Security education against phishing: A modest proposal for a major re-think

Iacovos Kirlappos, M. Angela Sasse
Department of Computer Science, University College London, Malet Place, London. WC1E 6BT
Email: iacovos.kirlappos.09@ucl.ac.uk, a.sasse@cs.ucl.ac.uk

Abstract
Online shoppers are targeted by many scams. To date, user education on phishing has tried to persuade them to check URLs and a number of other indicators, with limited success. We evaluated a novel anti-phishing tool in a realistic setting -participants had to buy tickets under time pressure, and lost money if they bought from bad sites. While none of our participants bought from sites the tool clearly identified as bad, 40% of participants risked money with sites flagged as potentially risky, but offering "bargains". The analysis of post-session interviews with participants revealed that - when tempted by a "good deal", they did not focus on the warnings. Rather, they looked for signs they thought *confirm a site's trustworthiness*: familiar designs or brands, trust seals, ads, reference to social networking sites and professional-looking design were mentioned as reliable indicators of a legitimate site. We argue that user education needs to focus on challenging and correcting the misconceptions that guide current user behavior, and present an outline such an approach.

1 The Phishing Problem

Phishing – tricking computer users to disclose personal information, credit card details, usernames and passwords - has been a major problem for the past 15 years. The probability of an online shopper coming across a phishing website is alarmingly high, since many show up as results in popular web search engines. In a recent UK police operation, 7 out of the top 10 Google results for a popular brand of boots were found to be fraudulent websites[1]. In addition, 1 in 12 buyers of tickets for events reported having been caught out by a scam ticket website, with the average loss for each victim being £80. Most sites are taken down quickly once identified, but new ones are springing up every day, making the process of identifying and closing all of them impossible.

The disclosure of financial details to scam sites can not only lead to immediate monetary losses, but identity theft and its consequences (damage to a person's credit rating, or being linked with illegal activities conducted using their credentials). Even though some banks cover customers who had their credit card details stolen, this is unlikely to be a sustainable solution. These problems can lead to an overall loss of trust in online shopping, and deter consumers from engaging in any online financial transactions.

2 User Education about Phishing

Two major approaches have been used to protect users against phishing: Anti-Phishing Indicators and User Education. Dhamija et al. [1] explain that the first approach is ineffective because passive indicators are ignored by a significant percentage of users. Even when users notice the indicators, they often do not understand what they signify, and the inconsistent positioning on different web browsers makes the task of identifying a phishing site difficult. Schechter et al. [12] report that 53% of their participants still attempted to log into a site after their task was interrupted by a strong security warning. In the same study, removing the HTTPS indicator had no effect on the willingness of participants to enter their personal details in a site and removing site authentication images resulted to 97% of participants entering their personal details. The findings of both papers allow us to conclude that any technical anti-phishing measures need to be complemented with effective user education to improve on users' ability to detect phishing sites.

Significant effort has been put into user education, both by governmental organizations and academic institutions. To improve on the understanding of security by the public, the US Computer Emergency Readiness Team offers *"advice about common security issues for non-technical computer users"* on its site (*http://www.us-cert.gov/cas/tips/*).Kumaraguru et al. [7] developed the *PhishGuru* training system to teach users how to identify phishing attacks. The system sends out simulated phishing emails and delivers training messages when users click on the URLs included in those. Its effectiveness was tested with 515 participants; 28 days after the first email and, despite being given training more than once, 17.5% of participants still entered personal details into simulated phishing websites. This was a significant improvement from the 40.1% a control condition revealed on day 0, but still leaves 1 in 5users vulnerable. The same research group developed *Anti-phishing Phil*[13], an online game to teach users not to fall for phishing by explaining how to identify phishing URLs, where to look for cues in web browsers, and how to use search engines to find legitimate sites. They report improved user ability to detect phishing websites after receiving training: the false positive rate (phishing site identified as real) was reduced from 30% to

[1] Operation Papworth highlights the UGGly side of Google: http://www.browsermedia.co.uk/2009/12/08/operation-papworth-highlights-the-uggly-side-of-google/

14%, and the false negative (non-phishing site identified as spoof) was reduced from 34% to 17%. Despite those reductions, adding the two percentages together indicates that 31% of users are still not able to tell the difference between a good and a bad site.

