

Development and Evaluation of a Low-Cost Web-Based Phishing Simulation Platform for Cybersecurity Awareness in Educational Settings

Shaan Budhabhatti¹

Received June 8, 2026

Accepted September 3, 2026

Electronic access October 15, 2026

Phishing remains a leading cause of cybersecurity incidents, and schools are vulnerable given limited budgets for security training. This study developed and evaluated a free, web-based phishing simulation platform to test whether simulation and training can reduce students' phishing click rate. The platform was built using Python and Flask on Replit's free version, with features for delivering the campaigns and tracking clicks. 173 students completed two simulated phishing campaigns, with the training delivered to Phase 1 clickers before Phase 2. Since students were tested at both stages, McNemar's test was used to analyze the paired difference in click rate rather than a two-proportion z-test. The rate at which students clicked in the simulated phishing decreased from 34.7% (60/173) in Phase 1 to 11.6% (20/173) in Phase 2, a 23.1 percentage point absolute reduction (95% CI, 14.6-31.7) and a 66.7% relative reduction. McNemar's test confirmed that the paired change was statistically significant ($\chi^2(1) = 23.0$, $p < .0001$; paired odds ratio = 4.08, 95% CI 2.22-7.48); of 60 Phase 1 clickers, 53 (88.3%) did not click again in Phase 2, while 13 new students clicked. It is possible to construct and implement a low-cost phishing simulation platform using free tools in schools. Since the same email was used in Phase 2, the stimulus difficulty was held constant. However, in the absence of a control group, the reduction cannot be entirely separated from practice effects or peer warnings. Future work should include a control group and a new email.

Keywords: phishing simulation, simulated phishing click rate, cybersecurity awareness, cybersecurity education, phishing training, educational technology, social engineering

Introduction

Phishing involves cybercriminals posing as legitimate sources to lead people into disclosing sensitive information or installing malware¹. It is still cited as the primary means of gaining initial access in reported data breaches², and industry statistics show that billions of phishing emails are sent each day. One widely quoted analysis of global email-authentication traffic estimates that about 3.4 billion phishing emails are sent every day worldwide³; the figures on breach costs come from a different source⁴.

Phishing succeeds mainly because it exploits predictable patterns in how people make decisions rather than targeting technical weaknesses. Simulation-based training is a commonly recommended component of security-awareness programs, and previous research has shown that realistic, practical experience combined with feedback enables people to identify phishing signs more reliably than passive instruction⁵⁻⁷.

The literature can be traced to PhishGuru, one of the earliest embedded training systems, which provided a brief les-

son whenever a user clicked a simulated phishing link. It was found that people retained what they had learned for at least a week, and in later work for at least 28 days⁸⁻¹⁰. A real-world application involving 515 participants repeated these retention effects beyond the laboratory setting and showed that younger users (ages 18-25) were consistently more susceptible than older users both before and after the training^{11,12}, a finding in line with the high-school-age group examined in this study, even though this study did not break down the results by student age or grade. Moreover, a game-based version of embedded training, known as Anti-Phishing Phil, demonstrated that interactive, feedback-based formats can outperform static, text-based tutorials in improving detection accuracy¹³.

More recent large-scale field studies cast doubt on the earlier optimism. A 15-month study involving over 14,000 employees revealed that voluntary embedded training (training offered but not mandatory or monitored, the same approach used in the current study) showed no consistent benefit and, under certain conditions, was linked to increased susceptibility, possibly because low completion rates limited its effectiveness¹⁴. An even larger 8-month randomized controlled trial with 19,500 healthcare workers also found only a 2-

¹ Farmington High School, Farmington, Connecticut, USA

percentage-point decrease in click-through rate due to embedded training and detected no measurable advantage from the annual awareness modules¹⁵. A scoping review of the literature on embedded training carried out in 2024 revealed that the reductions in click-through rates varied from study to study and were mostly influenced by the intensity of the simulation and the design used¹⁶. Moreover, a 2024 study found that adapting the training content to an individual's existing knowledge of phishing yielded better results than using a one-size-fits-all method¹⁷. Since the larger, controlled studies found effect sizes that were much smaller than the one observed in this study and since this project did not include a control group, the results presented below should be interpreted in light of that wider and more conservative body of literature, a point that is made again in the Discussion and Limitations section.

