How Vulnerable Are We To Scams?

Markus Jakobsson
ZapFraud Inc.

Ting-Fang Yen
DataVisor Inc.

Abstract

The number of Internet scams has increased in recent years. According to a survey by the Federal Trade Commission, more than one out of every ten adult Americans fall victim to scams every year, where a third of these scams originated on the Internet. However, it is well understood that surveys of victimization and losses severely underestimate the problem, since victims are unwilling to come forward due to embarrassment or resignation. This paper attempts to gain a better understanding of the problem by directly quantifying the extent to which users are vulnerable to scams.

We design and carry out experiments to estimate the fraction of scam messages that bypass commercial spam filters (i.e., messages that land in the user's inbox); and to assess the probability that a delivered message will be considered harmless by its recipient. The latter experiment provides evidence that recent scams – many of which are targeted – are substantially more credible to typical users than "traditional" scam.

1 Introduction

In April of 2015, alleged Russian hackers gained access to White House computers using social engineering methods [5]. This, of course, is not the first breach that uses social engineering, and while government and enterprise cases are poorly documented, the consumer-facing problems are better understood. Every year, more than one out of every ten adult Americans report falling victim to scams, according to the latest FTC Consumer Fraud Survey [13], and a third of the scams originate on the Internet. According to the FBI [7], the average reported loss exceeds $2,000.

In terms of consumer losses, the reported statistics are alarming, but the problem is likely to be even worse. Chances are that most scams are never reported. Victims are commonly embarrassed ("How could I fall for that?"), resigned ("There is no way I will get this money back") or depressed [3, 8]. Since insurance does not cover scam losses, there is no need to file a police report, and law enforcement is largely incapable of addressing Internet crimes.

The scam problem is not going away. In part, this is due to the high profitability and low risk of this crime [4], but it is also likely due to an increased

degree of targeting. When an attacker uses information about a target including name, address or affiliation to create a customized scam message for the victim, this helps make the message credible. Whereas untargeted phishing attacks have a yield on the order of a percent, it has been shown that targeted phishing attacks can have yields above 75% [10] – provided the phishing emails get delivered to their intended recipients.

This paper aims at quantifying the vulnerability associated with social engineering attacks, whether targeted or not. There are two parts to this effort. The first part estimates the likelihood that commercial spam filters let through scam emails. The second part assesses the probability that a delivered email would be considered harmless to its recipient. By combining these two measures, we obtain a rough estimate of the typical vulnerability to scam.

We show that scam messages were blocked with a probability between 10% and 70% for Gmail, Hotmail and Yahoo. The emails[1] were sent in quantities of no more than a few hundred emails per sender, and using newly created sender accounts. We also identify *human* vulnerabilities to a collection of eight commonly occurring scam messages, and find yields ranging between 7% and 63% per message. Only 12% of our subjects were found not to be at risk when faced with a sequence of seven scam messages.

Whereas one must be cautious not to jump to conclusions, combining these two results does indicate that scammers that are skilled at targeting their attacks could potentially expect a yield on the order of 50% *per scam message*, provided the recipients decide to act on their (lack of) instincts.

Outline: After reviewing the related work in section 2, we describe our experiment on the scam block rate at three popular web email service providers in Section 3, followed by a second experiment on the credibility of scam messages in Section 4.

Two examples of emails used in the scam block experiment are shown in Appendix A; for a copy of all the emails, please contact the authors. The full set of emails used in the credibility of scam messages are shown in Appendix B.

2 Related Work

One common type of scam is the *Nigerian* scam. This is a type of social-engineering attack that is typically geared towards convincing the victim – using a credible reason – to send money to the scammer. This type of scam – also referred to as a 419 scam[2] – is currently becoming increasingly targeted, which makes it believable to a larger group of potential victims and increases its yield.

In a 2012 publication, Herley [9] analyzed Nigerian scams, posing and answering the question "Why do scammers volunteer the information that they are in Nigeria, when this is likely to tip recipients off that it is a scam?" Herley

[1] The scam messages were not created for the purpose of this study, but were randomly selected from a repository of several years old scam messages.

[2] The naming comes from the article in the Nigerian criminal code dealing with fraud.

argued that this strategy helps the scammers filter out all but the most gullible victims, thereby increasing the return on investment on those that do respond.

Whereas the routine of deterring all but the most gullible is likely to still be a effective strategy for some types of fraud – or scammers would have stopped this practice – we believe that this type of scam is being replaced by more realistic scams, in which the victims are *not* intentionally tipped off. Scammers increase the credibility of scams using an increased degree of targeting. Park et al. [11] studied scams that use a greater degree of targeting than those studied by Herley. Park et al. identified common scammer techniques and tracked scammers by setting up "magnetic honeypots" on Craigslist. These are advertisements that are written to repel legitimate users but appeal to scammers – in fact, working very much along the lines of how Herley suggests that scammers repel non-gullible users.

One form of online scam that is commonly targeted is the *romance scam*. This is a variant of the 419 scams in which users looking for love enter an online relationship with a scammer – posing as a single man or woman. After a few weeks of interaction (and deepening interest on behalf of the victim), the scammer requests a loan – whether to solve a temporary problem, or to come to visit the victim. Rege [12] studied romance scams and quantified their impact by assessing the commonality of media reports relating to this scam.

None of the studies described above measured directly the reaction of would-be victims, but focused on the actions of the scammers or media coverage, since these presumably reflect the vulnerabilities of the victims. To our knowledge, there are no prior publications aimed at *directly* quantifying the vulnerabilities to scam of typical users. While there is an array of studies quantifying the vulnerability to phishing attacks (e.g., [10]), most of these are several years old, and do not consider scams in which the victim ends up voluntarily transferring money to the scammers – which is an increasingly common type of scam. One contribution of this paper is to directly assess vulnerabilities, using an improved type of security IQ tests.

Whereas security IQ tests (e.g., [6]) are common as a consumer awareness technique, and may have some benefits in terms of user education, it has been found [1] that they are not effective in terms of assessing real risks. There are several reasons for this. One is that the questions in typical security IQ tests are commonly not very subtle. That, in combination with the fact that the user knows that he or she is looking for signs of danger, helps people identify risks they may not have spotted in a real-life setting. At the same time, many tests have questions that are difficult to answer without knowledge of the supposed service providers, and rely on knowledge of the actual domains used by commonly spoofed companies. This means that the results of traditional security IQ tests do not represent the actual abilities of the test takers, and so, have a very limited utility within security research. We address these shortcomings by introducing – and successfully using – a more subtle approach to testing.

Nigerian scams is at the center of this paper, but we are also addressing phishing attacks. Our reason is that the line between these two types of scams is increasingly becoming blurred, making it meaningful to consider the two as

simple versions of one and the same principle.