Herley [4] argues that teaching users to check URLs is the wrong strategy because even diligent application of what is being taught offers users only limited protection against phishing. In his view, the effort/benefit ratio means they should ignore this advice, especially if the actual risk of financial losses is low.

Another reason why current education and training efforts may not be effective is because they assume that users are keen to avoid risks, and thus likely to adopt behaviors that may protect them. But in reality, most online shoppers are looking for good deals. They start from a search engine and are presented with links to various websites that present – often very tempting - offers. The opportunity to save a significant amount of money on something they need, or acquire something they might normally not be able to afford, makes users vulnerable. Stajano& Wilson [14] identify this as the *Need and Greed* principle that scammers exploit successfully: once scammers know what users want, they can easily manipulate them. To address this problem the UK Office of Fair Trading (*http://www.oft.gov.uk*) launched campaigns aiming to increase consumer awareness of fake shopping websites. The slogan *"If it sounds too good to be true, then it probably is"* appears here, and regularly in communications by law enforcement officers – so far with little success.

In line with Herley [4], we argue that current security education on phishing ([7] [13]) offers little protection to users who assess a potentially malicious site in this frame of mind. Security education needs to know what drives user behavior in this situation – what cues they are looking for, and how they interpret them. Successful security awareness, education and training has to do more than warn users of dangers – it has to target the misconceptions that underlie user actions. Whilst the results presented here focus on phishing, a shift in perspective could help to develop more effective security awareness, education and training in other areas of computer security.

3 Trust Cues in Online Transactions

Users shopping online face a situation of risk and uncertainty: they have to provide payment details and personal information to websites, and cannot be certain they will receive the goods they expect in return. Many online shoppers will take risks to gain benefits, and they look for trust cues to reduce the degree of uncertainty about the outcome – a trustworthy transaction partner is more likely to deliver. Riegelsberger et al. [9] developed a framework of trust signals that both transaction partners can emit, focusing on ways of incentivizing trustworthy behavior in transaction partners, but also incorporating signals that can be used to assess trustworthiness, such as "*professionalism"* of a site (e.g. absence of technical failures, breadth of product palette and usability) and *social embeddedness* (e.g. a retailers reputation amongst friends and relatives of a consumer).Combined with Kim et al.'s [5] findings that consumer trust directly and indirectly affects purchasing intentions, we can assume that a user's willingness to engage in a transaction is increased if the perceived risk is low.

Koufaris & Hampton-Sosa [6] conducted a study on the development of trust in online companies by first-time customers, identifying four factors affecting users' purchasing decisions:

1. Perceived reputation of the company.
2. Perceived usefulness and ease of use of the website.
3. Perceived security control.
4. The selection of products available (if wide range then more trustworthy).

However, the use of closed, specifically aiming to confirm those four factors, did not allow revealing any additional ones. Kim et al. [5] also discuss the effectiveness of third-party seals as an assurance of trust, concluding that they decrease the risk *perceived* by consumers, but that consumers know very little about their purpose, and what protection they offer.

Trust development principles are exploited by scammers, both in the real world and online [14], but the implications of users' trusting behavior have not been considered in the phishing context.

4 Study Description

Our study was originally designed to evaluate the effectiveness of a new active anti-phishing tool, SOLID (www.solidauthentication.com), which uses traffic-light security indicators to signal whether a website is genuine or fake (in a small box sitting outside the browser, Figure 1). *Green*, accompanied by the logo of the owner of the website, indicates a website's details match those expected. *Yellow* appears when the webpage fails some part of the authentication test. When a webpage is identified as malicious, a pop-up window appears before the webpage loads and the tool window turns *red*, explaining that a security risk exists, suggesting a redirection to the real website of the registered retailer. The user also has the option to close the current tab/window and proceed to the risky website if they want. If the user chooses to do that, the color of the indicator remains red. If the website is not registered with the tool, the color of the indicator is *gray*.