There has been another area of research concerned with why click rates differ so much across studies and even between separate phishing emails. The Phish Scale project, carried out by NIST, found that click rates are strongly influenced by how well the premise of a phishing email fits a person's usual situation and suggested a standardized method for rating email difficulty so that comparisons of click rates across studies and organizations become more meaningful^{18–20}. This point is directly relevant to the design of the present study: since both Phase 1 and Phase 2 used the same email (see Methods Section), the stimulus difficulty was held constant across phases, thereby eliminating this source of confusion, even though other confounding factors remain. A comprehensive review of phishing user studies over sixteen years also found that reported click rates and the sizes of training effects vary greatly across study populations, designs, and phishing templates, making it difficult to carry out direct comparisons between studies such as this one²¹. With regard to how to combine information with hands-on practice, a large field experiment involving more than 10,000 employees found that both giving people information about phishing and providing them with a simulated experience of a phishing attempt separately led to a reduction in the rate at which credentials were disclosed, although combining the two did not result in any additional benefit²². This suggests that the combination of simulation and instruction in a program may be less important than ensuring that at least one active, hands-on component is included, which aligns with this platform's approach of combining simulated email with a self-paced lesson.

In K-12 environments, the body of research is much thinner than it is in corporate settings. A systematic review of 24 academic studies on cybersecurity education in K-12 schools worldwide found that most of the published work focused on general digital citizenship and online safety rather than phishing or email security, and that secondary-school-aged students were studied far more frequently than younger children²³. A related review of cybersecurity-awareness interventions di-

rected at children also reached a similar conclusion: few of the interventions were assessed using outcome measures beyond immediate post-test knowledge, and it called for more studies that measure behavior change rather than just self-reported awareness²⁴. Moreover, a broader review of cybersecurity education research presented at computing education conferences found that empirical evaluation of student learning outcomes remains uncommon compared with the number of papers that describe tools and curricula in the field²⁵. The present study is one of the few to provide an outcome-based, behavior-level (as opposed to self-reported) evaluation of phishing-specific training carried out with a high-school population, although, in line with the calls made in these reviews for more rigorous evaluation designs, it too shares the field's common shortcoming of not having a control group. One of the few existing studies that had directly tested high school students used signal detection theory with a small number of students and staff and found an overconfidence bias in the self-rated ability to detect phishing regardless of technical background²⁶, a result which is consistent with the present study's decision to measure objective click behavior rather than self-reported detection ability.

Regarding the tools available, the most widely used free alternative to commercial phishing-simulation platforms is Gophish, an open-source, self-hosted framework that allows a security team to create email templates, run campaigns, and monitor opens and clicks via a web dashboard²⁷. Although Gophish and other similar open-source frameworks offer strong campaign and tracking facilities, by design they come without any built-in educational material, deliverability assistance, or user-facing training resources; administrators have to either produce the lesson content themselves or obtain it from outside sources, and also need to handle the landing pages as well as the deliverability and allowlisting work, typically requiring experience in server administration in order to set up and keep the tool running. The platform developed for this project is aimed at the same free-tier, resource-constrained market segment as Gophish but differs in scope in that it includes an integrated, self-paced educational module and has a built-in administrator dashboard that does not require any knowledge of the command line or server administration, with the aim of reducing the technical difficulties for a school's IT department when it carries out its first phishing-awareness pilot, rather than trying to match the scale or level of customization offered by dedicated red-team tools.

This project aims to narrow that usability gap by building a simple, free, web-based phishing-simulation tool using only Replit's free tier, to give schools and other resource-limited organizations a practical way to run realistic phishing simulations and build hands-on threat-recognition skills.

Methods

Research Design

The entire application was developed on Replit using Python and Flask, with no paid services required. To compare click rates from simulated phishing before and after an awareness-training intervention among a group of high school students, a single-group pre-post design was used. A control group was not included in this design; the implications of this decision are discussed in the Limitations section.

Platform Architecture and Core Components

The platform has five main components. It includes an administrator dashboard that provides a summary of the platform's statistics and grants access to the other modules. There is a user management system that lets administrators enter their targets' email addresses and assign participants to campaigns. The library of customizable phishing email templates features typical real-world social engineering techniques, such as creating a sense of urgency, impersonating authority, and using curiosity-driven hooks. A campaign management interface enables administrators to select recipients and templates before sending the simulations. Finally, a response-tracking system records each campaign's link clicks and email opens.