More remotely related to this work is the study of phishing campaigns in enterprise data breaches. The latest Verizon Data Breach Investigations Report [14] found that 23% of the recipients open phishing messages, and 11% click on attachments. In contrast to common scam messages, the goal of these phishing emails is to gain an initial "foothold" into the target organization, rather than money or personal information. As a result, they are much more personalized (focusing on specific companies or users).

3 How Much Scam is Blocked?

Commercial spam filters use a variety of ways to identify unwanted messages, such as based on the message body, the sender reputation, URLs in the message, attachments, etc. In this experiment, we attempt to estimate the extent to which common scam messages manage to bypass the built-in spam filters used by three popular web email service providers: Yahoo, Hotmail, and Gmail. These service providers, dominating in the webmail market, are likely to have deployed the state-of-the-art spam filtering technology, and hence allow us to measure the "best case" scenario in scam blocking.

3.1 Methodology

We obtained a corpus of over 400 scam messages, randomly selected from a repository of several years old scam messages. We intentionally selected old scam messages to make sure that we do not unintentionally measure the speed with which email service providers are able to respond to new threats.

A large fraction of the scam messages correspond to 419 scams, a form of advanced-fee fraud in which the fraudster attempts to extract money from the victim by promising a large sum of money in the future. There are many variants of this type of scam, involving the offer of investments, sale of products or services, lottery winnings, gifts, etc.

We measured the extent to which commercial spam filters block this type of scam messages. 12 email accounts were created for the purpose of this experiment: three to act as senders of scam messages, and nine as message recipients (three from each service provider). Each scam message was sent by one of the senders to all nine recipients. We examine the number of messages that made it to the recipients' inbox, i.e., bypassing the spam filter, to quantify the scam block rate for each service provider.

Our goal was to measure how well email service providers detect scam messages based on their text content – as opposed to the presence of known malicious URLs, for example. Similarly, to avoid measuring the impact of sender reputation, our recipients were made to automatically respond to all scam messages arriving in their inbox, with a simple message: "Thanks!" However, for two of the accounts we used, this auto-response feature failed, revealing to us the importance of the response to the scam filtering rate. This, in turn, sheds light

over why scammers commonly start a conversation by asking for an email – e.g., "Did you get my message?" and illustrates the relation between credible emails and low scam-block rates.

Initially, we had one sender account per service provider, similar to the setup for recipient accounts. However, the Hotmail and Gmail sender accounts were quickly disabled by captchas or phone number verification (possibly due to their scam messages being detected). As a result, we changed all three senders to Yahoo accounts, which did not have this security feature. While we have yet to investigate this further, it is possible that this – and the ease with which Yahoo accounts can be scripted (compared to Hotmail and Gmail) – is part of the reason for the high prevalence of scammers using Yahoo email accounts [2].

Over the course of two days, each sender account sampled the scam corpus and sent between 110 and 166 messages to the nine recipient accounts. The variability in the number of messages sent is due to a random wait time — from 1.5 to 7 minutes — between the sending of consecutive messages.

3.2 Analysis

Of the 427 scam messages sent by the three senders, all of them made it to at least one recipient's inbox by bypassing the spam filter. Table 1 shows the statistics for each service provider. The numbers correspond to a count of *unique* messages, with each row being the union of unique messages sent to, or received by, recipients associated with each service provider. The relatively low message count for the Yahoo recipients is due to those accounts being deleted by the service provider a few hours into the experiment. We are unsure of the reason: The accounts were identical in terms of how and how often they were accessed; and after all, they were *recipients* of scam messages – not the *senders*.

Service Provider	Messages Sent	Messages Received	Percentage Blocked
Hotmail	427	145	66%
Gmail	427	385	10%
Yahoo	331	99	70%

Table 1: Block rate statistics for the three web email service providers. The number of messages sent or received is computed as the union of unique messages sent to, or received by, recipients associated with each service provider.

Among the three service providers, Gmail has the lowest block rate (at 10%). Further investigation showed that the Gmail recipients have strikingly different results, with one recipient receiving 90% of the sent scam messages in its inbox, and the other two only receiving 2% (i.e., 9 messages). It turned out that our auto-response script was not working properly for all the Gmail recipient accounts, and that only one of the Gmail recipients was successfully responding to messages [3] — the one that received 90% of the scam messages. This shows that Gmail's spam filtering system weighs the sender reputation quite heavily.

[3]We were accessing those accounts from different geographic locations, and Gmail's phone verification system blocked access from our scripts.

This is very interesting, since it illustrates an important concept: *There is a direct relationship between the credibility of an email (as assessed in section 4) and the efficacy of the spam filter.* In particular, credible messages will result in low block rates. This acts to further amplify the effect of credible messages, since these will not be blocked to the same extent. It also indicates that the use of targeting of scams – a generalization of spear phishing – will be fueled not only by the increased credibility of these messages, but also by the effects this has on automated filtering.

We also noticed that a scam message is not always classified consistently across the three service providers. For example, of the messages that landed in the inbox of Hotmail or Yahoo recipients, only 12% (27 messages) showed up in both. Even among recipients belonging to the same service provider, the overlap is on average only 73% for Hotmail and 83% for Yahoo recipients. This suggests that there are many factors outside of the message content that plays into the filtering decision, which is likely probabilistic.

There are many variables affecting spam filtering that we do not measure, such as the sender age, location, the rate at which messages are sent, the method by which the messages are sent, the effect of spam poisoning, the content of the responses from the recipients, etc. That said, our small experiment is an attempt to quantify how well commercial spam filtering works against common scam messages in a relatively naive setting – we do not try to hide the fact that the accounts are controlled by scripts, though making that work did present some challenges.

4 How Credible are Scam Messages?

We now describe our experiment to quantify the probability that typical recipients fall for scam emails. In this experiment, we asked human subjects to review a collection of scam emails, assessing their credibility by performing a task in which the subjects were asked to indicate the nature of the primary risk associated with each message.

It should be noted that there is not a one-to-one correspondence between the messages in the first experiment and this one. The reason is that the first experiment assesses spam filters based on *hundreds* of messages, which is far too many to rate in the second experiment. Also, since scammers and spam filters are constantly reacting to each others' actions, it is not of importance how well any *particular* type of scam instance is blocked – it is the *typical* blocking rates that matter. Similarly, the exact human vulnerabilities found in the second experiment are of lesser importance than the general range, and the likely reasons underlying high yields. This is since sudden media attention to one type of scam may (at least temporarily) make people less likely to fall victim to this scam type – but that type of attention will not inoculate people against scam in a more general sense.

4.1 Methodology

Traditional security IQ tests pose test takers with relatively obvious cases of fraud (or its absence), and ask the test takers to classify email messages and webpages as either safe or not. This leads to test results that do not track the actual abilities [1].

