-Jan Overink en ik ben vierdejaars student Business & IT. Deze enquête neem ik af voor mijn afstudeeronderzoek naar phishing. Door middel van deze enquête probeer ik er achter te komen wat mensen van phishing weten en hoe gevoelig ze hier voor zijn. Verder doe ik ook onderzoek naar manieren om er voor te zorgen dat mensen geen slachtoffer worden van phishing aanvallen.

Mag ik beginnen met wat algemene vragen over uzelf?

1. Om te voorkomen dat deze enquête meerdere keren door dezelfde persoon wordt ingevuld vragen wij u om uw email-adres te noteren. Dit adres zal verder niet worden gebruikt, tenzij u aangeeft dat u de onderzoeksresultaten wenst te ontvangen.

............................................................................. @ ................................................

2. Wenst u na afloop van dit onderzoek de resultaten per email te ontvangen?
   ☐ Ja  ☐ Nee

3. Bent u man of vrouw?
   ☐ Man  ☐ Vrouw

4. Wat is uw geboortedatum?
   ........... - ........... - ...........

5. Wat is uw hoogste genoteerde opleiding? (Na keuze, doorstrepen wat niet van toepassing is)

   ☐ Basis onderwijs – lager onderwijs    Afgestudeerd/Mee bezig
   ☐ VMBO – mavo                        Afgestudeerd/Mee bezig
   ☐ MBO                                Afgestudeerd/Mee bezig
   ☐ HAVO                               Afgestudeerd/Mee bezig
   ☐ HBO                                Afgestudeerd/Mee bezig
   ☐ VWO                                 Afgestudeerd/Mee bezig
   ☐ Universiteit                        Afgestudeerd/Mee bezig
   ☐ Anders, namelijk........................ Afgestudeerd/Mee bezig

6. Hoe vaak bent u online?
   Uren per dag:........
   Dagen per week:........

7. Hoe goed kunt u met een computer overweg?

   1 ☐ Slecht
   2 ☐ Matig
   3 ☐ Gemiddeld
   4 ☐ Boven gemiddeld
   5 ☐ Goed

8. Bent u bekend met de term phishing?
   ☐ Ja  ☐ Nee

9. Bent u zich bewust van de hoeveelheid persoonlijke informatie die u op internet deelt en vervolgens publiekelijk toegankelijk is?
   ☐ Ja  ☐ Nee

10. Gebruikt u Facebook? Zo ja, wat zijn over het algemeen uw privacy-instellingen?
   ☐ Zichtbaar voor iedereen
   ☐ Zichtbaar voor vrienden
   ☐ Zichtbaar voor vrienden van vrienden
   ☐ Zichtbaar voor mezelf
   ☐ Weet ik niet
11. Bent u wel eens opgelicht via internet (bijvoorbeeld door middel van phishing)?

☐ Ja
☐ Nee


☐ Ja
☐ Nee


☐ Film/muziek
☐ Kleding
☐ Boek
☐ Meubilair
☐ Overig

14. Bij welke webwinkel heeft u dit product aangeschaft?

☐ bol.com
☐ Zalando
☐ Apple store
☐ wehkamp.nl
☐ Anders, namelijk ….

15. Gebruikt u over het algemeen iDeal (online bankieren) voor overschrijvingen, zoals aankopen op internet?

☐ Ja
☐ Nee

16. Sinds 1 Augustus is het volgens Europese wetgeving verplicht een IBAN nummer te gebruiken bij betalingen. In hoeverre weet u uw IBAN nummer al uit uw hoofd? Om privacy-redenen vragen wij u niet om alle getallen in te vullen. De getallen/letters die u niet hoeft in te vullen zijn aangegeven met X.

□□ XX □□□□ XXXXXX □□□□

Indien u een of meerdere lege vakjes heeft, ga dan verder. Ga anders naar de laatste vraag.

17. Weet u uw controlegetal uit uw hoofd? Dit is het derde en vierde getal (de twee getallen onder de eerste twee X’jes).

☐ Ja
☐ Nee

Er is een gemakkelijke methode waarmee u uw IBAN nummer kunt bepalen. Deze is hieronder weergegeven.

Als u uw controlegetal eenmaal heeft onthouden, is de rest heel gemakkelijk. Uw bankcode kunt u vinden op het bijgevoegde formulier. De laatste 10 getallen bestaan uit uw oude rekeningnummer. Wanneer uw oude rekeningnummer minder dan 10 getallen is, voegt u nullen ervoor om tot 10 getallen uit te komen.

18. Denkt u dat u uw IBAN nummer vanaf nu kunt onthouden?

☐ Ja
☐ Nee

Dat was het. Bedankt voor uw tijd!
Meld u a.u.b. bij de afnemer van de enquête voor een korte debriefing.