The emails were sent using a tiered delivery service in which the system first tried to deliver them via the Resend API, switched to SendGrid if the Resend API was not available, and then used direct SMTP (or entered a console-logging development mode) if neither of the API keys had been configured; this ensured that the tool could be used on Replit's free tier without needing a paid transactional email service. For click tracking, a unique and randomly generated identifier (a UUID4 token) was embedded in each recipient's personal link; when the link was accessed, a Flask route would search for the relevant student record, record the click (including a timestamp from the server to calculate the time-to-click and details about the browser and device) on only the first visit in order to prevent double-counting, and then display an immediate feedback page indicating that the email had been a simulated phishing exercise, together with an optional link to the full educational module described below.

Practical Deployment Considerations

Since the platform is designed for schools with limited resources, it is worth noting a few practical points regarding its deployment. Phishing emails simulated and sent from the free-tier version of the platform might be flagged by the institution's spam filters, which could affect click-rate figures if some students do not receive the message. It is advisable for administrators to work with IT staff to add the sending domain

to the allowed list when running a campaign. The free-tier hosting generally includes limits on the amount of outbound email that can be sent and on guaranteed uptime, meaning that for a large student population, sending emails in batches over time may be necessary. In order for the tool to remain useful throughout several school years, it is necessary to carry out regular maintenance: the templates have to be updated to take into account current phishing techniques, the simulated sender domains have to be rotated, and monitoring for possible platform downtime must be carried out. If the platform is to be used beyond a single school (e.g. across an entire district), it will likely be necessary to move off the free tier when the number of users or email volume exceeds the platform's limits. More generally, resource-constrained organizations can also draw on general cybersecurity advice for small businesses to implement practical, low-cost security measures that complement a phishing-simulation program²⁸.

Participants or Sample

A total of 173 students from the participating high school took part in both phases of the study.

Variables and Measurements

The independent variable was the stage at which the study took place: Phase 1 (this being the stage before the participants had received any training) and Phase 2 (this being the stage that followed the participants' having undergone the educational module). The dependent variable was a simple yes-or-no answer stating whether or not a specific student clicked on the link that was included in the simulated phishing email, an action which was automatically logged by the platform's click-tracking system and which the students did not report themselves. The system also automatically collected two additional measures for descriptive purposes, namely whether the email was opened and the time it took the students to click the link, as recorded by the server. These secondary measures were not analyzed in this study.

Data Collection and Procedure

Two phishing simulation campaigns were conducted. In Phase 1, the students were sent a simulated phishing email that employed urgency, impersonation, and a sense of curiosity to get them to click; the system recorded whether each student clicked the link in the email. Those students who had clicked in Phase 1 were then given specific educational material on how to recognize phishing signs, verify the sender's identity, check links before clicking, and follow safe email practices.

The educational module was provided individually and on a self-paced basis via a webpage (not through a live or classroom session); each student who moved from the information

page onto the lesson itself received a self-contained set of materials covering four areas: an explanation of what phishing is and its common forms (email phishing, smishing, vishing, and website spoofing), a checklist of red flags to watch out for (such as urgent or threatening language, general greetings, sender addresses that raise suspicion, poor grammar, unexpected attachments or links, and offers which sound too good to be true), best practices to adhere to before clicking on any link (which include hovering to see where the link leads, verifying the sender, checking if the message was expected, and using official channels or IT support if you are unsure), and statistics on the actual real-world impact of phishing. The module concluded with a short interactive quiz consisting of two questions and links to external resources (the CISA and FBI guidance pages). Since it was a single webpage allowing for self-paced learning rather than a timed session, no fixed length was set or recorded.

In Phase 2, the same simulated phishing email used in Phase 1 was sent to all 173 students one week later, regardless of their Phase 1 result, to see whether their clicking behavior changed as a result of the training. The use of the same email in both phases means that the comparison being made is direct and fair when it comes to click rate, rather than one email being more convincing than the other making the comparison biased; the drawback is that any improvement might simply be due to the students recognizing that particular email rather than them developing a general ability to detect phishing emails, a point which is covered in the Limitations section.