We introduce a new approach to security IQ tests in which test takers are posed with a collection of troublesome situations – but for which the risks are different – and in which the test taker is asked to identify what the *primary* risk is, from a collection of possibilities of varying degrees of correctness. Therefore, a failure to understand the risk causes the test taker to select an option that arguably *is* a risk – but certainly not the most prominent risk.

The answers to the survey questions can be grouped into three types: *correct*, *reasonable* and *naive* answers. The first and last type are self-explanatory. By reasonable, we mean neither correct nor naive. For example, consider an email that contains a clickable link. It is reasonable that the risk associated with the email is for the recipient's computer to get a computer virus – however, a person with expert knowledge may know that the type of email may primarily be used to steal credentials.

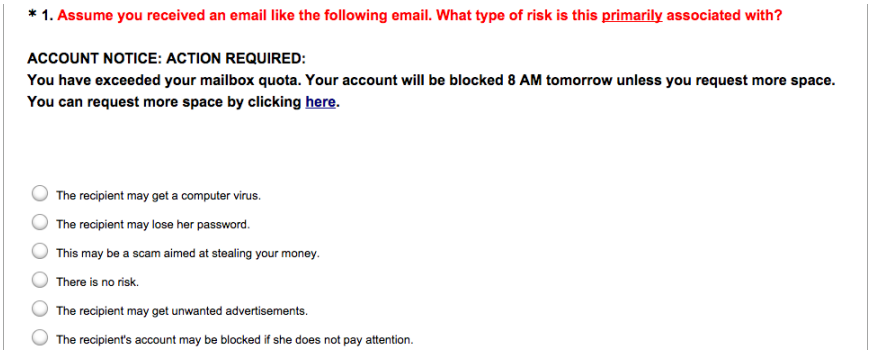


Figure 1: One of the eight questions the subjects were asked to review. The correct answer is "The recipient may lose her password", as typically, a victim of this type of scam is asked to sign in to what he is made to believe is his or her email service provider. As detailed in the next section, only 10% of the test takers selected this answer. However, since there is a clickable link, the answer "The recipient may get a computer virus" (40%) is also a reasonable response. A user who responds this way does not have a deep knowledge of scams, but still identifies this risk as a possibility. Similarly, the response "This may be a scam aimed at stealing your money" (19%) is reasonable, since the end goal of the scammer may be to do so. On the other hand, the responses "There is no risk" (3%), "The recipient may get unwanted advertisements" (5%) and "The recipient's account may be blocked if she does not pay attention" (23%) are naive.

We performed an experiment in which 107 subjects were asked to review

eight email messages (one example shown in Figure 1, and the remainder shown in Appendix B), and for each one indicate what the primary risk is – if any. In addition, the subjects were asked to respond to a multiple-choice demographic question.

The questions in our survey correspond to phishing emails; Nigerian scams; and (to decrease test-taker bias) one message that is simply spam, and of unlikely direct risk to a recipient. The phishing emails and the Nigerian scam messages are selected to reflect the recent increase in sophistication among scammers: some of the messages are "traditional scams" whereas others are of the type that have become increasingly common recently, in which attackers increase their yield by the targeting of recipients.

4.2 Analysis

Figure 2 shows the percentage of *naive* (in black) and *reasonable* (in grey) answers for each of the eight email messages that subjects were asked to review. Question 5 is left out since that does not correspond to a scam, but was simply added as a red herring.

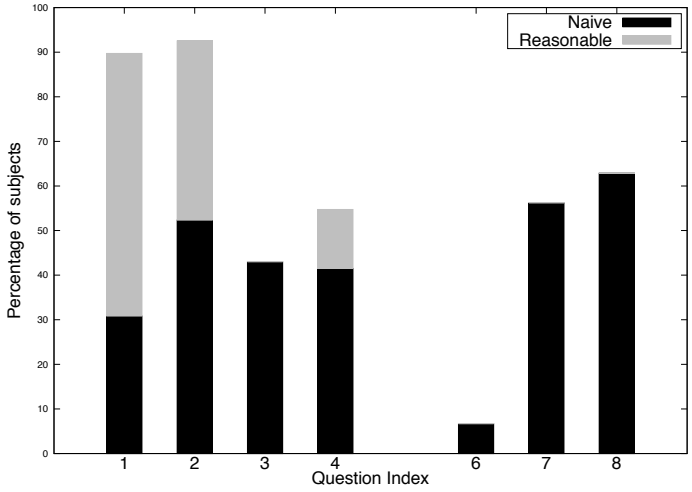


Figure 2: The figure shows the percentage of *naive* answers (in black) and of *reasonable* answers (in grey) for each email message. The remaining percentages correspond to correct responses. The full questions are shown in Appendix B.

The subjects selected the correct answers between 7% and 93% of the time, and had *reasonable* answers up to 59% of the time. The remaining portion of the answers – between 7% and 63%, with an average of 42% – corresponds to

naive answers. These, in turn, correspond to users at high risk. It is our belief that users with reasonable answers will be partially protected, since they would be cautious – although for the wrong reasons.

The two questions with the highest rates of correct answers (question 3 and 6) are "traditional" Nigerian scams. Questions 2, 4, 7 and 8 correspond to targeted attacks, e.g., sent in response to Craigslist ads or posts in job-seeking forums. These targeted emails also have the highest percentage of naive answers, ranging from 42% to 63%, confirming the power of targeted scams.

We also determined what percentage of users *never* selected a naive response to the any of the seven questions relating to scams. This was only 12% of the subjects. In other words, 88% of the subjects selected *at least* one naive answer, indicating that they are at risk. This is an eye-opening number to us.

In addition to being asked to identify the primary risk associated with the example emails, the subjects were asked to respond whether or not they: have free anti-virus software; have paid anti-virus software; are older than 40; have a total household income below $100k; have a total household income above $100k[4]; have purchased anything online in the last year; have ever used eBay; have ever used Craigslist; have suffered Internet scams in the last year; have suffered Internet scams in the past; whether they know how to identify scams; and whether they sometimes go through their spam folder to find good messages. [5] The distribution of responses is shown in Table 2.

Question	% of Subjects	# of Subjects
I have free Anti-Virus software on my computer	67%	70
I have paid Anti-Virus software on my computer	22%	23
I am at least 40 years old	33%	35
My household income is less than $100,000 per year	84%	88
My household income is more than $100,000 per year	7%	7
I make at least one online purchase every month	69%	72
I have used eBay	80%	84
I have used Craigslist	74%	78
I have lost money to an Internet scam in the past year	2%	2
I have lost money to an Internet scam	8%	8
I know how to spot online scams	77%	81
Sometimes good emails get placed in my spam folder by mistake, and I have to search for them there	54%	57
I would be interested in a scam IQ test (looking much like this survey) that would give me a score and tell me when I am wrong	23%	24

Table 2: Demographic information of the 107 test-takers.

