F for Fake: Four Studies on How We Fall for Phish

Mark Blythe
School of Design,
Northumbria University,
Newcastle, United Kingdom
mark.blythe@northumbria.ac.uk

Helen Petrie, John A. Clark
Department of Computer Science
University of York,
York, United Kingdom
Helen.Petrie|jac@cs.york.ac.uk

ABSTRACT
This paper reports findings from a multi-method set of four studies that investigate why we continue to fall for phish. Current security advice suggests poor spelling and grammar in emails can be signs of phish. But a content analysis of a phishing archive indicates that many such emails contain no obvious spelling or grammar mistakes and often use convincing logos and letterheads. An online survey of 224 people finds that although phish are detected approximately 80% of the time, those with logos are significantly harder to detect. A qualitative interview study was undertaken to better understand the strategies used to identify phish. Blind users were selected because it was thought they may be more vulnerable to phishing attacks, however they demonstrated robust strategies for identifying phish based on careful reading of emails. Finally an analysis was undertaken of phish as a literary form. This identifies the main literary device employed as pastiche and draws on critical theory to consider why security based pastiche may be currently very persuasive.

Author Keywords
Phish detection, human factors, visually impaired users, critical theory, persuasion.

ACM Classification Keywords
H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

General Terms

Experimentation, Human Factors

INTRODUCTION
Scam emails are so common and so old that there are now classics of the genre. Everyone who has an email address will, sooner or later, receive a request for their banking details so that untold millions can be deposited following the death of some wealthy client. Very often such "phishing" emails will be from Nigeria and indeed are known as "419" scams after the section number of the Nigerian penal code that deals with them [20].

There are numerous guides to spotting phish [e.g. 2, 16, 27]. They all agree that in order to avoid phish we must follow simple rules such as:

- Never click on a link in an email
- Never respond to an email asking for confirmation of banking details
- Only use up to date virus protection, spam filters, web browsers and operating systems

Forms of address in emails such as "dear valued customer" and poor spelling and grammar in emails from large organizations are also often mentioned as warning signs of possible phishes. The website of the Anti-Phishing Working Group offers extensive information on many aspects of phishing and how to counter them [1]. However, despite numerous education campaigns, enough people still fall for phish to make new phishing attacks worthwhile.

This paper attempts to understand how we fall for phish through four related studies using a variety of qualitative and quantitative methods. The first asks what strategies phishers are currently using. Although there are many educational campaigns, enough people are falling for them to make new attacks profitable, why? Are phish becoming more convincing? A content analysis of a phishing archive indicated that many phish contained no spelling mistakes and used convincing company logos. Are such phish more difficult to spot? Therefore the second study conducted an online survey asking people to distinguish phish from genuine emails. Would the use of logos make phish more convincing? The results indicated that phish were correctly detected in less than half of examples presented and significantly less so if they included a logo.

These two studies were complemented with two qualitative studies which addressed the question of what makes a successful phish. How are phish identified? In the third study, we conducted in-depth interviews with eight blind people in which we discussed their strategies for identifying phish. We suspected that blind people might be more vulnerable to phish because they might miss cues like visual warnings in security toolbars and browsers, but it became clear that these participants had robust strategies for identifying phish which were based on careful reading of suspicious emails. This led us to consider phish as

CHI 2011, May 7–12, 2011, Vancouver, BC, Canada.

literature. What are the main literary devices that phishers use? The final study therefore drew on critical theory to consider phish as a literary form.

PHISH AND HCI

Although many technical weaknesses remain in our computing systems, these often require some skill to find and exploit. This has left the human as the "weakest link" in the system and a prime target for attack. Phishing is the most common type of attack. Giani and Thompson [15] use the term "attacks against cognitive channels" to describe such attempts to deceive the user. Dhamija et al [9] suggest why users fall victim to attacks: a lack of understanding of how computer systems work; a lack of attention; and that phishers can be quite adept at using visual elements to instill a sense of trust. Wu et al [29] demonstrate the limited effectiveness of security toolbars and other browser security indicators. Jakobsson [17] summarised two investigations of the ability of average users to judge whether an email or webpage is trustworthy or phishy, concluding that: spelling and design matter; people do check URLs; people judge relevance before authenticity; personalization creates trust; emails are very phishy, web pages are a bit phishy, phones calls are not; padlock icons have limited direct effects; independent channels create trust; and people recognise common forms of attacks.

Downs et al [12] interviewed non-computer experts and found that people can manage the risks that they are most familiar with, but do not appear to extrapolate to be wary of unfamiliar risks. People are more aware of technical security attacks, but less aware of threats posed by social engineering. People have developed certain cues to detect phishing emails, but these develop over time and with experience. The attacks delivered by other media are less likely to be regarded as suspicious.

Few studies have looked at individual differences in susceptibility to phish attacks. However, Sheng et al [25] found that women are more susceptible than men and that young people are more susceptible than older people.

As people become more wary of "classic" scams, new ones may be emerging. Our first study undertook a content analysis of a phishing archive to gain an idea of the strategies currently in use by phishers.