Data Analysis

The 173 students who were evaluated in both Phase 1 and Phase 2 have click outcomes that constitute paired (matched) observations rather than independent samples. Since the paired change in click rate was to be assessed, McNemar's test for correlated proportions²⁹ was used rather than a two-proportion z-test, which assumes independent groups. Rather than relying on the raw marginal totals, McNemar's test examines the two groups of students whose behavior changed between phases (clicked-then-not-clicked vs. not-clicked-then-clicked). Along with the test statistic, the 95% confidence intervals for the paired difference in proportions and the paired odds ratio are provided to indicate the effect size. A result was considered statistically significant at $\alpha = 0.05$.

Ethical Considerations

The testing was carried out with the participants' informed consent, and no passwords, credentials, or any other personal information was collected. The students' responses were analyzed only in aggregate. Prior approval for the study was obtained from the IT director of the relevant school district, who

has authority to approve the use of the district's email systems and student accounts in this situation.

A formal institutional review board process was not undertaken; instead, the study was reviewed and approved by both the school principal and the district's IT director. This method of approval was considered suitable for a study at the high school level since the activity complied with the school's minimal risk, routine IT-security activity: no credentials, passwords, or other personal information were gathered at any stage, the participants used their own school email accounts through the district's own systems, and the outcome data were only analyzed and reported in summary form. The approval given by the principal included the educational and student welfare areas, while that provided by the IT director concerned permission to use the district's email systems for the exercise, thus ensuring the appropriate degree of administrative oversight for this kind of educational security exercise.

Results

173 students took part in both stages of the phishing simulation. In Phase 1, 60 students (34.7%) clicked on the simulated phishing link; in Phase 2, 20 students (11.6%) did so, which represents an absolute reduction of 23.1 percentage points (95% CI, 14.6 to 31.7 percentage points) and a 66.7% relative reduction in the click rate. Since the same students were evaluated at both time points, McNemar's test for paired proportions was used to assess the change rather than a two-proportion z-test. Of the 173 students, 53 clicked in Phase 1 but not in Phase 2, and 13 did not click in Phase 1 but did in Phase 2; these 66 discordant pairs form the basis of the McNemar comparison. The paired difference was found to be statistically significant (McNemar's $\chi^2(1) = 23.0$, $p < .0001$; exact binomial $p < .0001$), with a paired odds ratio of 4.08 (95% CI, 2.22 to 7.48), showing that students were about four times more likely to go from clicking to not clicking than the reverse.

Out of the 60 students who clicked in Phase 1, 53 (88.3%) did not click in Phase 2 and 7 clicked in both phases. Of the 17 students who clicked in Phase 1 and completed the targeted post-click training, only 3 clicked again in Phase 2, which means there was an 82.4% reduction in repeat clicks among that subgroup.

Of the 43 Phase 1 clickers who did not finish the educational module, the platform only counted a student as having completed the training when their browser had loaded the full lesson page, and it seems that a number of these students had stopped at the immediate reveal page (this page shows a short, one-screen summary of red flags) and had not gone on to complete the full lesson or had opened the lesson page only to close it before they could make use of the content. It is important to note this point when interpreting the subgroup re-

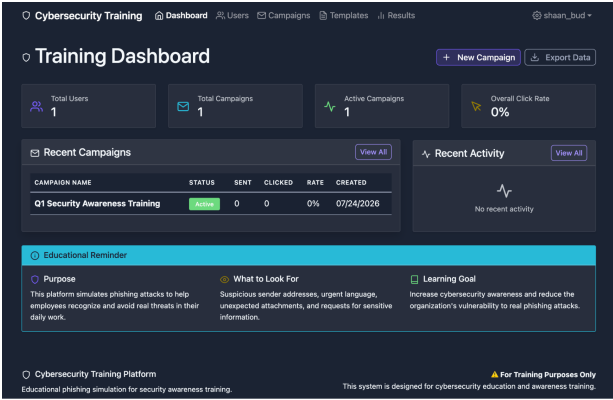


Figure. 1 Main dashboard overview

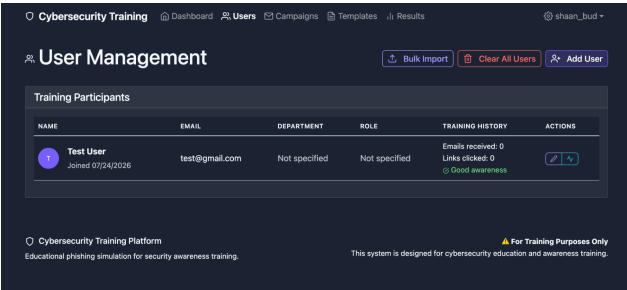


Figure. 2 User management interface

sult mentioned above: the 82.4% reduction among the trained students refers to the 17 students who engaged with the full training, not all 60 of the Phase-1 clickers. A summary of these results is given in Table 1.