We find no significant differences in terms of test-taker success based on *any* of these selections – including the one whether the test taker claims to know how to identify scams.

[4]This redundancy was added to identify users who did not pay attention to the survey, and who clicked indiscriminantly. No such subjects were identified.

[5]This constitutes a known risk to very gullible users, since this commonly implies that they do not receive any protection at all from their spam filters.

5 Conclusion

In this paper, we quantify the extent to which users are vulnerable to scams through two experiments: One to measure the scam block rate of commercial spam filters, and the other to measure the credibility of common scam messages to regular users. Our study presents several interesting results:

- Commercial spam filters used by Yahoo and Hotmail *fail* to block 30% and 34% of the scam messages, when sent in quantities of no more than a few hundred emails per sender, using newly created sender accounts. Gmail, on the other hand, can allow 90% of the sent scam messages through to the user's inbox, for the case of the one Gmail recipient whose auto-response features was working successfully.
- We find that the reputation of email senders plays an important role in the spam filtering system. This points to a direct relationship between the credibility of an email and the (in)efficacy of the spam filter, since credible messages will not be blocked to the same extent.
- We introduce a new approach to security IQ tests where the test takers are asked to identify the primary risk from a collection of possibilities of varying degrees of correctness. Our results show that the average rate of *naive* answers is 42%, and this is especially high (over 60%) for targeted messages. Only 12% of the users were free from naive answers, suggesting that 88% are potentially at risk.

Acknowledgments

Many thanks to Bill Leddy for inspiring discussions and helpful feedback, and to Arthur Jakobsson for help processing experimental data.

References

[1] V. Anandpara, A. Dingman, M. Jakobsson, D. Liu, and H. Roinestad. Phishing IQ tests measure fear, not ability. In *11th International Conference on Financial Cryptography and 1st International Conference on Usable Security*, 2007.

[2] M. Bates. If your Yahoo email account is sending spam. `http://www.batesline.com/archives/2013/01/yahoo-email-spam.html`, 2013.

[3] BBC News. Suicide of internet scam victim. `http://news.bbc.co.uk/2/hi/uk_news/england/cambridgeshire/3444307.stm`, 2004.

[4] K. Brulliard. Worldwide Slump Makes Nigeria's Online Scammers Work That Much Harder. `http://www.washingtonpost.com/wp-dyn/content/article/2009/08/06/AR2009080603764.html`, 2009.

[5] CNN News. How the U.S. thinks Russians hacked the White House. http://www.cnn.com/2015/04/07/politics/how-russians-hacked-the-wh/.

[6] DELL. SonicWALL Phishing IQ Test. http://www.sonicwall.com/furl/phishing/.

[7] FBI and IC3. 2013 Internet Crime Report. http://www.ic3.gov/media/annualreport/2013_IC3Report.pdf, 2013.

[8] L. Fulford. Autistic schoolboy hanged himself after falling for scam 'police' email saying he had looked at indecent websites. http://www.mirror.co.uk/news/uk-news/autistic-schoolboy-hanged-himself-after-5025351, 2015.

[9] C. Herley. Why do Nigerian Scammers Say They are from Nigeria? In *11th Annual Workshop on the Economics of Information Security (WEIS)*, 2012.

[10] T. N. Jagatic, N. A. Johnson, M. Jakobsson, and F. Menczer. Social Phishing. *Communications of the ACM*, 50(10), 2007.

[11] Y. Park, J. Jones, D. McCoy, E. Shi, and M. Jakobsson. Scambaiter: Understanding Targeted Nigerian Scams on Craigslist. In *Network and Distributed System Security Symposium (NDSS)*, 2014.

[12] A. Rege. What's Love Got to Do with It? Exploring Online Dating Scams and Identity Fraud. *International Journal of Cyber Criminology*, (2), 2009.

[13] Staff Report of the Bureau of Economics, Federal Trade Commission. Consumer Fraud in the United States, 2011: The Third FTC Survey. https://www.ftc.gov/sites/default/files/documents/reports/consumer-fraud-united-states-2011-third-ftc-survey/130419fraudsurvey_0.pdf, 2013.

[14] Verizon RISK Team and others. Verizon 2015 Data Breach Investigation Report. Technical report, 2015.

A Emails in Spam Filter Experiment

Below are two examples of the scam emails used in our first experiment (Section 3) to measure the scam block rate for commercial spam filters. Please contact the authors for a copy of the full collection.

Message 1:

DEPT OF ENERGY AND MINERAL RESOURCES,
PRETORIA, SOUTH AFRICA.
ATTN:

It is my great pleasure to write you this letter on behalf of my colleagues. Your information was given to me by a member of the South African Export Promotion Council (SAEPC) who was with theGovernment delegation on a trip to your country for a United Nations bilateral conference talk on sustainable development to encourage foreign investors. I have decided to seek a confidential co-operation with you in the execution of a deal hereunder for the benefit of all parties and hope you will keep it confidential because of the nature of the business. Within the Department of Minerals and Energy where I work as an assistant Director of Audit, with the co-operation of two other top officials, we have in our possession an overdue contractor payment in US Dollars funds. The said funds represent certain percentage of the contract value executed on behalf of my Department by a foreign contracting firm, (Pearls Ltd) which we the officials over-invoiced by the amount of US$15,200,000 {Fifteen Million Two Hundred Thousand US Dollars). Since the present elected Government is determined to pay foreign contractors all debts owed, so as to maintain good relations with foreign governments and non-governmental agencies, we included our bills fo approvals with the Department of Finance and the Reserve Bank of South Africa (RBSA). We are 100+% sure of funds approvals to anyone or company we (The Audit Committee) recommend as part of the sub-contractors who did jobs for the Department. We are seeking your assistance to front as the sub-contractor of the unclaimed funds, since we are not allowed to operate foreign accounts. Details and change of beneficiary information upon application for claim to reflect payment and approvals will be secured on behalf ofYou=2Fyour Company. My colleagues and I are prepared to give you US$2.28m while we take US$7.4m and the balance of US$5.52m for taxes and miscellaneous expenses incurred. This business is completely safe and secure,provided you treat it with utmost confidentiality. It does not matter whether You=2Fyour Company does contract projects, as a transfer of rights will be secured in favor of You=2Fyour Company through the Federal high Court of South Africa before we can proceed. I have reposed my confidence in you and hope that you will not disappoint us. Kindly notify me immediately for further details upon your acceptance of this proposal.

Yours Faithfully,
Alex Brown.