Figure 1: The SOLID window displaying the traffic-light convention color code

Participants were recruited using the UCL Psychology subject pool, which is open to the public. The requirements

for participation were being over 18, regularly use online shopping, and being able to visit the lab for one hour of testing. The standard reward for their participation was a £15 Amazon voucher, and an additional reward was provided to participants who chose safe websites in the experiment. Thirty-Six participants were tested in total:

- 17 (47%) were male and 19 (53%) female.
- Average age was 24 years (SD = 3.8)
- Average computer experience was 12 years (SD = 3.6).
- Average daily internet browsing: 4.5 hours (SD = 2.1)
- They receive 14 (SD = 7.7) emails per day.
- 35 (97%) of them had checked their account balance online at least once.
- 34 (94%) had transferred money to other people's accounts using online banking services.
- All had bought goods online in the past.
- 19 (53%) had configured a firewall.
- 18 (50%) had designed a website
- 8 (22%) had registered a domain name.
- 7 (19%) recalled using SSH in the past.
- 9 (25%) had been victims of phishing, or knew someone who had been.
- 12 (33%) had been victims of Internet scams, or know someone that has been.

Participants were equally divided between two conditions: 18 used the active anti-phishing tool, and 18 did not. They were asked to buy tickets for a music festival, presented with 6 websites, and asked to decide within 5 minutes which one to buy from. This timeframe was used to replicate the "Time Principle" identified by Stajano and Wilson [14] as a tactic used by attackers- very plausible in this case, since tickets for popular events tend to sell out quickly. To replicate the risk that ticket buyers face when buying from unknown retailers online, the reward given to participants varied, depending on which website they chose, based on the following scenario:

"You want to buy tickets for Friday 27th of August for the LED electronic music festival at the moment they will go on sale. You have £60 available. You know that festivals sell out very quickly, so you only have 5 minutes to buy those. You have searched in Google for "LED festival tickets" and came upon 6 websites that claim to sell tickets. You now need to choose from which one to buy. Your additional reward from the experiment is the amount of money you initially have available (60 pounds) minus the price of the tickets on the website you chose to buy from. If you buy from a fraudulent website then you get no extra reward (only the 15 pounds that are paid for you participation in the experiment). You can browse in the websites with no limitations. Warnings will be given to you when 2 and 1 minute are left."

All the websites used in the experiment were local copies of legitimate retailers downloaded from the Internet. Our DNS server was modified so that the sites appeared to the participants in the same way as if they were browsing online (URL structure and website appearance). SOLID was modified to display the colors shown in Table 1.

Table 1: The websites used in the experiment with the corresponding prices and colors

Website	Ticket price	Tool Color
Gigantic (www.gigantic.com)	£50	Green
HMV Tickets (www.hmvtickets.com)	£50	Green
See (www.seetickets.com)	£25	Red
Skiddle (www.skiddle.com)	£20	Gray
Sold-out ticket market (www.soldoutticketmarket.com)	£40	Gray
View London (www.viewlondon.co.uk)	£20	Yellow

Table 2 shows that most participants who used SOLID chose the safe options (green), and none chose the website marked as *Red* ($X^2(1)$ p = 0.03324). Whilst this could be argued to be a success, a significant number still chose sites labeled as potentially risky (*gray* or *yellow*) over the ones clearly labeled as safe. Why did so many participants ignore the potential risks when a safe alternative existed?

Table 2: Distribution of participants' potential rewards based on the color of the website they have chosen

Potential Payoff	Number of participants	
	Control Condition	SOLID
£10	5	10 (green)
£35-40	12	8 (gray/yellow)
£20	1	0 (gray)

5 Identified trust factors and user misconceptions

In the debrief interviews following the experiment, each participant was asked to explain what affected their choice of website. No guiding questions were used - participants were free to report any factors that affected their final choice. During this discussion, the websites were left open so participants could refer back to them. The interviews were audio-recorded, transcribed and analyzed using Grounded Theory [3] coding techniques. The results show that security indicators were only one amongst several different signals that our participants used to assess the legitimacy of a website. We identified eight factors that affected the participants' choice of websites (detailed below).