Table 1 | Phishing Simulation Outcomes

Metric	Phase 1	Phase 2
Participants	173	173
Students Clicking Phishing Link	60 (34.7%)	20 (11.6%)
Students Not Clicking	113 (65.3%)	153 (88.4%)
Overall Click Rate Reduction	66.7% relative / 23.1 pp absolute	

Table 2 | Individual Behavior Changes Between Simulations

Phase 1 → Phase 2	Students
Clicked → Clicked	7
Clicked → Did Not Click	53
Did Not Click → Clicked	13
Did Not Click → Did Not Click	100

Figures 1–5 show the dashboard, user-management interface, phishing-template library, campaign interface, and an example simulated phishing email. Figure 6 compares the simulated-phishing click rate between Phase 1 and Phase 2.

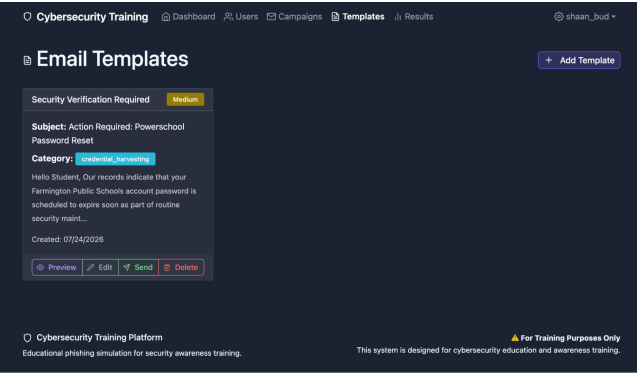


Figure. 3 Phishing email templates library

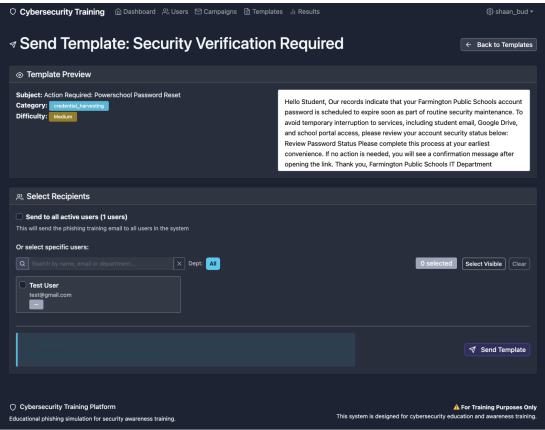


Figure. 4 Campaign sending page

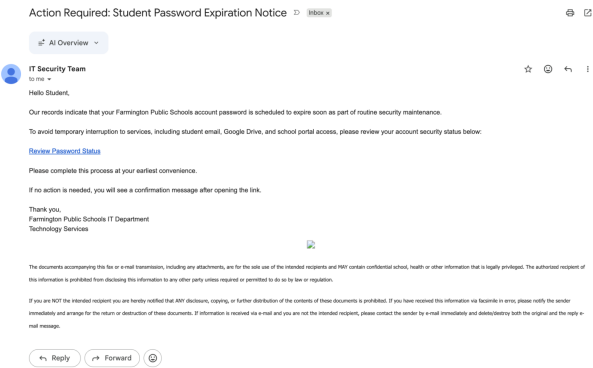


Figure. 5 Example simulated phishing email received

Discussion

Restatement of Key Findings

A large decrease in the rate at which people clicked on simulated phishing messages was observed when repeated phishing

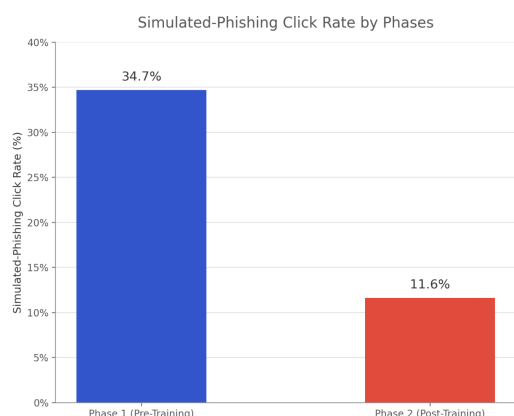


Figure. 6 Simulated-phishing click rate, Phase 1 vs. Phase 2

simulations were combined with focused awareness training, the click rate falling from 34.7% to 11.6%, a 66.7% relative reduction that was confirmed by a paired (McNemar's) test. Furthermore, 88.3% of the students who had clicked in Phase 1 did not click in Phase 2.