Message 2:

Mrs. Mariam. Abacha (Dr.)
345 ABACHA STREET
FAX/TEL: 234-17593830
Dear Friend,
URGENT AND CONFIDENTIAL

I am Mrs. Mariam. Abacha, the widow of Sani Abacha the Late Nigerian Head of State. I am presently in distress and under house arrest while my son Mohammed is undergoing trial in

Lagos and Abuja though he has just been recently granted bail under the condition that my family refunds to the Federal Government some amount of money. The government has frozen all the family account and auctioned all our properties. Refer to this website about my husband's loot and you will understand what I mean http://news.bbc.co.uk/1/hi/world/africa/1935646.stm managed to ship through an undercover courier company,the sum of US24, 000,000.00. kept by my late husband. Themoney was disguised to beat the Nigerian Security and it is currently deposited in a security company which I will disclose the name and contacts receive the money and pay it into your account for the family safely. I am offering you 30% for assisting me secure this money Contact me immediately with my email address so that I can forward to you all necessary details. Endeavour to send your phone and fax numbers for easy Communications. This project is not risky. my private email address, `removed`

Sincerely Yours,

Mrs. Mariam. Abacha (Dr.)

B Emails in Scam Credibility Experiment

We show the questions and answer options given to the subjects in the scam credibility experiment described in Section 4, with response percentages in parentheses. Question 1 is shown in Figure 1.

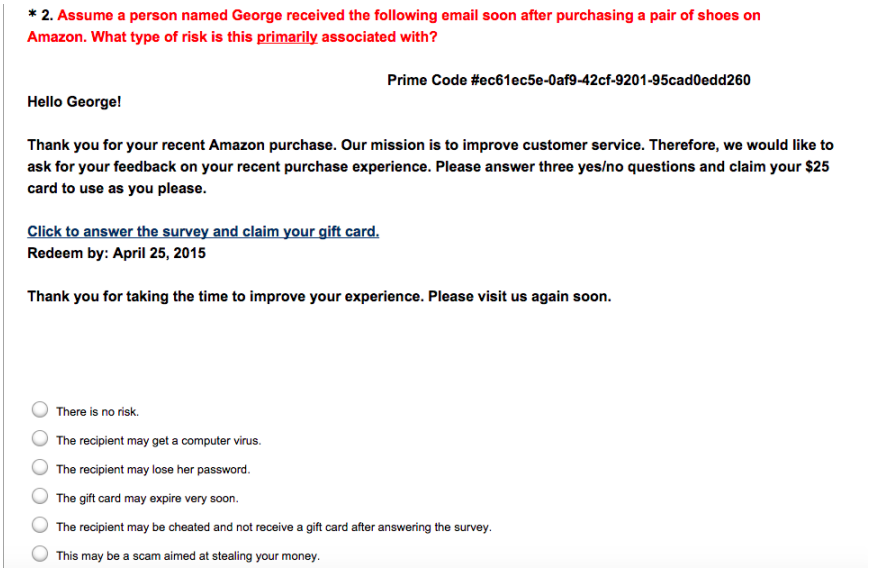


Figure 3: Question 2 shows a new variant of a scam that has existed for about ten years, in which the recipient is asked to respond to a survey in return for a financial award. To receive the award, the user has to enter personal information, including her password, on a site she believes is associated with the surveying organization. In the past, this attack has not used any targeting, but recently, it has been seen to target (and correctly name) Amazon Prime members – including accounts used uniquely for Amazon purchases. This suggests a potential breach or misbehavior of at least one Amazon seller. The "prime code" is likely to be spam poison, i.e., a technique used to circumvent spam filters. Correct answer: The recipient may lose her password (8%). Reasonable answers: The recipient may get a computer virus (15%); This may be a scam aimed at stealing your money (24%). Naive answers: There is no risk (13%); The gift card may expire very soon (7%); The recipient may be cheated and not receive a gift card after answering the survey (33%).