All 8 participants in the SOLID condition who chose potentially unsafe yellow and gray sites said the potentially higher reward was an incentive to ignore the green site – confirming the *Need and Greed* principle [14]. Participants mentioned on average 3 (SD = 1.35) additional factors each that their decision:

1. Previous experience with website.

Previous experience with a website and familiarity with a brand induces users' willingness to trust it. With the exception of 1 participant, they had never shopped from any of the six websites, but 18 (50%) said they had heard of the brand names and this played a key role in their choices – suggesting a *"trust halo effect"* [8]. An example of this is the *View London* website, which five (14%)

participants had used to read venue reviews, but never before to buy tickets for events. Brands like *View London* and *HMV* are popular in the UK, the first because of its review pages, and the second because of high street retail outlets, which sell music and gaming products - but none of our participants was familiar with their ticket-selling operations. This very broad concept of "being familiar" with a brand can be exploited by scammers by creating fake websites, claiming to be online outlets of familiar brands.

2. Logos and certifications

Five websites displayed some form of trust logo, and 10 participants said those played a major role in their decisions. The "VeriSign Secured" logo (Figure 2) turned out to be the most popular one. Six participants (17%) said they trusted this sign because they had seen it on other trusted websites. But none of them could explain what the logo stands for, and why a website displaying it should be secure. Only two participants checked whether the logo was a clickable link, and what information about the merchant it was providing.

Figure 2: "VeriSign Secured" logo

The *Internet Shopping Is Safe* (ISIS) logo (Figure 3) was displayed in one website, and 4 (11%) participants reported it affected their choice. Another logo, the *Hitwise No1 Award Winners* logo (Figure 4) was displayed on one website, but only one participant mentioned it affected their choice. Logos of financial organizations – like the credit/debit cards accepted –were mentioned as trust factors by 3 (8%) participants: *"They accept VISA, MasterCard and AMEX so they must be real".* Again, only2 (6%) participants checked whether the logos were clickable links, or displayed a valid certificate or registration number with the relevant authority. In total, 13 participants claimed that logos affected their choices, but no one could explain why those signal trustworthiness.

Figure 3: ISIS logo

Figure 4: Hitwise logo

3. Advertisements

There are other forms of affiliation a website can suggest with known entities that are interpreted as trust signals: advertisements by well-known companies were mentioned by five (14%) participants, who argued that a reputable company would not pay scammers to advertize on their website. It had not occurred to them that scammers might include ads to make their site look legitimate, and that the companies advertised may not be aware of this.

4. Social networking references

The growing popularity of social networking websites is starting to affect online commerce in many ways; scammers exploit this by suggesting their site is associated with those. Inclusion of links to *Facebook* and *Twitter* pages boosted 7 (19%) participants' confidence in a site – they believed that links to those sites cannot be fraudulent, because any scam victims could post negative feedback to warn others. The iconic status that social networking sites have for many users is a key weakness if they do not understand how easily an association with the sites can be suggested or faked.

The presence of user feedback can also contribute to trust development, and received positive comments by 4 (11%) participants. This was particularly clear in the case of a website which included pictures of users who left feedback, or of other members of the website who "are planning to attend an event", and confirms past findings that richer media representations can be used to induce a positive trust bias [10].

5. Inclusion of charity names

The inclusion of the name of a charity (Oxfam) on one website (*www.gigantic.com*), accompanied by a claim that they donate 10% of their profits to it, lead 2 participants (6%) to believe the site was genuine. Benevolence – signaled by the website through a charitable donation – is an intrinsic trust property [9].Real-world scams exploit this, using charities as a pretext – e.g. collecting donations of money or clothes that the scammers actually keep for them. Online scammers can also exploit this, because users are not aware of the potential misuse of charity names, and do not attempt to verify the claims they see on websites.

6. Amount of information provided

The amount of information the website included on the event of interest was reported as an important factor by 6(17%) participants. All websites included information on the event (gate opening times, facilities, instructions how to get to the venue etc.), but those that displayed the information on the main event page attracted participants more. Again, addition of rich media like maps made them appear "more real" and trustworthy. In general, participants seemed to follow the maxim that the more effort is put into the development of a website, the less likely it is to be the site of scammers, who want to make money fast.

