

Cyber Security Training Game Teaches People to Avoid Phishing Attacks

Wombat Security Technologies

<http://wombatsecurity.com/>

July 2010

Executive Summary

A number of scientific studies have demonstrated that employees are not motivated to pay attention to traditional cyber security training or read online training materials. However, **Wombat's Anti-Phishing Phil™** was designed to be fun and engaging to entice users to learn about anti-phishing strategies in context. The game was built on learning science principles aimed at enhancing learning and retention of practical knowledge and is the result of a number of iterations and refinements. This white paper summarizes the results of extensive tests of Anti-Phishing Phil conducted to evaluate the effectiveness of the game when it comes to teaching users how to best protect themselves from phishing attacks. The tests involved a total of 4,517 participants. Results from the study clearly show that playing the Anti-Phishing Phil game for ten minutes or less can greatly reduce the likelihood of users falling for phishing attacks. The effectiveness of the game is greatest with those users who are most prone to falling for phishing attacks in the first place, namely those employees or customers who represent the greatest source of vulnerability for an organization. For these users, the study showed reductions of 50 percent or more in the chance of falling for a phishing attack.

Key Findings

- **Anti-Phishing Phil training succeeds where traditional training fails.** We present the results of a study in which over 4,500 people played Anti-Phishing Phil. Results of the study show that people who played the game are able to distinguish phishing websites **more accurately and more quickly** than those who have not.
- **Users find Anti-Phishing Phil engaging and the game requires minimal time.** The interactive design of Phil makes it fun for people to play, and it only takes ten minutes to complete the game. In other words, the game is not just more effective than traditional training approaches but it is also more likely to be used by a significantly greater number of people, thereby resulting in a substantially higher ROI for an organization.
- **The study shows 50 percent reduction in people falling for subsequent phishing attacks.** In our study, participants who fell for a phishing email and received Anti-Phishing Phil training were 50 percent less likely to fall for subsequent phishing emails.
- **Users remember Anti-Phishing Phil training.** Users trained with Anti-Phishing Phil have been found to retain the knowledge they learned and be able to apply it longer than those trained with more traditional approaches.

The Anti-Phishing Phil Game

Anti-Phishing Phil is an online interactive game designed to teach users how to identify phishing URLs and avoid phishing scams that may compromise their computers, their data or that of their enterprise. During the game, users assume the identity of a fish named Phil who swims through the ocean deciding whether to eat or reject worms associated with URLs. The game is split into four rounds, each of which is two minutes long. Before each round begins, Phil's friend PhishGuru presents you with tips to help you successfully navigate web addresses.

- In each round, Phil is presented with eight worms, each of which carries a URL that is displayed when Phil moves near it, as shown in Figure 2. Each worm is associated with a URL, and Phil's job is to eat all the real worms (legitimate URLs) and reject the bait (phishing URLs) before running out of time.
- The player is given a stipulated time in which they have to perform (and thereby learn) the things that are presented in the game.
- Each round increases in complexity, and by round two, there are enemies that you need to avoid so as not to lose a life.

- Phil is rewarded with 100 points if he correctly eats a good worm or correctly rejects a bad one. He is slightly penalized for rejecting a good worm (false positive) by losing 10 seconds off the clock for that round. He is severely penalized if he eats a bad worm and is caught by phishers (false negative), losing one of his three lives.
- Players have to correctly recognize at least six out of eight URLs within two minutes to move on to the next round. As long as they still have lives, they can repeat a round until they are able to recognize at least six URLs correctly. If a player loses all three lives the game is over.
- At the end of every round a review screen is displayed, showing all of the URLs from that round and tips for identifying them correctly, as shown in Figure 3.

The game is implemented in Flash 10.0, and the content for the game (including URLs and training messages) is loaded from a separate data file. In each round of the game, four good worms and four phishing worms are randomly selected from a customizable collection of URLs in the data file for that round. Sound effects are used to provide audio feedback, and background music and underwater background scenes help keep the users engaged.