STUDY I: PHISHING ARCHIVE CONTENT ANALYSIS

MillerSmiles [22] is an anti phishing website with an archive of more than one and a half million reports. It stores examples of phishes by type (e.g. "419" scams), but it also has an archive which presents the full text of scams reported for given periods. It is an invaluable resource for investigating the kinds of strategies employed by phishers. The following analyses are based on scams stored on the site collected during a five day period from 26th to 31st October 2009. There were 100 individual scams archived for the period. A content analysis [21] was performed on the purported sender and premise for each phish, the number of spelling mistakes and the use of logos.

Sender and Premise

The vast majority (82%) of the emails purported to be from banks or building societies. Other purported senders included: PayPal, eBay, Facebook, and email providers, as well as private individuals. The most frequently used premise (75%) for the phish were "security updates" followed by alerts caused by invalid logins. Unspecified "security checks" were also common, as were general "account updates". There were four emails which purported to indicate that new security messages were waiting to be opened in a secure website. Each of these strategies draw on real security protocols in order to direct victims to phishing sites. Just four of the phishes were not banking security scams. One purported to be a complaint from an eBay user who had sent money and received no goods. Two were happy to announce a big lottery or prize win based on radom email selection. One was a call to click agreement with racist sentiments.

Spelling and Grammar

Although poor spelling is often suggested as a good indicator of phishing attacks, just 11% of these phish contained three or more obvious spelling errors. Half contained one or two. The most common mistakes were missing or incorrectly used definite articles; mistakes with tense forms "after this will be completed" and incorrect forms of words "we advice you" rather than "we advise you". But 38% contained no obvious errors at all. Many of the correctly spelled emails had text from genuine advertising for banking or security services. Banks and secuirty advisors tell us to beware of poor spelling and grammar. But what happens when the phishers learn to present themselves in more convincing ways?

Figure 1: Phishing email purporting to be from HSBC Bank

Logos

It has long been understood in advertising that there is tremendous value in a recognisable logo. Brand recognition is an important part of establishing trust. Logos have never been easier to copy and paste into documents and the use of such visual aids may lend credence even to poorly written phishes. 64% percent of the phishes contained a logo, some featured other visual aids such as colour schemes, photographs and graphically designed layout copied from official campaigns of the targetted company.

Style

Although the vast majority of phish purported to be business letters from banks, another source for imitation was advertising. Some of the most convincing phish took this form, adopting logos, formatting and images from plausible campaigns on security. For instance, Figure 1 shows the body text of a phish purporting to be from HSBC Bank.

A small minority of these phish adopted a conversational and humorous tone. With a subject line of "DO I EVER AGREE!!!!!" this phish appeals to racism through (a sort of) humour:

> How all business phones should be answered!
> GOOD MORNING, WELCOME TO CANADA
> Press '1' for English; and,
> Press '2' to disconnect until you learn to
> speak English.

Many phish employed forms of brevity (e.g. "You have a new security message from HSBC. Click here."). Even briefer emails contained links that promised further information. Some of these emphasized the fact that they were automated messages, two were little more than lists of automatically generated text on when an email was sent and received, with a link for the curious to find out more. These are interesting because rather than attempt to come up with a plausible reason for contact they rely on a blank appeal to curiosity.

This study suggested then that spelling and grammar could not be relied on to give away a phish attack, as 38% were spelled correctly and 68% made themselves look convincing by copy pasting logos indistinguishable from the "genuine" online article. Are phish getting harder to spot? We designed the following study asking participants to distinguish phish from spam.

STUDY II: THE ONLINE PHISHING SURVEY

An online survey was undertaken in which participants were presented with phishing emails and genuine emails and asked to distinguish which were which.

The survey

The study was conducted as an online survey using QuestionPro. The phishing emails were taken from the MillerSmiles phishing archive and the "genuine" emails were items of genuine spam – advertisements or mass mailout charity appeals. Five of the 10 phishing emails contained logos and five did not. An introductory page explained that phish are scam emails attempting to trick people into giving away banking details or visiting malicious sites whereas spam, though annoying, are genuine advertisements. Participants were initially presented with an email subject line and asked whether it was a phish or not. They were then presented with the email text and the question was repeated. In each case, respondents were asked to answer on a four point scale: Definitely, Probably, Probably Not and Definitely Not.

After the email questions, respondents were asked a series of questions about their email habits and attitudes, particularly in relation to detecting phishing attacks, and some demographic questions. The survey took approximately 30 minutes to complete. A £10 Amazon voucher was offered to the first 200 participants who completed the survey, which proved a very effective incentive. The survey was sent to staff and student email lists at Departments of Biology, Computer Science, History and Psychology, friends and colleagues of the authors, and to an email list of an organization of blind computer users, so the sample is biased towards highly educated, computer literate people.

Respondents

224 people responded to the survey, 121 men and 113 women (10 people did not state their sex). Ages ranged from 18 to over 65 years, with the most common age group being 18 – 30 years. 87 respondents were students (covering a wide range of disciplines, with the most computer being biology, computer science and psychology), the rest either working (128) or unemployed (20) (9 respondents did not provide information about their employment status). 45 respondents had a high school education, 50 had a first degree, 75 had a Masters degree or diploma and 56 had a PhD (11 did not provide information about their education). 183 were native speakers of English, 51 speak English as a second language and 10 did not provide information about their native language. 32 respondents were blind or partially sighted.