7. Website layout

7(19%) participants mentioned that the structure of the website design appeared familiar, because it was similar to other legitimate websites. This similarity led them to assume the site ought to be genuine. Interacting with particular websites leads to 'mental anchoring' of the design and appearance of the trustworthy sites, against which they assess trustworthiness of a new site on a first-time interaction. Participants were also re-assured by indicators of routine business – in this case, availability of tickets for a variety of events. They simply assumed a scam site would try to target a particular event.

8. Company information

The level of detail the website provided on the company behind it also affected participant decisions. 5(14%) participants mentioned the presence of the registration

number of the company; tax reference numbers, direct telephone numbers, ticket delivery information, and claims that they are official ticket outlets increased their confidence in the website. But as with logos and privacy policies, none of the participants knew how to verify this information, and did not attempt to do so.

6 Effective anti-phishing education

6.1 What should we teach users?

The results of our analysis reveal a significant gap between the signals security experts would like users to consider when assessing the legitimacy of a website, and those they actually use when faced with a tempting offer. Our findings – which unite and confirm a set of observations from previous studies – suggest that advice given in current user education is largely ignored because it focuses on indicators users do not understand or trust. To help users we need to explain how and why the indicators of trustworthiness they use successfully in the real world fail them online. As Wash [15] puts it, users form their own *'folk models'* when dealing with computer security issues, and use those to justify their decisions to ignore expert advice. Our participants ignored SSL locks and URLs, and used their own heuristics to assess the legitimacy of a site.

Reliance on indicators and models from the physical world leave users vulnerable in many ways:

- Participants were surprised when told after the experiment that fake versions of real websites can be uploaded by anyone online, or that someone can create a website claiming to be someone else.
- The fact that 13 participants used trust logos to guide their choices may seem encouraging, but only two checked whether those logos were clickable links, seeking more information on the certification and the merchant. None of our participants could explain what protection those logos might offer: they reacted to the mere presence of those as safety indicators.
- The 'blind trust' users place in sites which suggest a link with popular Social Networking sites demonstrates their popularity, and a worrying potential for exploitation by scammers. Our participants did not consider that anyone can create a page or profile in those sites, and claim anything they want, or that logos of social networking can be added by scammers in their fraudulent sites.
- The other design elements participants reported (amount of information provided, website layout and company information) can also be easily mimicked. Our participants seemed unaware that - whilst signals of high levels of investment are reliable indicators of the motivation of real-world retailers - design elements can be copied in a matter of seconds, and website design outsourced to developers in low-wage countries.

In summary: whilst many of our participants were confident in their ability to assess the legitimacy of a website, their assessment relied on trust signals that scammers can easily fake. Users also respond to mere references to entities they trust from their everyday experiences – names of companies, charities, etc. mentioned on sites. They do not check the legitimacy of such claims or detailed information posted, which makes it easy for scammers to defraud them.

Current security education approaches do not target the misconceptions we identified. Rather than telling users to look for broken links, we ought to tell them that online, they cannot rely on trust indicators that work for them in the real world. Users do not understand how scammers operate, and make assumptions about how the online environment operates based on their real-world experiences. Effective security education needs to:

1. Challenge users' assumptions about trust signals, and their decision processes, and
2. Replace them with trust signals and strategies for assessing risks in the online environment.

Security education needs to be mindful that – just as in the physical world – some users want to take risks in the online environment. So it should:

3. Equip users to assess the potential risks and benefits correctly, rather than tell them to avoid going to any potentially risky site.