Students Who Clicked Only in Phase 2

Table 2 also indicates that of the 113 students who did not click on the email in Phase 1, 13 later clicked on the same email in Phase 2. This group should be mentioned specifically: even though there was an overall improvement, a portion of the students who had not been affected before either remained or became susceptible to the same simulated email. Possible reasons for this could be natural variations in people's level of attention or caution on a particular day, the email being forwarded or discussed among students in a way that unintentionally increased its apparent legitimacy, or the students who did not click in Phase 1 never having received any awareness material (since only those who clicked in Phase 1 were sent into the training program). That means that if a training program is aimed only at students who fail the first test, it will not reach a large number of students who are still at risk, which in turn supports the view of giving all participants basic awareness material rather than restricting it to those who fail on the first attempt. This pattern is also in line with broader research on phishing susceptibility, which shows that individual differences, rather than training alone, determine which individuals remain vulnerable in a given simulation³⁰.

Implications and Connection to Objectives

The findings are generally in line with previous studies which have indicated that hands-on, simulation-based practice is

more effective than passive materials alone for improving people's ability to recognize phishing attempts^{7,31}. It is nevertheless essential to be clear about what was measured in this study: participants' actions in response to two simulated emails, not a validated assessment of overall cybersecurity awareness or actual phishing risk. The main outcome is therefore referred to throughout this paper as the simulated phishing click rate rather than 'phishing risk' or 'awareness' in the general sense. A more comprehensive evaluation of awareness would ideally also include phishing-reporting rates, time-to-click, students' ability to spot red flags in emails they have not seen before, the extent to which these skills are retained after several weeks or months, and the rate of false-positive suspicion of legitimate email, all of which were not measured in this study.

Recommendations

These results are consistent with the recommendations issued by NIST and CISA that security awareness training should be a key element in reducing organizational risk^{5,6}. Since this study employed a single-group pre-post design and used the same email in both phases, further research should include two changes: a waitlist or delayed-training control group, consisting of a group of students who do not receive the training between Phase 1 and Phase 2, and a new phishing email that has not been seen before at the follow-up instead of using the same one, in order to confirm that the improvement is due to a transferable ability to detect phishing emails rather than simply recognizing that particular message. This would also help with generalizability of the findings if the study were carried out using larger and more diverse samples from a number of schools and if the follow-up period were extended in order to see how well the skills are retained.

Limitations

The study has several limitations. It was conducted at a single high school and involved a sample of 173 students. Since the students were not randomly allocated to the different training conditions, any claims regarding the effect of the training should be made with caution. There was no control group and all the participants went through the Phase 1 simulation before taking Phase 2, so some of the improvement that was observed might be due to becoming familiar with the concept of a phishing test or to receiving informal peer-to-peer warnings following the first campaign, rather than as a result of the training material itself. This is not just a possibility: a large, multi-center study of six US healthcare organizations found that repeated phishing campaigns were linked to lower odds of clicking on a later simulation even when no formal training was provided with each individual campaign, which supports

the idea that repeated exposure on its own, independent of the training content, could have been responsible for part of the improvement seen in this study³². Phase 2 used the same simulated email as in Phase 1, which prevents the results from being affected by a change in the difficulty of the stimulus, but it also means that the study cannot tell the difference between a real, transferable skill in detecting phishing emails and simple recognition or memory of that particular email; in order to confirm transfer, a follow-up study using a new and previously unseen phishing email would be required. As the sample came from a single suburban high school, the findings may not be applicable to other age groups, different types of schools, or to different starting levels of awareness about cybersecurity; universities, workplaces, and schools that have different norms regarding the use of technology might display different click-rate patterns and might react in different ways to the same training program. For comparison, a large observational study at a university found that susceptibility varied according to class year, previous cybersecurity training, and the amount of time spent using computers³³, showing how demographic and institutional factors that differ between a high school and a university environment could cause the click-rate patterns reported here to change in either direction.