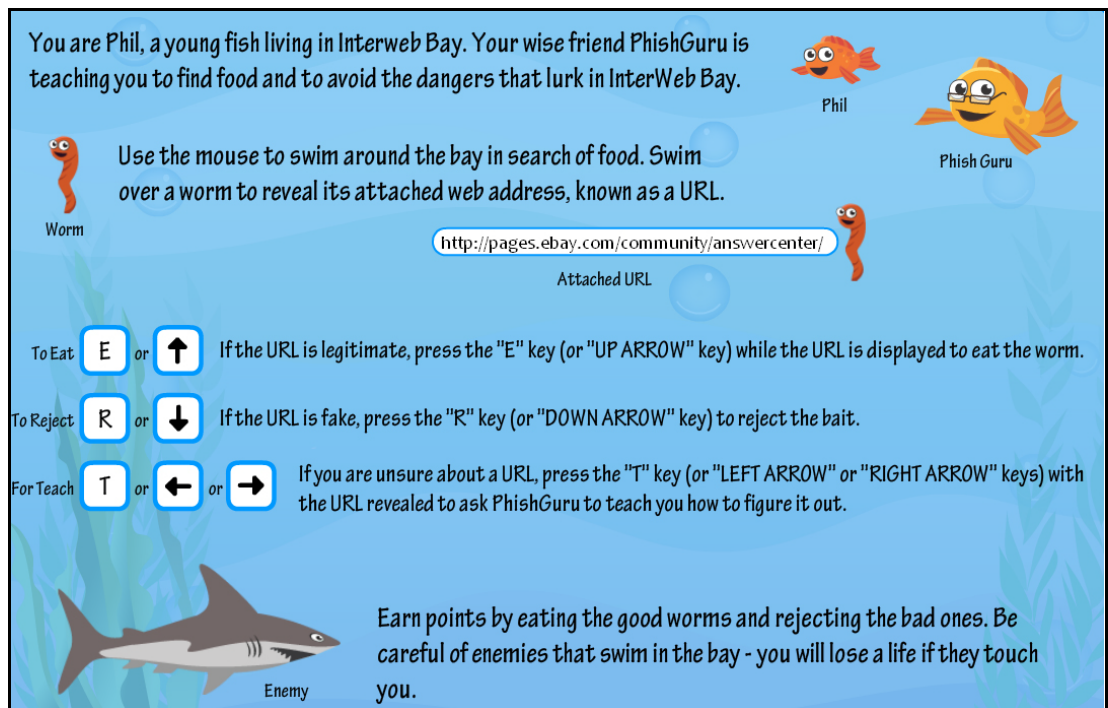


Figure 1. The instructions screen for Anti-Phishing Phil.