Results

Overall accuracy and precision of phishing detection

The task of trying to correctly distinguish the phishing emails from the genuine emails can be conceptualized as a binary detection problem. The four possible outcomes are summarized in Table 2. If we take the desired outcome to be the correct detection of phishing emails (if we took it to be the correct detection of genuine emails to be the desired outcome, we would get a different but totally complementary matrix), then a True Positive is when a respondent correctly detects an email as a phish. Similarly, a True Negative is when a respondent correctly detects an email as a genuine one. However, a False Positive is when a respondent thinks an email is a phish, but it is actually a genuine email. So they are overly cautious, and reject a genuine email. On the other hand, a False Negative is when a respondent thinks an email is genuine when it is actually a phish, so they are taken in by the phish; this is of course, the worst outcome.

Two measures derived from these possible outcomes are typically calculated to summarize performance on detection tasks [28]:

- Accuracy = (Number of True Positives + Number of True Negatives)/(Number of True Positives + Number of True Negatives + Number of False Positives + Number of False Negatives)

Accuracy measures the proportion of correct responses in the total set of responses

- Precision = Number of True Positives/(Number of True Positives + Number of False Positives)
  Precision measures the proportion of correct positives in all the positive responses

		Respondents think the email is:	
		Genuine	Phish
But it is actually:	Genuine	TRUE NEGATIVE Respondent correctly detects a real email	FALSE POSITIVE Respondent is over cautious, thinks a real email is a phish
	Phish	FALSE NEGATIVE Respondent is taken in, thinks a phish is a real email	TRUE POSITIVE Respondent correctly detects a phishing email

Table 2: Possible outcomes of detecting a phish email

Table 3 shows the mean number of emails correctly detected in each category from viewing only the subject line of the email. The analysis was conducted using strict criteria for correct detection: respondents had to answer appropriately that the email was "Definitely (not) a phish" rather than "Probably (not) a phish" for their answer to be deemed to be correct. It can be seen that respondents were not very accurate at correctly detecting phishing emails from the subject line, with on average less than 4 out of 10 correctly identified; respondents were even less accurate at detecting genuine emails, with only on average just over 1 out of 10 correctly identified.

		Respondents think the email is:	
		Genuine	Phish
But it is actually:	Genuine	All: 1.25 Men: 1.33 Women: 1.18	All: 8.74 Men: 8.67 Women: 8.82
	Phish	All: 6.46 Men: 6.12 Women: 6.90	All: 3.54 Men: 3.88 Women: 3.10

Table 3: Detection outcomes from viewing the email subject line only

Table 4 shows the mean number of emails correctly detected in each category from viewing the subject line and then the email text. It can be seen that respondents were much more accurate at correctly detecting phishing emails from the body of the message than from the heading only, with on average nearly 8 out of 10 correctly identified; respondents were still less accurate at detecting genuine emails, with only on average just under 5 out of 10 correctly identified.

		Respondents think the email is:	
		Genuine	Phish
But it is actually:	Genuine	All: 4.31 Men: 4.57 Women: 4.07	All: 5.69 Men: 5.43 Women: 5.93
	Phish	All: 2.81 Men: 2.07 Women: 3.68	All: 7.19 Men: 7.93 Women: 6.32

Table 4: Detection outcomes from viewing the email subject line and body

Two three way Analysis of Variance [18] were conducted to investigate the effects of viewing the subject line versus the body of the message, as well as the age and gender of respondents on respondents' Accuracy and Precision of detection. For these analyses, respondents were grouped into 18 – 30, 31 – 45, and over 45 years of age.

For Accuracy, there was a significant difference between viewing subject line and message body ($F = 621.49$, $df = 1, 208$, $p < 0.001$), with Accuracy being higher after viewing the message body. As in Sheng et al's 2010 study [27] there was also a significant difference between men and women, with men being significantly more accurate than women ($F = 6.83$, $df = 1, 208$, $p < 0.01$). There was no significant effect for Age ($F = 0.66$, $df = 2, 208$, n.s.), however there was a significant interaction between Age and Sex ($F = 3.96$, $df = 2, 208$, $p < 0.05$). The interaction is illustrated in Figure 2, which shows that the difference between men and women is most marked for the youngest age group and disappears entirely in the oldest age group.

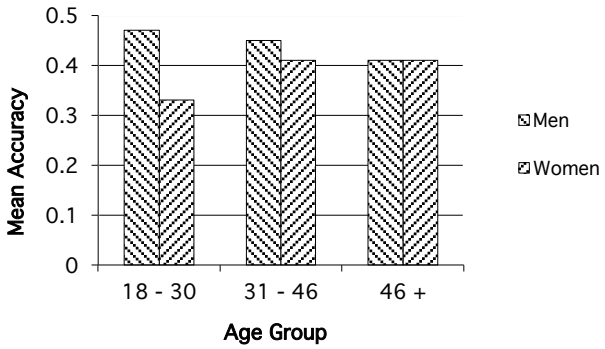

Figure 2: Accuracy of phishing detection

The results for Precision were very similar. There was a significant difference between viewing subject line and message body ($F = 503.23$, $df = 1, 208$, $p < 0.001$). There was also a significant difference between men and women, ($F = 5.73$, $df = 1, 208$, $p < 0.02$). There was no significant effect for Age ($F = 1.15$, $df = 2, 208$, n.s.), however there was a significant interaction between Age and Sex ($F = 3.69$, $df = 2, 208$, $p < 0.03$).