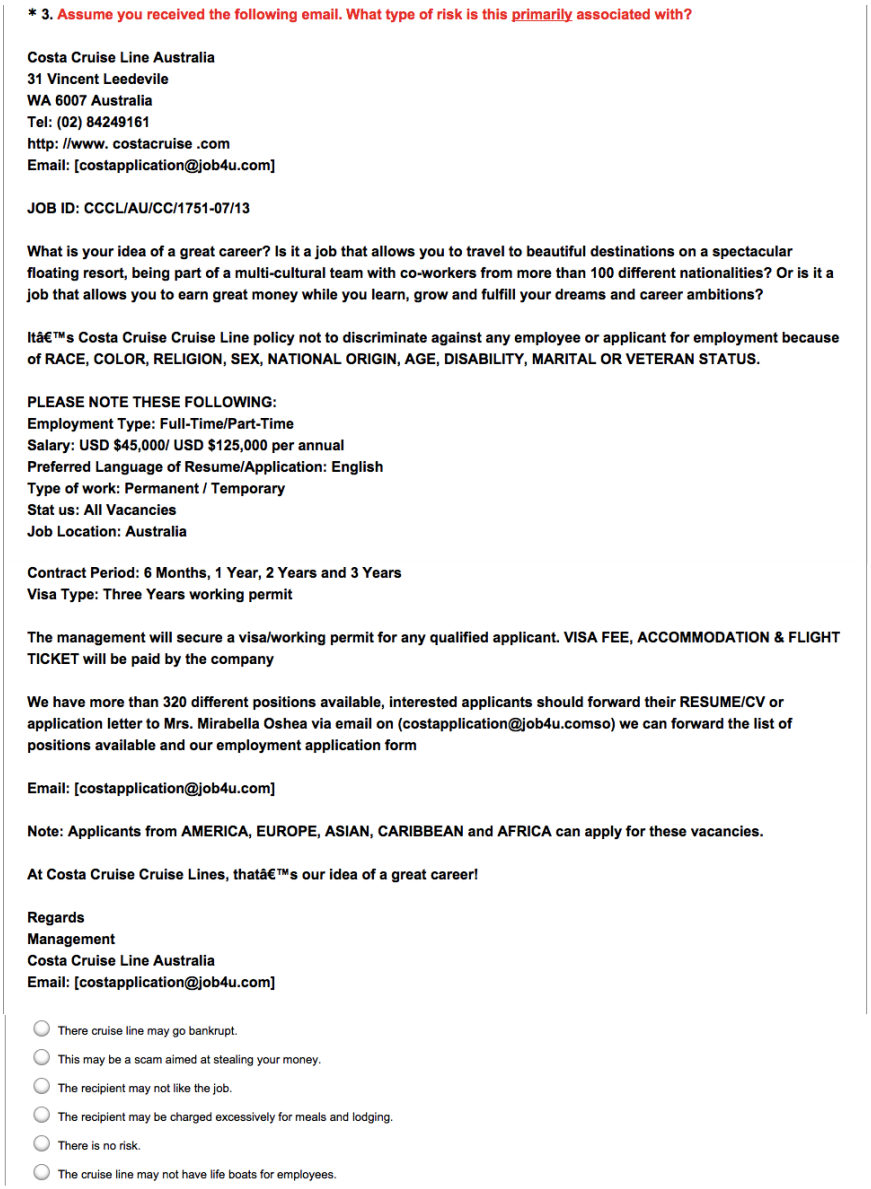


Figure 4: Question 3 shows a fairly traditional Nigerian scam. Sometimes it is emailed out en masse, with no targeting, and sometimes, it is sent to users of job seeking forums, where it is likely to have a much higher yield. Please notice the spam poison (the "job id"); this is a semi-unique value to circumvent spam filters; this is indicative of the scam having been sent at a large volume. Correct answer: This may be a scam aimed at stealing your money (57%). Reasonable answers: (none) Naive answers: There cruise line may go bankrupt (3%); The recipient may not like the job (12%); The recipient may be charged excessively for meals and lodging (12%); There is no risk (15%); The cruise line may not have life boats for employees (1%).

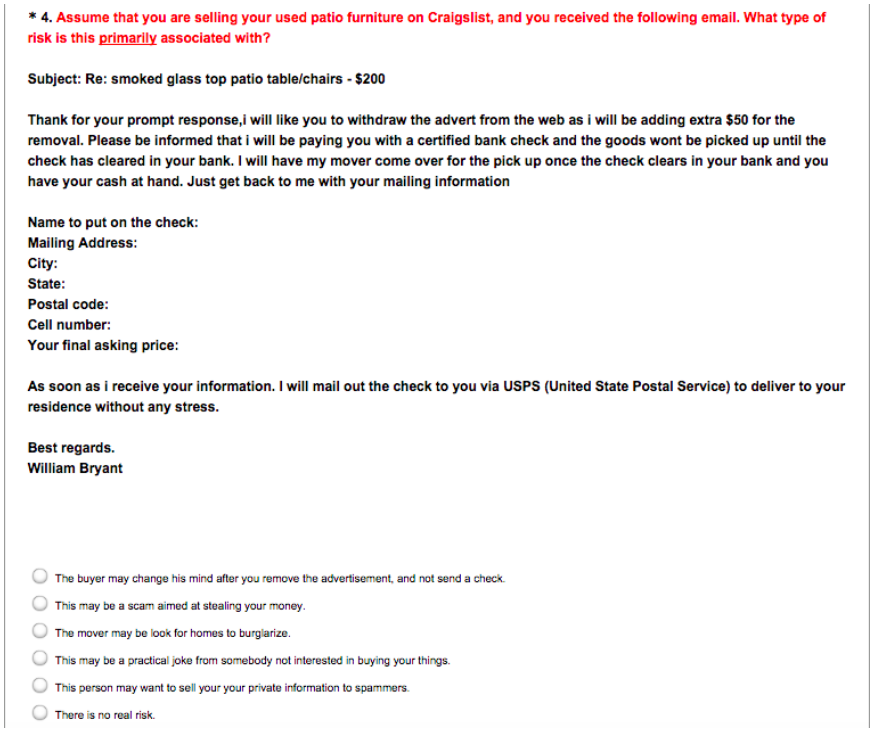


Figure 5: Question 4 shows a targeted scam that is sent in response to a user having put an advertisement on Craigslist; this is a very common type of scam [11]. Correct answer: This may be a scam aimed at stealing your money (47%). Reasonable answer: This person may want to sell your your private information to spammers (13%). Naive answers: The buyer may change his mind after you remove the advertisement, and not send a check (16%); The mover may be look for homes to burglarize (16%); This may be a practical joke from somebody not interested in buying your things (2%); There is no real risk (5%).

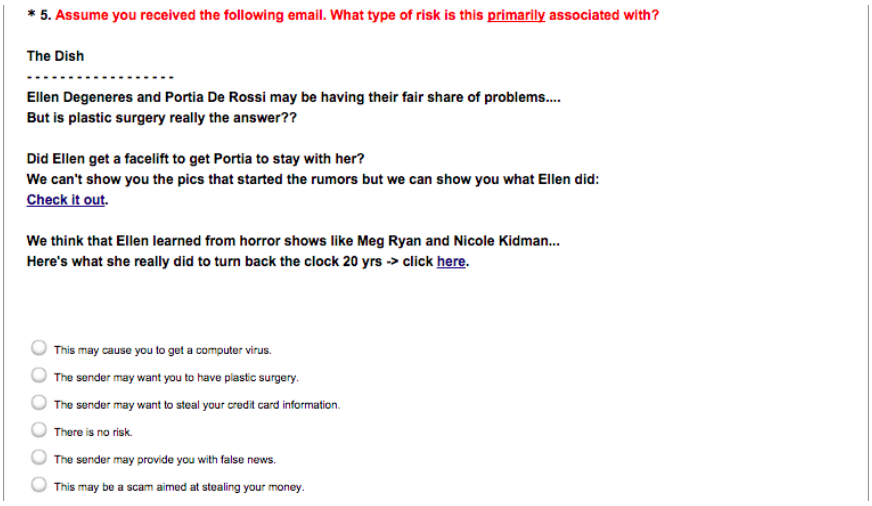
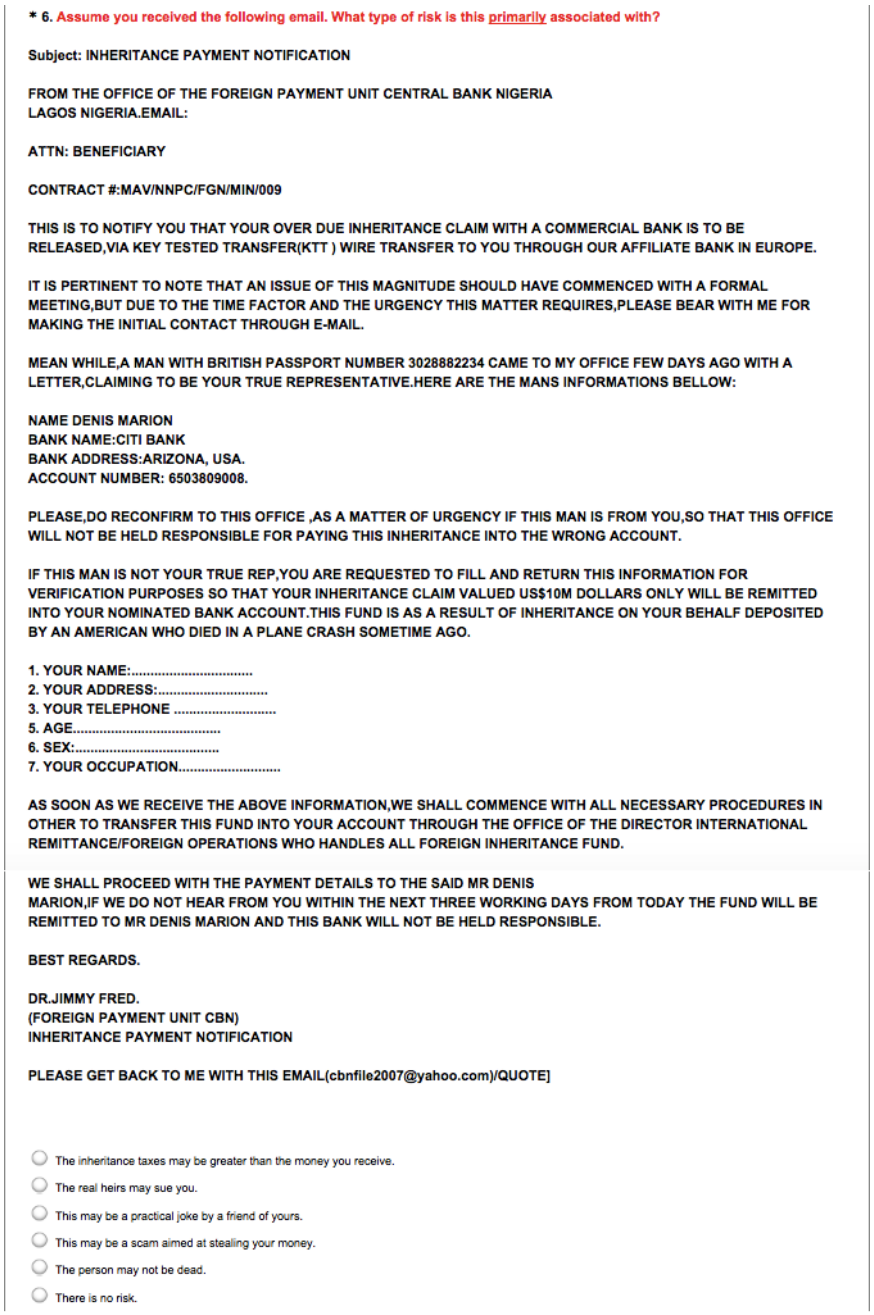