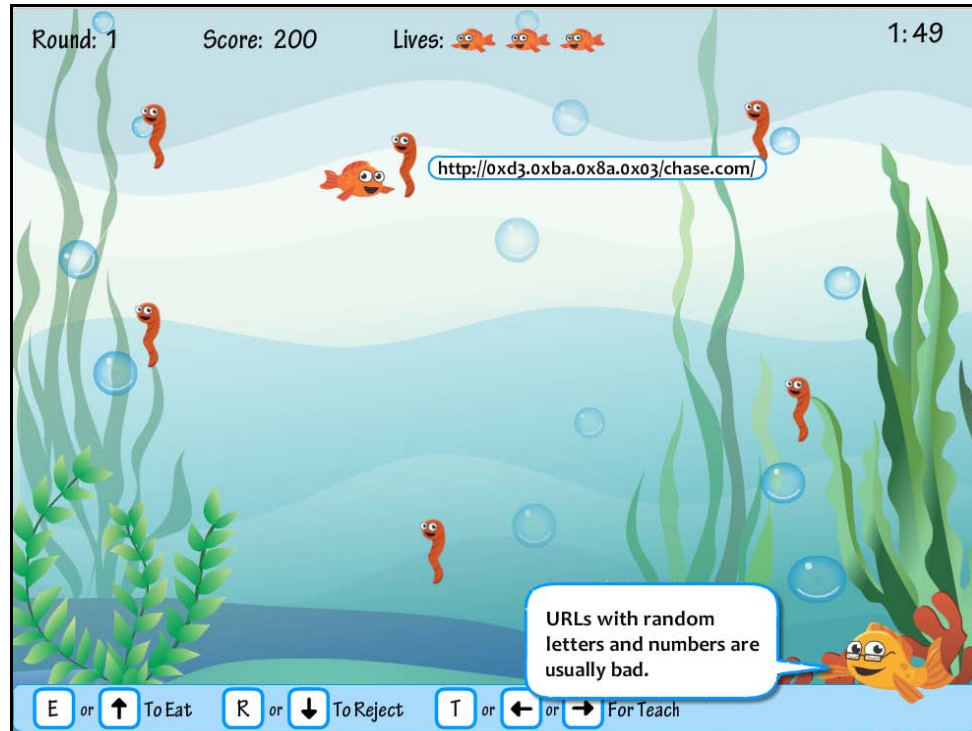


Figure 2. A screenshot of the Anti-Phishing Phil game. Players must eat good worms (ones that have good URLs) and avoid bad ones (those with phishing URLs). PhishGuru offers advice to users and helps them improve over time.

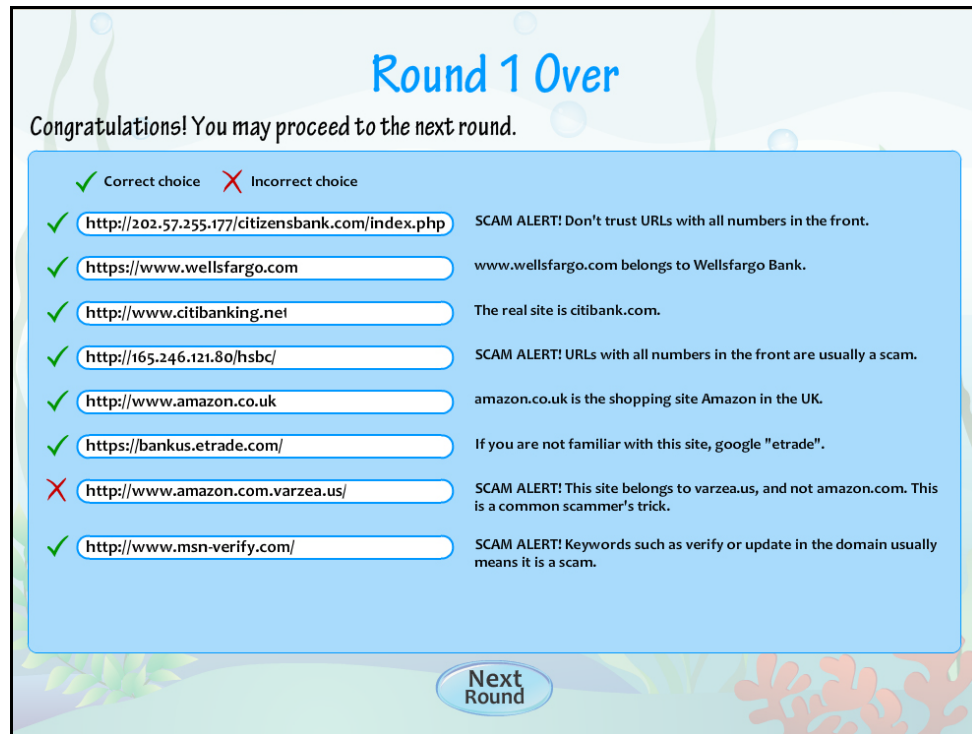


Figure 3. The round over screen of Anti-Phishing Phil provides feedback on the player's choices.

Performance Evaluation

In this section, we discuss new results from data we collected in a deployment of Anti-Phishing Phil. Earlier studies published in the *Scientific American* article “How to Foil ‘Phishing’ Scams” had already shown that the Anti-Phishing Phil game is significantly more effective than traditional training approaches (such as posters and PowerPoint presentations). Our results from this evaluation provide additional evidence that Anti-Phishing Phil training is more effective for knowledge acquisition and knowledge retention than traditional training methods.