The analyses of variance were repeated using less strict criteria for correct detection: respondents had to answer

appropriately that the email was "Definitely (not) a phish" or "Probably (not) a phish" for their answer to be deemed to be correct. With this change in criteria, the significant sex and age differences disappeared, but the significant difference in both Accuracy and Precision of detection between viewing the subject line only and the message body remained (for Accuracy; $F = 112.15$, $df = 1, 208$, $p < 0.001$; for Precision: $F = 72.72$, $df = 1, 208$, $p < 0.001$). This change suggests the age and gender effects may be due to differences in respondents' confidence in their decisions. However although the gender difference confirms earlier research [25], there may be a confounding of gender and area of study/expertise of our participants in our study: our computer science participants are predominantly male, whereas our psychology participants are predominantly female. Participants from computer science may be more confident (possibly erroneously) about their ability to detect phish than participants from psychology. Our data was not extensive enough to explore these questions more fully.

Finally, we investigated the accuracy and precision of the sighted and visually impaired respondents. Based on the results of the interviews in Study II, we predicted that blind respondents would be more accurate and more precise in detecting phish emails than sighted respondents. However, there was no significant difference between the two groups in either accuracy or precision, using either the strict or less strict definitions of correct detection.

The effect of a logo

The effect of an appropriate logo in a phish email on respondents' ability to detect the phish was investigated. Five of the phish emails had logos and five did not. A three way analysis of variance was conducted on number of correct detections of the phish emails. This showed a significant difference in detection rates between phish emails with and without a logo ($F = 55.19$, $df = 1, 19$, $p < 0.001$), with better detection rates for phish emails without logos. There was also a significant interaction between the logo condition and the respondents' sex. Figure 3 shows that men detected more phish in both conditions, but the difference between the sexes was greater for phishes with logos.

Finally, and again contrary to our expectations, there was no difference between blind and sighted respondents in their reactions to the logos ($F = 0.29$, $df = 1, 229$, n.s.).

Attitudes and habits in relation to phishing

Respondents were asked a series of questions about their attitudes and habits in relation to detecting phishing emails, see Table 5. Each of these questions was based on the good practice guidelines suggested on anti-phishing sites, public service providers such the BBC and warnings issued by banks [2, 16, 27].

All these questions were answered on a 5 point rating scale from "Strongly agree" to "Strongly disagree". The responses to these questions were investigated to see whether they could predict respondents' Accuracy and Precision in detecting the phishing emails.

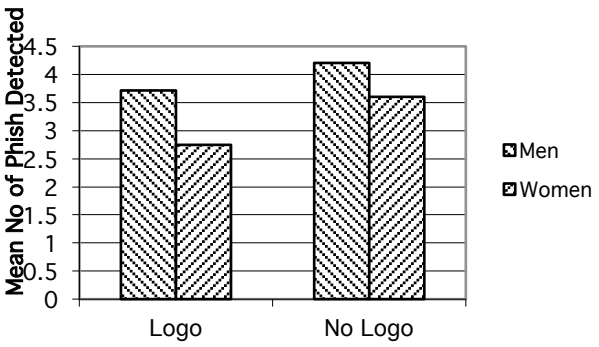

Figure 3: Effect of logo on phish detection

A principal component analysis [18] was conducted on respondents' answers to these questions, which resulted in three clear Components:

- Personal connection: questions about being suspicious/not opening emails/attachments from senders not personally known to the respondent
- Up-to-date Technology: questions about whether respondents have latest versions of browser, OS, anti-virus software
- Well-presented emails: questions related to suspicion of emails which do not have good spelling and grammar

A three way analysis of variance was conducted on the scores on each of these Components to investigate whether they were affected by age and sex. The only significant effect was between Component and sex ($F = 8.9$, $df = 2, 216$, $p < 0.001$). Figure 3 shows that women were more concerned about emails from unknown senders than men (note that low scores indicate agreement with the area of concern), whereas men are more concerned about having up-to-date technology. There is little difference between the sexes in attitudes to well-presented emails. But note that all means a below the mid-point of the 5 point scale, indicating that all respondents think all three Components are important and relevant to them.

Scores on each of these Components were also used in a linear regression analysis [8] to predict their Accuracy and Precision scores on phishing email detection. None of the Components significantly predicted Accuracy or Precision of detection from viewing the subject line only, but there were significant predictions for the Accuracy and Precision of detection from viewing the subject line and then the message body. These are summarized in Table 6.

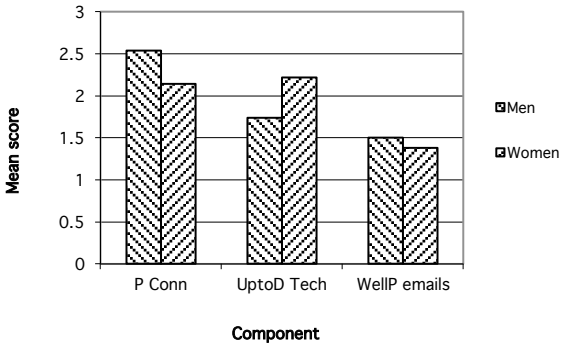

Figure 4: Mean scores for men and women on the three phishing attitudes components

Component	Question	Component loading
Personal connection	I would be suspicious of an email that did not address me by name	0.525
	I would not open an email from someone I did not know	0.734
	I would not follow a link in an email from someone I did not know	0.805
	I would not open an attached file from someone I did not know	0.807
Up-to-date technology	I update to the latest version of my web browser	0.863
	I update to the latest version of my operating system	0.860
	I update to the latest version of my anti virus software	0.660
Well-presented emails	I would be suspicious of an official email with spelling mistakes in it	0.886
	I would be suspicious of an official email with grammatical mistakes in it	0.890

Table 5: Components extracted from questions on attitudes and habits in relation to phishing

This quantitative analysis of the ability to detect phish indicates that the interpretation of email is no simple matter.