6.2 How should we teach users?

The first step towards effective user education is to recognize that awareness, education and training are three distinct steps of a process to improve user competence [11]: The role of security awareness is to attract users' attention, and help them realize that there is a problem that might affect them. This is a necessary first step to render them receptive to education and training measures. Security awareness measures need to capture users' attention using strong visual elements, surprise, or humor. In the case of phishing, existing perceptions need to be challenged - users' perceptions of their ability to assess the risks involved in online transactions, and what reliable indicators of trustworthiness are. An example would be an advertisement – online or in print – that shows two very similar websites with the caption like *"One of these websites belongs to {a famous bank}; the other is run by a criminal gang in Elbonia waiting to steal your username and password and empty your account at {famous bank}.* Can you tell which is which?" Once users realize they cannot tell the difference, or chose the wrong site, they are more likely to pay attention to a subsequent pointer to a site that offers education (to improve their knowledge) or training (to improve their skills).

An example for delivering security education in this particular context would be a game in which users can collect or lose points by answering questions about the trust and assurance indicators (identified above) on a professional-looking website. For instance, if they point at an ad on the site, they would be presented with the statement "*The presence of and ad by {famous brand) indicates this is a legitimate site, because {famous brand)*

would not pay to advertise on a phishing website" and asked to rate it as True or False. Explanations of why an answer is true or false can help to correct misconceptions, and re-enforce correct statements; high scores or badges can motivate individual users, or groups in an organizational setting.

6.3 How could we reach users?

Another fundamental aspect of delivering effective security education is the choice of communication channel to disseminate awareness, education and training information to users. To date, two different approaches have been used:

1. General public awareness and education campaigns (both online and offline), and
2. Context-specific warnings and indicators (online).

In public awareness campaigns, users are informed about the risk of scams, and sometimes told about possible ways of protecting themselves from those, but no training is delivered. The effectiveness of those campaigns is questionable. Approaches like UK police campaigns, promoting general truisms such as*" If it looks too good to be true, then it probably is"*, do not provide any useful information or skills to consumers. Many legitimate online retailers sell goods at significantly lower than high street prices, and that is a major draw for online shopping. So how can consumers tell when "a good deal" becomes" too good to be true"? Generic warnings like this will deter many who would most benefit from lower online prices – people on lower incomes – from shopping online altogether, since they can least afford to take a risk [9].

A more promising approach is to provide awareness, education and training in the context of the services the users aim to access. Consumers are more motivated if warnings are specific to risks they know and care about, and more likely to be accessible when explained by peers who have a similar perception of risks and pitfalls. An example worth following is eBay (*www.ebay.com*), which has created an online community where users can post tutorials on how to identify counterfeit goods, or how to be careful not to fall for scams (and this is often featured from eBay's homepage). Another context-specific approach is used by a UK bank that asks its customers for partial PINs and passwords to access their online banking accounts (e.g. digits 2, 1 and 4 of the PIN and digits 2, 6 and 9 of the password): In its login page it explicitly mentions that users should never disclose their full PIN and password to a website, aiming to teach their customers the principle of not disclosing their full password, thus protecting themselves from password capture using phishing attacks.

Both the above measures increase user awareness on how scammers may target them, when using those specific websites, but also aim to educate them by explaining how to avoid falling for those attacks. But this is still no training, which is about not only presenting correct behaviors to users, but also testing users understanding on the communicated information and correcting any identified misconceptions [11].

A potential user-training approach is to create short tutorials, included in the retailer's and bank's websites, which could be used to assess and improve users' understanding of the information communicated to them. After we have informed users about the potential of a criminal gag in *Elbonia*, we need to draw their attention to the differences between a legitimate and a scam website (e.g. *your bank would never ask you to disclose your full PIN and password*). To ensure correct skill acquisition at the final steps of the tutorial users should be asked to distinguish between a few examples of legitimate and scam sites, based on the principles presented to them. To encourage user participation retailers could launch competitions with prize draws, incentivizing their customers with potential rewards for the time they spent taking the tutorial.

6.4 Lessons learnt from misconceptions

Trust symbols like logos and certifications are currently either misinterpreted, or go unnoticed. Only a third of the participants in the study reported that trust seals affected their decision, and none of those knew what they mean, who has the authority to certify that the site is genuine, and what protection they would receive in case of problems.