Study design

4,517 people participated in the study over the course of two weeks. We used a between-subjects design to test two conditions: a control condition and a game condition. In the control condition, participants saw 12 websites and were asked to identify whether each URL was legitimate or fraudulent. There were 2,496 participants in the control condition.

In the game condition, participants were shown six websites before playing the game (pre-test) and another six websites after they finished playing the game (immediate post-test). To measure retention, we emailed participants seven days later and asked them to take a similar test (referred below as the “delayed post-test”). In total, we tested each participant in the game condition on 18 websites divided into three groups of three phishing websites and three legitimate websites. We randomized the order of the websites within each group, and the order in which the groups of websites were shown to each participant. In the game condition, 2,021 users completed both pre-test and immediate post-test, 674 of whom also came back one week later for the delayed post-test. In our analysis we focus on people who completed pretest, immediate post-test, and delayed post-test.

Results

Our results demonstrate that users are able to more accurately and quickly distinguish phishing websites from legitimate websites after playing the game, and that they retain knowledge learned from the game for at least one week.

We classified the game condition participants into three categories based on their pre-test scores: novice (0 - 2 URLs correct), intermediate (3 - 4 URLs correct) and expert (5 – 6 URLs correct). We had 46 participants in the novice group, 256 in intermediate, and 372 in expert.

Our results demonstrate that users are able to more accurately and quickly distinguish phishing websites from legitimate ones after playing the game.

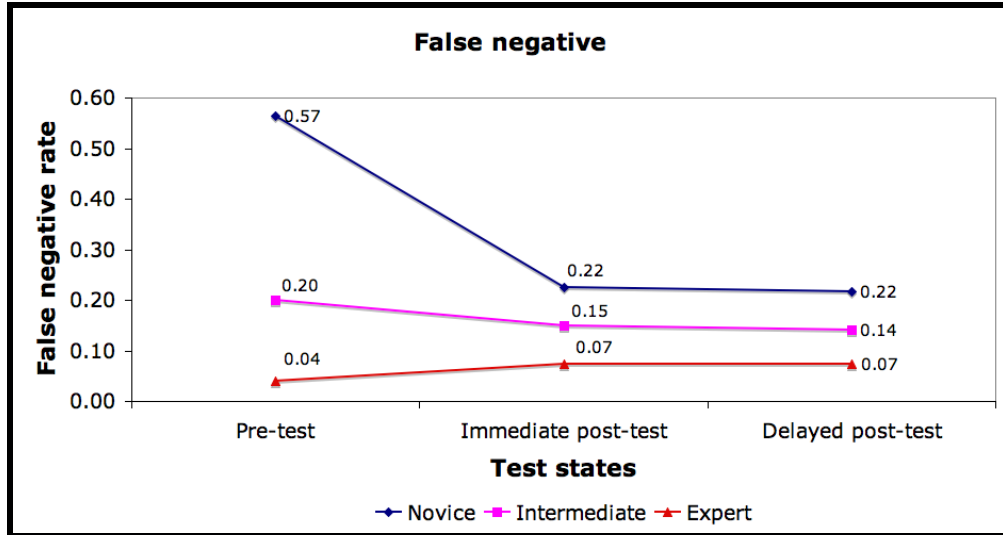


Figure 4. False negatives for users who played Anti-Phishing Phil (“game condition”). False negatives are situations where people incorrectly label a phishing site as legitimate. Novices saw the greatest reduction in false negatives, and retained what they had learned.

Novice users showed the greatest improvement, with false positive rate decreasing from 0.84 to 0.22, and false negative rate decreasing from 0.57 to 0.22.