It is interesting to note that a number of participants thought that the survey itself was a phishing attack. This email was posted to a departmental discussion board:

> "It had a couple of the features - it was the first I'd heard about them collaborating on research of this sort, and offering 200 £10 Amazon vouchers seemed too good to be true […]"

Predicting	Significant effect?	Significant predictor variables
Accuracy of detection from subject line	Adjusted $r^2 = 0.016$ $F = 2.11$ df = 3, 206 n.s.	
Accuracy of detection from email text	Adjusted $r^2 = 0.042$ $F = 4.07$ df = 3, 210 $p < 0.01$	Up-to-date Technology Beta = -0.151 $t = 2.24$, $p < 0.03$ Personal Connection Beta = 0.136 $t = 2.02$, $p < 0.05$
Precision of detection from subject line	Adjusted $r^2 = 0.007$ $F = 1.51$ df = 3, 206 n.s.	
Precision of detection from email text	Adjusted $r^2 = 0.033$ $F = 3.39$ df = 3, 210 $p < 0.02$	Personal Connection Beta = -1.81 $t = 2.21$, $p < 0.03$

Table 6: Results of linear regressions predicting Accuracy and Precision of phishing email detection from Components

Following re-writes of the text surrounding the instructions about the Amazon vouchers, there were several inadvertent inconsistencies and errors that led to the following suspicion:

> "It does seem to be a deliberate attempt to make us suspicious going on here. I wonder if this is one of those think-you're-doing-one-thing-while-doing-another psychology experiments?"

Some of the participants who completed the survey did not supply their email address necessary to receive the Amazon voucher because they believed this was the phishing bait finally revealed. As the content analysis of the archive indicated, it is very common for phish to pose as security measures. It is not unreasonable to expect phishers to pose as researchers either. The form security conscious phish take will be returned to in the following sections.

Two of the interesting results of the online survey were that nearly 80% of phish were correctly detected but that the use of logos made them significantly more difficult to detect. To consider how phish were detected we conducted two qualitative studies. The first looked at the unusual situation of blind email users.

STUDY III: INTERVIEWS WITH BLIND EMAIL USERS

Anti-phishing technologies often rely on visual cues: for instance, Google Chrome's toolbar turns yellow when it visits a secure website. Researchers have experimented with systems to educate people about the possibilities of phish: for instance, a pop up screen shot might demonstrate that the URL a user thought they clicked was not the same as the URL that their web browser would take them to [19]. Though ingenious and helpful to sighted people, such visually based systems may be of limited use to blind or partially sighted users.

A small qualitative study was undertaken to investigate whether blind people were more vulnerable to phishing attacks. Eight blind people were interviewed about their email routine and asked to read 10 example phish and identify any points in the text that would arouse suspicion. Participants were recruited from members of the British Computer Association of the Blind (www.bcab.org.uk), so the sample of interviewees was highly computer literate. Nevertheless it was striking during the interviews how proficient and accurate they were in identifying phishing strategies. It also became clear that in some respects assistive technologies such as screen readers make identifying phish easier.

Screen reading technologies which read out information for blind users made spelling mistakes within the subject line and the body of emails very obvious: a word which a sighted reader might not notice as misspelled while scanning a document sounded quite garbled when read out loud by a screen reader. Even where the screen readers were used at great speed (very common amongst proficient blind users), these users spotted spelling mistakes immediately. Reflecting on this, one of the participants noted: "I've never thought of my screen reader as a security device before but I suppose in a way it is". Phishing sites occasionally register domain names to visually resemble their targets. For instance, someone posing as Lloyds bank might register a domain name that replaces the second letter "l" in the name with the number "1" - L1oyds. The number "1" bears a strong visual resemblance to a letter "l" but bears no auditory resemblance at all. These screen reader users would be far less likely to fall for such visual tricks. As previously noted, phishers are including logos which are exact copies of genuine logos. Copying logos has never been easier and the unsophisticated phisher can extract a genuine logo from web material and include it in an email. While such devices might deceive the sighted reader, blind users' screen readers merely registered the presence of a graphics file.

The blind interviewees were extremely cautious about computer security and described email routines in which anything remotely suspicious would be immediately deleted. None of the sample scams drawn from MillerSmiles as examples for this study fooled any of the blind interviewees. This is perhaps because these participants were highly computer literate but several participants noted that being blind meant having to be cautious and security minded.

The participants were highly sensitive to context and very quickly identified phish, sometimes laughing with recognition at old friends like the Nigerian "419" scam. Far from being more vulnerable to phishing attacks because they were blind it seemed that they might in fact be less vulnerable to increasingly visual emails that rely on logos and images. Each of the participants were careful and astute readers of emails. Often a phish would be identified in the first few lines following the omission of a personal address, the first spelling or grammar mistake or an unlikely sounding premise. It became clear that careful reading is central to the process of identifying phish. What then can literary theory add to an understanding of how phish work?