Figure 6: Question 5 was added as a red herring. It is spam rather than scam. The responses were ignored.



*** 6. Assume you received the following email. What type of risk is this primarily associated with?**

Subject: INHERITANCE PAYMENT NOTIFICATION

**FROM THE OFFICE OF THE FOREIGN PAYMENT UNIT CENTRAL BANK NIGERIA
LAGOS NIGERIA.EMAIL:**

ATTN: BENEFICIARY

CONTRACT #:MAV/NNPC/FGN/MIN/009

THIS IS TO NOTIFY YOU THAT YOUR OVER DUE INHERITANCE CLAIM WITH A COMMERCIAL BANK IS TO BE RELEASED,VIA KEY TESTED TRANSFER(KTT) WIRE TRANSFER TO YOU THROUGH OUR AFFILIATE BANK IN EUROPE.

IT IS PERTINENT TO NOTE THAT AN ISSUE OF THIS MAGNITUDE SHOULD HAVE COMMENCED WITH A FORMAL MEETING,BUT DUE TO THE TIME FACTOR AND THE URGENCY THIS MATTER REQUIRES,PLEASE BEAR WITH ME FOR MAKING THE INITIAL CONTACT THROUGH E-MAIL.

MEAN WHILE,A MAN WITH BRITISH PASSPORT NUMBER 3028882234 CAME TO MY OFFICE FEW DAYS AGO WITH A LETTER,CLAIMING TO BE YOUR TRUE REPRESENTATIVE.HERE ARE THE MANS INFORMATIONS BELLOW:

**NAME DENIS MARION
BANK NAME:CITI BANK
BANK ADDRESS:ARIZONA, USA.
ACCOUNT NUMBER: 6503809008.**

PLEASE,DO RECONFIRM TO THIS OFFICE ,AS A MATTER OF URGENCY IF THIS MAN IS FROM YOU,SO THAT THIS OFFICE WILL NOT BE HELD RESPONSIBLE FOR PAYING THIS INHERITANCE INTO THE WRONG ACCOUNT.

IF THIS MAN IS NOT YOUR TRUE REP,YOU ARE REQUESTED TO FILL AND RETURN THIS INFORMATION FOR VERIFICATION PURPOSES SO THAT YOUR INHERITANCE CLAIM VALUED US$10M DOLLARS ONLY WILL BE REMITTED INTO YOUR NOMINATED BANK ACCOUNT.THIS FUND IS AS A RESULT OF INHERITANCE ON YOUR BEHALF DEPOSITED BY AN AMERICAN WHO DIED IN A PLANE CRASH SOMETIME AGO.

**1. YOUR NAME:.....
2. YOUR ADDRESS:.....
3. YOUR TELEPHONE .....
5. AGE.....
6. SEX:.....
7. YOUR OCCUPATION.....**

AS SOON AS WE RECEIVE THE ABOVE INFORMATION,WE SHALL COMMENCE WITH ALL NECESSARY PROCEDURES IN OTHER TO TRANSFER THIS FUND INTO YOUR ACCOUNT THROUGH THE OFFICE OF THE DIRECTOR INTERNATIONAL REMITTANCE/FOREIGN OPERATIONS WHO HANDLES ALL FOREIGN INHERITANCE FUND.

**WE SHALL PROCEED WITH THE PAYMENT DETAILS TO THE SAID MR DENIS
MARION,IF WE DO NOT HEAR FROM YOU WITHIN THE NEXT THREE WORKING DAYS FROM TODAY THE FUND WILL BE REMITTED TO MR DENIS MARION AND THIS BANK WILL NOT BE HELD RESPONSIBLE.**

BEST REGARDS.

**DR.JIMMY FRED.
(FOREIGN PAYMENT UNIT CBN)
INHERITANCE PAYMENT NOTIFICATION**

PLEASE GET BACK TO ME WITH THIS EMAIL(cbnfile2007@yahoo.com)/QUOTE]

- The inheritance taxes may be greater than the money you receive.
- The real heirs may sue you.
- This may be a practical joke by a friend of yours.
- This may be a scam aimed at stealing your money.
- The person may not be dead.
- There is no risk.

Figure 7: Question 6 corresponds to a traditional Nigerian scam. While the passport number could have been used as scam posion, a quick search for "mean while,a man with british passport" reveals that it probably was not – all reported instances of the scam have the same number following the word "passport". Correct answer: This may be a scam aimed at stealing your money (94%). Naive answers: The inheritance taxes may be greater than the money you receive (2%); The real heirs may sue you (2%); This may be a practical joke by a friend of yours (1%); The person may not be dead (0%); There is no risk (1%).