Trust seals can only be effective if users are able to recognise them, know what protection they offer, and check their legitimacy [9]. Since this is not the case, broader awareness campaigns using a range of information channels are needed. First attract people's attention to the presence of those seals, then explain what the problem is and what measures are in place to protect them (in this case a browser add-on) and provide them with information on what needs to be done on their side. SOLID had a significant effect deterring participants from known bad sites. Active anti-phishing tools, which interrupt the user's primary task only when a threat is identified, seem to be an effective measure against phishing attacks - confirming [2]. But to improve user defenses against future scams, an additional step is required: Whenever unauthorized use of trust symbols is detected, users should be presented with information on what went wrong, increasing their awareness on the problem and the potential risks they face while shopping online. This needs to be done in the browser when users visit sites that carry those seals, so that it does not require users to download and install additional software to be protected. In addition, whenever a risk is identified, short tutorials with strong eye-catching visual artifacts should be used, ensuring users understand the nature of the problem, what the messages delivered to them mean and also correct any potential user misconceptions. The information delivered should be short and descriptive so that it does not appear as "too much effort" to users, as they may then ignore it.

Users seem to trust sites that appear familiar. This can be used to the retailer's advantage - established brands can provide easy recognition and reassurance to customers. But customers will expect trusted institutions guaranteeing a transaction, and help them if they go wrong. This can enable consumers to engage in transactions where the perceived level of risk is higher than what they would

otherwise accept. An example of a well-trusted organization is PayPal (*www.paypal.com*): It is a payment method that provides the user with the advantage of having their card details not visible to the seller, and also guarantees to refund its customers when transactions go wrong. Including support for payment methods like this on a site could increase the overall willingness of consumers to buy from it. The presence of those mechanisms alone is not enough though: Users again need to be made aware of the potential problems they may encounter when shopping online (e.g. receive counterfeit products, receive nothing, have credit card details compromised etc), and to what extent they are protected, provided they comply with a manageable set of rules. This can be achieved by getting big retailers on board to use those mechanisms, and provide visual elements to explain to users how they are protected. Statements like: "Paying by *EasyPay* ensures your card details are not shared with anyone when buying online" can increase customer confidence in e-commerce. This could be accompanied by short tutorials labeled with phrases like: *"how am I protected?"* that explain to the users in more detail what can go wrong in an online transaction (e.g. not receive the goods) and how they are protected. Any approach attempting to do this should be consistent across online retailers/service providers to avoid flooding the users with varying information, causing confusion instead of aiding their education and skill acquisition.

In terms of the visual cues users use to trust a site, they need to be made aware how easy it is for attackers to mimic these elements [9]. An engaging, - though perhaps controversial – approach to achieve this would be to create a YouTube video demonstrating "How to create your own phishing website in 10 mins and 5 easy steps", and spreading the word through social networking sites.

7 Conclusion

Our findings suggest the need for a change of direction in security awareness, education and training. Instead of flooding users with warnings, and keep telling them to behave as security experts would like them to, effective security awareness starts with the users' perspective and decision-making processes, imperfect they may be. Users form their own models of risk, and use a set of heuristics to assess the trustworthiness of the websites they interact with. These heuristics are currently influenced by the way trust signaling works in the physical world, and scammers have been able to successfully exploit this [14]: a well thought-out scam can create a propensity to follow it, even before users start thinking about security, providing them with an incentive to comply with the scammer's instructions that is too strong to be ignored, either by giving them good deals and/or creating professional-looking online stores. Having identified the misconceptions users form, we need to connect with them through specific awareness, education and training campaigns. First, attract users' attention to the problem, explain how they can be targeted (e.g. by fake trust symbols) and explain what makes them vulnerable (*"trust symbols alone do not signify trustworthiness"*). This will improve on the users' ability to accurately perceive the risks they face when shopping online, making them more receptive to education measures. Education should explain to the users what mechanisms exist to protect them and how to use those (e.g. automatic verification tools). Finally, training can improve specific user knowledge and skills in the context of the sites and services they use.

Campaigns need to address the retailer side as well. We identified some examples of bad practice among legitimate retailers, who do not provide reliable trust signals, or allow scammers to exploit potential vulnerabilities in their website design. They need to be made aware on how their websites – and their customers – are attacked, and how they can help customers distinguishing between their legitimate website and scam ones. This could help them protect their customer base and their reputation against this type of attack.