STUDY IV: LITERARY ANALYSIS

Our fourth study draws on literary and critical theory to consider why some phishing strategies remain effective despite high levels of computer literacy and frequent publicity campaigns warning users of the dangers.

Dourish et al [11] note that detecting and deleting phish is part of the more general practice of managing junk mail. The user experience of phish is primarily a literary one, reading the subject lines and texts of emails. The term "literary device" refers to particular techniques of writing. This includes technical devices such as the use of onomatapeia in comic books (*boom!*, *crash!*, *pow!*). The first thing to notice about phish from a literary perspective is that they are all forms of *pastiche*. Pastiche is a form of imitation: the style and form of a particular author or a genre are drawn upon to create a new piece of writing. The vast majority of the phishes sampled from Miller Smiles were pastiches of circular business letters from banks. Most of them were bad pastiches, but a worrying minority are quite convincing. Not merely because they are spelled correctly and use convincing logos, they are also stylistically convincing. Because the form of writing being pastiched (or directly copied) is that of the business letter, most of the devices are associated with formality.

The forms of address are courteous, often exaggerated to the point of being courtly. These attempts at formality often fail because of poor spelling or grammar, but they can also fail by going too far (e.g. "attention honourable beneficiary" or "Esteemed customer"). The strategies draw on previous limitations of legitimate circulars. Mass mailouts mean that an anonymous address form is used for legitmate reasons (e.g. "dear householder"). The absolute impersonality of the email is stressed, it is merely part of a routine check. Similarly when threats are made they are rarely personally menacing, the threat is a matter of bureaucratic routine and the consequence of ignoring it is usually inconvenience rather than calamity.

As computing technology has become integral to almost every aspect of culture literary and critical theory has become increasingly relevant to studies of HCI [e.g. 3, 4, 5, 6, 23]. Literary and critical offers a number of perspectives and analytical styles to interpret texts. The following sections draw on structuralism and psychoanalysis.

STRUCTURALIST READING

Structuralism drew on mid twentieth century anthropology and its accounts of ritual and myth in European colonies [14]. Like formalism, structuralism sought to break narrative into its constituent elements and this has been used to make story engines [e.g. 7]. Almost before it began it was superseded by post-structuralism which argued that meaning is not constructed in discreet units. What then

might a structuralist analysis of phishing emails look like? The constituent parts of a phishing email could be formulated as: interpellation, premise and instruction.

Interpellation

For the structuralist critic Louis Althusser "interpellation" is the process by which the state constitutes a subject. His most famous example is a policeman hailing a man on the street with the words "Hey, you there!" The man is unsure whether the policeman is addressing him or not but stops nevertheless and in so doing constitutes himself in a power relationship where he is subject to authority. Typical interpellations in phishing scams would be – dear valued customer, or client, or Sir or Madam. As previously noted phishers are subject to and exploit the same resources as junk mailers. They are also subject to the same limited forms of direct address to strangers. But as advertisers find it increasingly easy to use direct rather than anonymous address so too will phishers and other spammers. As other sources of contact from strangers develop it is likely that phishers will adapt to it – Facebook messages, online game systems like Xbox, bulletin boards and so on will provide new opportunities for more convincing interpellations

Premise

All phishing scams have a premise of one kind or another. Sometimes it is quite elaborate - a clerk at a bank discovers a way of accessing the account of a wealthy client who has died intestate; or quite simple - someone has tried to guess the password on your online bank account so it has been suspended. The main body of an email then will be exposition of one kind or another. There is a structural resemblance here to the joke, in that the set up is always a mis-direction. As with a joke the initial "feedline" raises expectations about how the rest of it is going to go. In a joke expectations are reversed with the punchline and the surprise causes laughter. In a phishing scam expectations appear to be fulfilled but results, at some later date, in fraud. As with jokes, if you've heard it before it will not be as effective. One of the main counter measures to phishers are programs of education. If people are educated about the way the scams work then they will not fall for them. However, the form of the security warning is now itself one of the most popular forms of phishing premise – there are new dangers, here are the ways to avoid them. Why this may be so will be returned to in the final section.

Instruction

The final basic constituent of a phishing attack is a call to action, some form of imperative command, an instruction - confirm details, respond within forty eight hours and so on. For this to be successful the plot must be plausible to the victim: maybe someone really has tried to access their online bank account, there might after all be a problem with their PayPal account details, maybe, just maybe, their email has been randomly selected in a lottery. For an instruction to be carried out there must either be belief or a willing "suspension of disbelief". Primarily the calls to action play on fear. The genre then could be thought of as horror, but it isn't quite horror. It is not a desperate sort of fear, there is seldom real menace or the threat of physical harm. Perhaps suspense is a better way of framing it. The instructions create suspense – what if it's real? What if there is a problem? The suspense is relieved when the link is followed or the bait is otherwise taken.

Decision making theory has been used to model user–phishing interaction into three stages:

"Construction of the perception of the situation; generation of possible actions to respond; generation of assessment criteria and choosing action" [10].