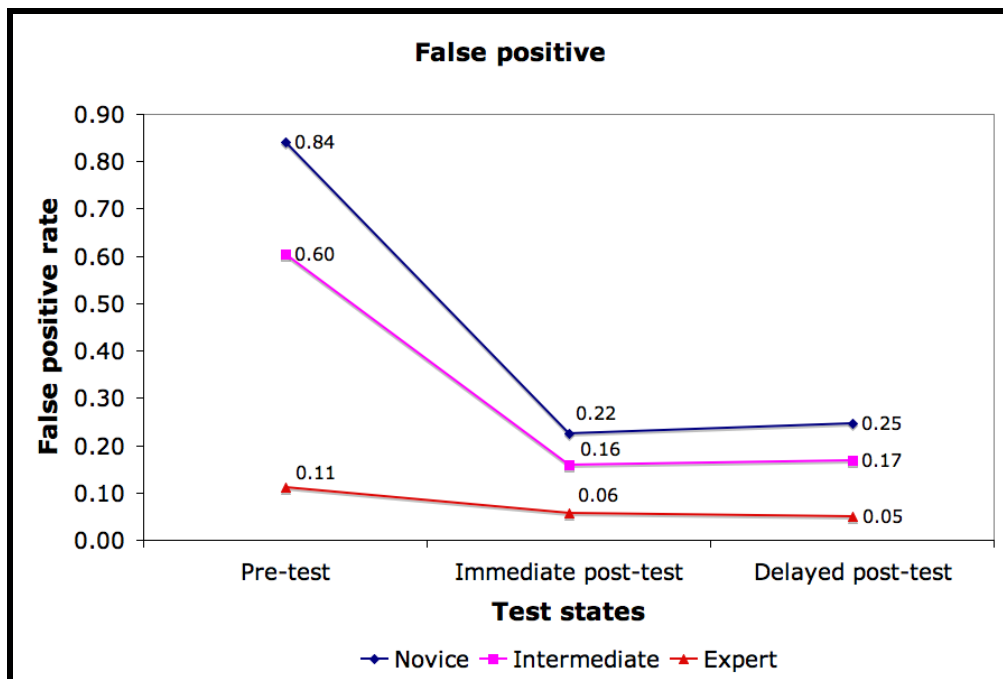


Figure 5. False positives for users who played the Anti-Phishing Phil game. False positives are situations where people incorrectly label a legitimate site as phishing. Again, novices saw the greatest improvement in reducing false positives, and retained what they had learned.

As illustrated in Figure 4, **novice users showed the greatest improvement, with the false positive rate decreasing from 0.84 to 0.22, and the false negative rate decreasing from 0.57 to 0.22.** The intermediate group also showed statistically significant improvements, although not as large as the novice group. Finally, we observed no statistically significant improvements for the expert group. Delayed

post-test scores did not decrease from immediate post-test scores; demonstrating that participants retained their knowledge after one week.

Participants were also able to determine website legitimacy more quickly after playing the game. The mean time it took users in the game group to determine a website's legitimacy before the game was 21.2 seconds. After the game, it decreased to 11.2 seconds.

Conclusions

Security education plays an important role in increasing users' alertness towards security threats. Alert users are cautious, and less likely to make mistakes that will leave them vulnerable to attack (false negatives). However, cautious users tend to misjudge non-threats as threats (false positives) unless they have learned how to distinguish between the two. Thus, good user security education should not only increase users' alertness, but also teach them how to distinguish threats from non-threats. Results from our evaluation found that people retained what they had learned for at least one week without significant degradation in performance.

Security experts agree that it is unlikely that any filter will ever be able to stop all phishing attacks. In fact many targeted spear-phishing attacks have been shown to make it past the best filters commercially available today. Effective training has to be an integral part of any organization's defense against phishing. While traditional approaches to cyber security training have proven ineffective, this and other studies have shown that well-designed training games like Anti-Phishing Phil can dramatically boost an organization's overall defense against attacks such as phishing.

About Wombat Security Technologies

With millions of users across North America, Europe, and Asia, Wombat Security Technologies is a global leader in cyber security awareness training and also offers unique anti-phishing filtering solutions. Wombat's products have been licensed for use in sectors as diverse as finance, government, and healthcare to name just a few.

For more information about Wombat Security Technologies, visit our website at www.wombatsecurity.com or email us at info@wombatsecurity.com.