Our proposed new approach to security education can be generalized beyond anti-phishing, to the extended security community. Flooding users with large amounts of information on what to do to stay secure does not seem to work as they:

1. Do not understand the details or the purpose of that information (due to lack of accurate threat understanding) and
2. Do not care about security when using technology as is seems to be *"too much effort"* to them.

What needs to be done instead is to consider how users make decisions in their everyday activities (both in business and personal settings) and try to tailor newly-proposed security solutions based on this, accommodating their work or personal goals when interacting with technology and the *folk models* they form on the virtual world based on their real-world experiences.

8 References

[1] Dhamija, R., Tygar, J. D. and Hearst, M. Why phishing works. In CHI '06: Proceedings of the SIGCHI conference on Human Factors in computing systems, pages 581-590, 2006.

[2] Egelman, S., Cranor, L. F. and Hong, J. You've been warned: an empirical study of the effectiveness of web browser phishing warnings. In Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems, pages 1065-1074, ACM, New York, NY, USA, 2008.

[3] Glaser, Barney G. and Strauss, Anselm L. The discovery of grounded theory: strategies for qualitative research. Chicago.: Aldine, 1967.

[4] Herley, C. So long, and no thanks for the externalities: The rational rejection of Security advice by users. In Proceedings of the New Security Paradigms Workshop 2009, pages 133-144, 2009.

[5] Kim, D., Ferrin, D., and Rao, H. A trust-based consumer decision-making model in electronic commerce: The role of trust, perceived risk, and

their antecedents. In Decision Support Systems, 44(2):544-564, 2008.

[6] Koufaris, M. and Hampton-Sosa, W. The development of initial trust in an online company by new customers. In Information & Management, 41(3):377-397, 2004.

[7] Kumaraguru, P., Cranshaw, J., Acquisti, A., Cranor, L., Hong, J., Blair, M. A., and Pham, T. School of phish: a real-world evaluation of anti-phishing training. In SOUPS '09: Proceedings of the 5th Symposium on Usable Privacy and Security, pages 1-12, New York, NY, USA. ACM, 2009.

[8] Leuthesser, L., Kohli, C. S., Harich, K. R. Brand equity: the halo effect measure, In European Journal of Marketing, Vol. 29 Issue: 4, pp.57 – 66, 1995.

[9] Riegelsberger, J., Sasse, M. A., and McCarthy, J. D. The mechanics of trust: a framework for research and design. In International Journal of Human-Computer Studies, 62(3):381-422, 2005

[10] Riegelsberger, J., Sasse, M. A., and McCarthy, J. D. Rich Media, Poor Judgement? A Study of Media Effects on Users Trust in Expertise. In Proceedings of British HCI Conference, pages 267–284, 2005.

[11] Sasse, M.A, Ashenden, D., Lawrence, D., Coles-Kemp, L, Fléchais, I., Kearney, P. Human Vulnerabilities in Security Systems, Human Factors Working Group, White Paper, Cyber Security Knowledge Transfer Networks, 2007.

[12] Schechter, S. E., Dhamija, R., Ozment, A., Fischer, I. The Emperor's New Security Indicators. IEEE Symposium on Security and Privacy, 20-23 May 2007.

[13] Sheng, S., Magnien, B., Kumaraguru, P., Acquisti, A., Cranor, L. F., Hong, J., and Nunge, E. Anti-phishing Phil: the design and evaluation of a game that teaches people not to fall for phish. In SOUPS '07: Proceedings of the 3rd symposium on Usable privacy and security, pages 88-99, New York, NY, USA. ACM, 2007.

[14] Stajano, F. and Wilson, P. Understanding scam victims: seven principles for systems security. In Communications of the ACM, 54 (3), pages 70-75, New York, NY, USA, 2011.

[15] Wash, R. Folk models of home computer security. In SOUPS 2010: Proceedings of the 6th Symposium on Usable Privacy and Security, SOUPS '10, pages 1-16, New York, NY, USA. ACM, 2010.