These stages correspond to the structural elements of the email content: the interpellation and premise construct a perception of the situation, the instruction generates possible actions. The model in decision theory is one of a user engaged in rational cognitions: "A user generates the criteria to evaluate the resulting gains and losses of possible actions, then evaluates those actions and chooses the best one" [10]. But what about irrational action and unconscious motivations? It is possible that such cognitive accounts of decision making may be supplemented with insights from psychoanalysis.

Psychoanalytic Reading

Psychoanalysis is now almost wholly ignored in departments of psychology. The success of cognitive psychology in both research and the treatment of mental illness has ensured that Freud and his factious successors have been consigned if not quite to the dustbin then at least to the literature department. In fields other than academia however, it has never gone away. The Public Relations industry was founded by Edward Bernays, the nephew of Sigmund Freud. Bernays explicitly drew on his Uncle's theories to make appeals to consumers which aimed not at their rational cognitions but their hidden desires. Woody Allen famously spent half a lifetime on the couches of psychoanalysts and like many other famous cases, remains avowedly neurotic. However psychoanalysis has been and remains successful in exploiting our neuroses in order to persuade us to watch films or buy products or both. Yet psychoanalysis is almost wholly absent from the persuasive computing literature.

Slavoj Zizek is a philosopher and critic who is a card carrying disciple of the French psychoanalyst Lacan and has written several books explaining Lacanian theoretical concepts through examples drawn from film and popular culture. Lacanian theory is based on a set of specialised terms and like all critical theory is often dismissed as jargon. Zizek defies those that reproach Lacan with being difficult with examples drawn from mass media and everyday experience. What might a Zizekian reading of phish look like?

Zizek is fond of a particularly gruesome story by Patricia Highsmith that is relevant here. "*The Pond*" is the story of a newly widowed woman who moves house with her young

son. At first she loves her new home but at the bottom of the garden there is a dark pond clogged with strange weeds and she worries that her child will fall into it. She hires a firm to kill the weeds but the roots grow back almost immediately and stronger than ever. She tells her son not to go near the pond and warns that if he should ever fall in he must pull at the weeds to get to the side. But her son remains attracted to it and her fears become unbearable. She asks the company to put more and stronger weed killer into the pond but when she gets off the phone she discovers that her son is missing. She finds him face down in the pond entangled in the weeds. After the funeral she returns to the pond and wades in to pull them out by hand. They now seem to be alive and the more she struggles against them, the more they drag her down into the dark water.

For Zizek the pond is "the sinthome" "the kernel of enjoyment that simultaneously attracts and repels us" [30]. Our fears do not simply appall us they also exert a strange fascination, we are uneasy about them, we return to them. We are uneasy about online security in just the way Highsmith's heroine is anxious about the pond. We return to it, we worry at it, and in doing so sometimes open ourselves to the possibility of becoming ensnared. The successful phish both repels and attracts its victim. It keys into existing unease about online security and creates suspense. There could be no more perfect form for a phishing attack then, than a warning about a phishing attack with instructions on how to avoid it.

Literary theory is particularly sensitive to the relationship between form and content. The form here is a pastiche of the discourse of personal online security as espoused not only by banks but also those who sell protection. Here the focus is relentlessly on the individual, we must protect ourselves and if we don't we have nobody else to blame, certainly not the bank or the security firm. We must take some action or other and continually anticipate new kinds of attack. The standard warnings and tips demand eternal vigilance and constant updates. We are to be continuously afraid and endlessly fascinated; we must be cautious and yet quick to act. It is this impossible position which makes the language of security the perfect medium for fraud.

CONCLUSION

This paper has reported findings from four related studies of phish. The content analysis suggested that phish are beginning to look more convincing with better spelling, grammar and visual appeals like logos. An online survey investigated whether participants could distinguish phish from spam. Although participants were well educated and computer literate, phish were not always detected. Detection rates for phish with logos were significantly lower than for those without.

While quantitative work on phishing is relatively common, qualitative studies are more unusual. In order to better understand what strategies people use to identify phish we also conducted in depth interviews with eight blind people. It was thought that participants might be more susceptible to phish, however they demonstrated robust reading strategies for identifying phish.

The central role of reading in identifying phish led us to consider phish as literature. The final study then considered phish as a literary from and drew on critical theory to understand phish as pastiche. It argued that the focus on the individual from banking and security services make banking and security service notifications an excellent form for phish to take as they play so neatly on our anxieties in terms of both form and content.

Like warnings about not sharing passwords, warnings about how to avoid phish look exclusively to the individual. A small but growing body of work on usable security is beginning to argue that theoretical security is not enough. Singh et al's recent study of password sharing amongst married couples argues strongly that actual practice must be the basis for definitions of security [26]. Warnings from banks about not sharing passwords are there as much to protect the banks as customers, who have little choice but to ignore them.

The emphasis on personal responsibility for online security is one of technological hegemony, the banks dictating the technological solutions that are most appropriate, regardless of the convenience and habits of the customers. Although we are all well aware that we should update our anti-virus software, our firewalls, operating systems and web browsers we do not necessarily do it. Or at least not as often as we know we should. Like Highsmith's heroine we try but always fail to protect ourselves. And the sense of continuous anxiety which the situation necessitates is the one which phishers are currently exploiting.