Closing Thought

The project demonstrates that it is possible to construct a practical and inexpensive phishing-simulation platform using free tools and deploy it in a school environment, and that the rate at which people clicked on the simulated phishing messages decreased considerably between the two testing periods. However, it will be necessary to carry out a controlled follow-up study if we are to find out whether it was the training itself, not repeated exposure or informal warnings from peers, that caused this change; until such a study is done, the results should be regarded as an encouraging, low-cost pilot rather than as confirmed evidence of a causal effect of training.

Acknowledgments

The author thanks the Farmington Public Schools Technology Department for supporting the implementation of the phishing awareness pilot program. The author also thanks Mr. Jeff Daddio for project mentorship, Mr. Matt Ross for technical guidance and review of ethical implementation procedures, and Mr. Russell Crist for being the Sponsor of this research study. The author also thanks Dr. Vahid Behzadan from University of New Haven for guidance through my research work.

References

- 1 Federal Trade Commission. How to recognize and avoid phishing scams. 2024. <https://consumer.ftc.gov/articles/how-recognize-avoid-phishing-scams>.
- 2 Verizon. 2024 data breach investigations report. Verizon, 2024.
- 3 Valimail. Spring 2019 Email Fraud Landscape. Valimail, San Francisco, CA. June 11, 2019.
- 4 IBM Security. Cost of a data breach report 2024. 2024. <https://www.ibm.com/reports/data-breach>.
- 5 National Institute of Standards and Technology. Building an information technology security awareness and training program. NIST Special Publication 800-50. National Institute of Standards and Technology, 2003.
- 6 Cybersecurity and Infrastructure Security Agency. Avoiding Social Engineering and Phishing Attacks, 2021 <https://www.cisa.gov/news-events/news/avoiding-social-engineering-and-phishing-attacks>.
- 7 D. Jampen, G. Gür, T. Sutter, B. Tellenbach. Don't click: towards an effective anti-phishing training. A comparative literature review. *Human-Centric Computing and Information Sciences*. Vol. 10, Article 33, 2020. <https://doi.org/10.1186/s13673-020-00237-7>.
- 8 P. Kumaraguru, Y. Rhee, A. Acquisti, L.F. Cranor, J. Hong, E. Nunge. Protecting people from phishing: the design and evaluation of an embedded training email system. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '07)*. pg. 905–914, 2007.
- 9 P. Kumaraguru, Y. Rhee, S. Sheng, S. Hasan, A. Acquisti, L.F. Cranor, J. Hong. Getting users to pay attention to anti-phishing education: evaluation of retention and transfer. *Proceedings of the Anti-Phishing Working Group's 2nd Annual eCrime Researchers Summit*. pg. 70–81, 2007.
- 10 P. Kumaraguru, S. Sheng, A. Acquisti, L.F. Cranor, J. Hong. Teaching Johnny not to fall for phish. *ACM Transactions on Internet Technology*. Vol. 10, No. 2, Article 7, 2010.
- 11 P. Kumaraguru, J. Cranshaw, A. Acquisti, L.F. Cranor, J. Hong, M.A. Blair, T. Pham. School of phish: a real-world evaluation of anti-phishing training. *Proceedings of the 5th Symposium on Usable Privacy and Security (SOUPS '09)*. 2009.
- 12 S. Sheng, M. Holbrook, P. Kumaraguru, L.F. Cranor, J. Downs. Who falls for phish? A demographic analysis of phishing susceptibility and effectiveness of interventions. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10)*. pg. 373–382, 2010.
- 13 S. Sheng, B. Magnien, P. Kumaraguru, A. Acquisti, L.F. Cranor, J. Hong, E. Nunge. Anti-Phishing Phil: the design and evaluation of a game that teaches people not to fall for phish. *Proceedings of the 3rd Symposium on Usable Privacy and Security (SOUPS '07)*. pg. 88–99, 2007.
- 14 D. Lain, K. Kostainen, S. Capkun. Phishing in Organizations: Findings from a Large-Scale and Long-Term Study. *2022 IEEE Symposium on Security and Privacy (S&P)*. 2022.
- 15 G. Ho, A. Mirian, E. Luo, K. Tong, E. Lee, L. Liu, C.A. Longhurst, C. Dameff, S. Savage, G.M. Voelker. Understanding the Efficacy of Phishing Training in Practice. *2025 IEEE Symposium on Security and Privacy (S&P)*. 2025.
- 16 N. Marshall, D. Sturman, J.C. Auton. Exploring the evidence for email phishing training: A scoping review. *Computers & Security*. Vol. 139, 2024.
- 17 L. Schöni, V. Carles, M. Strohmeier, P. Mayer, V. Zimmermann. You Know What? Evaluation of a Personalised Phishing Training Based on Users' Phishing Knowledge and Detection Skills. *Proceedings of the 2024 European Symposium on Usable Security (EuroUSEC '24)*. 2024.
- 18 M. Steves, K. Greene, M. Theofanos. A Phish Scale: Rating Human Phishing Message Detection Difficulty. *Workshop on Usable Security and Privacy (USEC) at NDSS*. 2019.
- 19 M. Steves, K. Greene, M. Theofanos. Categorizing human phishing difficulty: a Phish Scale. *Journal of Cybersecurity*. Vol. 6, No. 1, tyaa009, 2020.