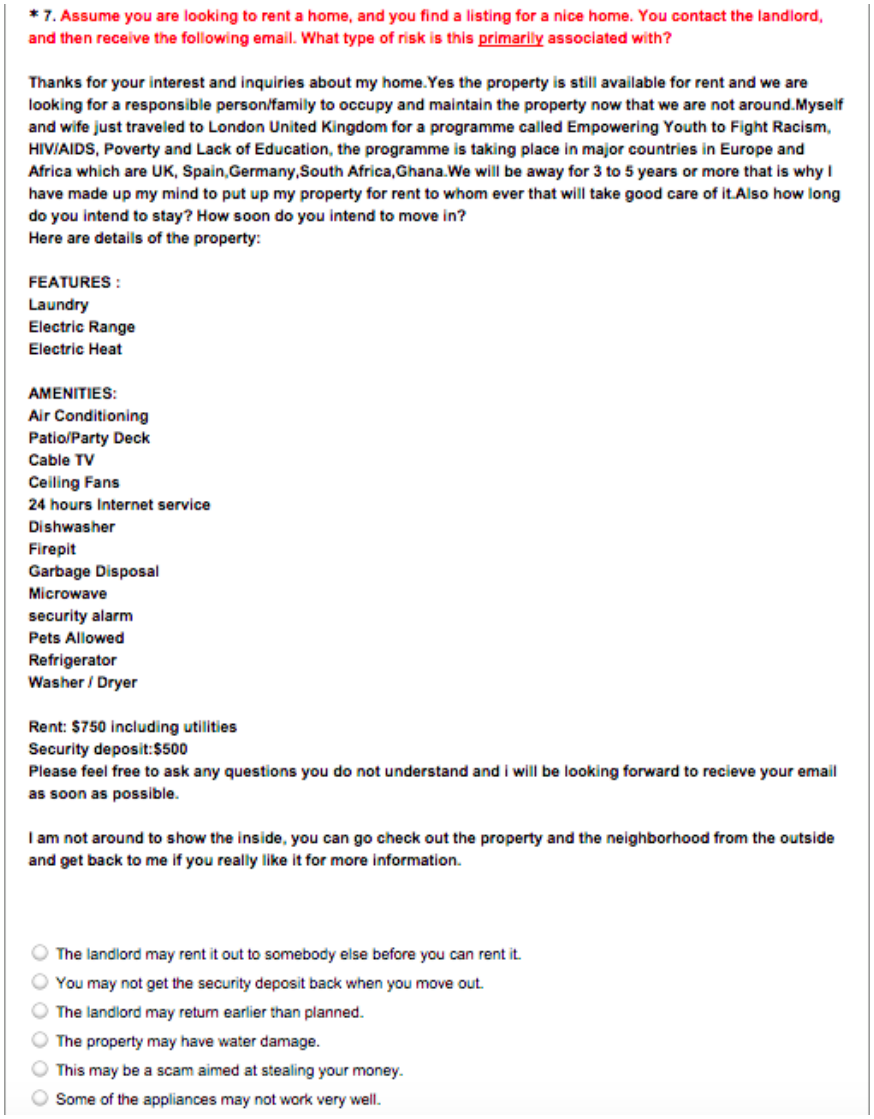


Figure 8: Question 7 corresponds to a targeted scam, commonly seen on housing forums. Correct answer: This may be a scam aimed at stealing your money (44%). Naive answers: The landlord may rent it out to somebody else before you can rent it (29%); You may not get the security deposit back when you move out (10%); The landlord may return earlier than planned (10%); The property may have water damage (5%); Some of the appliances may not work very well (2%).

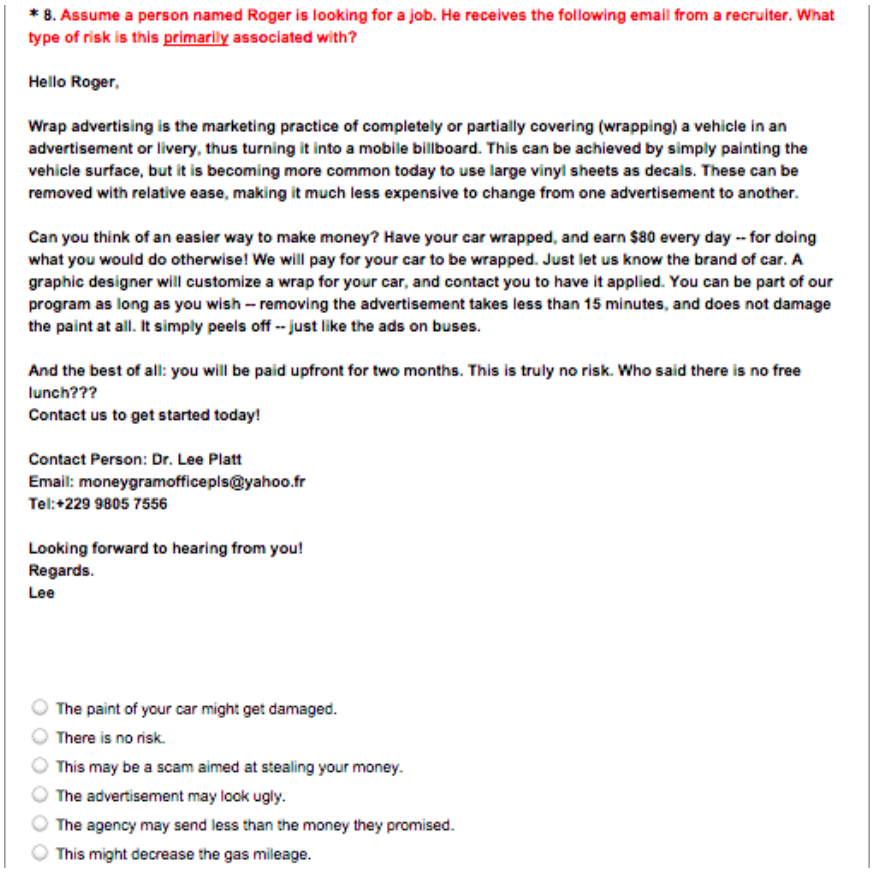


Figure 9: Question 8 corresponds to a targeted scam, commonly seen on job seeking forums. This particular instance was edited for length by the authors, to make it more suitable for the survey; however, this gist of the original message was closely followed. Correct answer: This may be a scam aimed at stealing your money (37%). Naive answers: The paint of your car might get damaged (26%); There is no risk (8%); The advertisement may look ugly (13%); The agency may send less than the money they promised (16%); This might decrease the gas mileage (0%).