This paper has taken an interdisciplinary multi-method approach to the problem of phish. Although quantitative and qualitative methods such as surveys and interviews are common in HCI it is more unusual to draw on literary and critical theory. For the most part the large body of work on persuasive computing has drawn on cognitive accounts of psychology. However there is a long tradition of psychoanalytic work that may be of value in thinking about the ways that computers might be made more or indeed less persuasive.

REFERENCES

1. Anti-Phishing Working Group (APWG). http://antiphishing.org/.
2. Bank Safe Online: Protecting Yourself. http://www.banksafeonline.org.uk/protecting_yourself.html.
3. Bardzell, J. (2009). Interaction criticism and aesthetics. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '09)*. New York: ACM Press.
4. Bell, G., Blythe, M., Gaver, B., Sengers, P. & Wright, P. (2003). Designing culturally situated technologies

for the home. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '03) Extended Abstracts*. New York: ACM Press.

5. Blythe M. Reid J., Wright P. & Geelhoed E. (2006). Interdisciplinary Criticism: Analysing The Experience Of Riot! A Location Sensitive Digital Narrative. *Behaviour and Information Technology*, 25(2), 127-139.
6. Blythe, M., McCarthy, J., Light, A., Bardzell, S., Wright, P., Bardzell, J. & Blackwell, A. (2010). Critical dialogue: interaction, experience and cultural theory. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI 2010). New York: ACM Press.
7. Braun, N. (2004). Storytelling and Conversation to Improve the Fun Factor in Software Applications. In M. Blythe, K. Overbeeke, A. F. Monk, & P. C. Wright (Eds.), *Funology: From Usability to Enjoyment*. Dordecht, NL: Kluwer.
8. Cohen, J., Cohen, P., West, S.G. and Aiken, L.S. (2003). *Applied multiple regression/correlation analysis for the behavioural sciences*. Hillsdale, NJ: Lawrence Erlbaum.
9. Dhamija, R., Tygar, D. & Hearst, M. (2006). Why phishing works. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '06)*. New York: ACM Press.
10. Dong X., Clark J. A. & Jacob J. (2008). Modelling user-phishing interaction. *Proceedings of Human-System Interaction*, May 25-27, 2008, Kraków, Poland.
11. Dourish, P., Grinter, E., Delgado de la Flor, J. & Joseph, M. (2004). Security in the wild: user strategies for managing security as an everyday, practical problem. *Personal and Ubiquitous Computing*, 8(6), 391-401.
12. Downs, J. S., Holbrook, M.B. & Cranor, L. F. (2006). Decision strategies and susceptibility to phishing. In *Proceedings of the Second Symposium on Usable Privacy and Security (Pittsburgh, Pennsylvania, July 12 - 14, 2006) (SOUPS '06)*. New York: ACM Press.
13. Eagleton T. (2003). *After theory*. London: Penguin Books.
14. Easthope, A. & McGowan, K. (1992). *A Critical and Cultural Theory Reader*. Milton Keynes: Open University Press.
15. Giani, A. & Thompson, P. (2007). Detecting deception in the context of Web 2.0. In *Proceedings of Web 2.0 Security and Privacy 2007*. http://w2spconf.com/2007/.
16. HSBC Phishing Scams. http://www.hsbc.com/1/2/online-security/phishing.
17. Jakobsson, M. (2007). The human factor in phishing. In *Proceedings of Privacy & Security of Consumer Information '07*. http://markus-jakobsson.com/papers/jakobsson-psci07.pdf.
18. Jolliffe, I. T. (1986). *Principal Component Analysis*. Berlin: Springer Verlag.
19. Keppel, G. & Wickens, T.D. (2004). *Design and analysis: a researcher's handbook*. Upper Saddle River, NJ: Pearson Prentice-Hall.
20. Kumaraguru, P., Rhee, Y., Acquisti, A., Cranor, L. F., Hong, J. & Nunge, E. (2007). Protecting people from phishing: the design and evaluation of an embedded training email system. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '07). New York: ACM Press.
21. Mayring P. (2004). Qualitative Content Analysis in Flickr. In Kardorff, U. & Steinke, E. (Eds.), *A Companion to Qualitative Research*. London: Sage.
22. MillerSmiles.co.uk 419 scams. http://419.millersmiles.co.uk/.
23. Propp, V. (1968). *Morphology of the Folk Tale*. Texas: University of Texas Press.
24. Satchell C. (2008) Cultural Theory and Real World Design: Dystopian and Utopian Outcomes. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '08). New York: ACM Press.
25. Sheng, S., Holbrook, M., Kumaraguru, P., Cranor, L. F., and Downs, J. 2010. 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*. New York: ACM Press.
26. Singh, S., Cabraal, A., Demosthenous, C., Astbrink, G., and Furlong, M. 2007. Password sharing: implications for security design based on social practice. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI '07)*. New York: ACM Press.
27. The One Show: Phishing in Your Bank Account? http://www.bbc.co.uk/blogs/theoneshow/consumer/2008/10/30/phishing.html.
28. Wickens, T.D. (2002). *Elementary signal detection*. New York: Oxford University Press.
29. Wu, M., Miller, R. C. and Garfinkel, S. L. (2006). Do security toolbars actually prevent phishing attacks? In *Proceedings of the Conference on Human Factors in Computing Systems*. New York: ACM Press.
30. Zizek S. (1992). *Looking Awry: an introduction to Jacques Lacan through popular culture*. Cambridge, MA: October Books.