- 20 K.K. Greene, M. Steves, M. Theofanos, J. Kostick. User Context: An Explanatory Variable in Phishing Susceptibility. *Workshop on Usable Security and Privacy (USEC) at NDSS*. 2018.
- 21 S. Baki, R. Verma. Sixteen Years of Phishing User Studies: What Have We Learned? *IEEE Transactions on Dependable and Secure Computing*. Vol. 20, No. 2, pg. 1200–1212, 2023.
- 22 A. Baillon, J. de Bruin, A. Emirmahmutoglu, E. van de Veer, B. van Dijk. Informing, simulating experience, or both: A field experiment on phishing risks. *PLOS ONE*. Vol. 14, No. 12, e0224216, 2019.
- 23 A. Ibrahim, M. McKee, L.F. Sikos, N.F. Johnson. A Systematic Review of K-12 Cybersecurity Education Around the World. *IEEE Access*. Vol. 12, pg. 59726–59738, 2024.
- 24 F. Quayyum, D.S. Cruzes, L. Jaccheri. Cybersecurity awareness for children: A systematic literature review. *International Journal of Child-Computer Interaction*. Vol. 30, 100343, 2021.
- 25 V. Švábenský, J. Vykopal, P. Čeleda. What are cybersecurity education papers about? A systematic literature review of SIGCSE and ITiCSE conferences. *Proceedings of the 51st ACM Technical Symposium on Computer Science Education (SIGCSE '20)*. pg. 2–8, 2020.
- 26 P. Unchit, S. Das, A. Kim, L.J. Camp. Quantifying Susceptibility to Spear Phishing in a High School Environment Using Signal Detection Theory. *HCI International 2020, Lecture Notes in Computer Science*. Springer, 2020.
- 27 Gophish. Gophish - Open Source Phishing Framework. Accessed 2026. <https://getgophish.com/>.
- 28 U.S. Small Business Administration. Cybersecurity. 2024. <https://www.sba.gov/business-guide/manage-your-business/strengthen-your-cybersecurity>.
- 29 Q. McNemar. Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika*. Vol. 12, No. 2, pg. 153–157, 1947.
- 30 N. Beu, A. Jayatilaka, M. Zahedi, M.A. Babar, L. Hartley, W. Lewinsmith, I. Baetu. Falling for phishing attempts: An investigation of individual differences that are associated with behavior in a naturalistic phishing simulation. *Computers & Security*. Vol. 131, 103313, 2023.
- 31 R. Wash, M. M. Cooper. Who provides phishing training? Facts, stories, and people like me. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. Paper 492, pg. 1–12, 2018. <https://doi.org/10.1145/3173574.3174066>.
- 32 W. J. Gordon, A. Wright, R. Aiyagari, L. Corbo, R.J. Glynn, J. Kadakia, C. Kufahl, C. Mazzone, E. Noga, M. Parkulo, B. Sanford, A. Scheib, A. Landman. Assessment of Employee Susceptibility to Phishing Attacks at US Health Care Institutions. *JAMA Network Open*. Vol. 2, No. 3, e190393, 2019.
- 33 A. Diaz, A.T. Sherman, A. Joshi. Phishing in an academic community: A study of user susceptibility and behavior. *Cryptologia*. Vol. 44, No. 1, pg. 53-67, 2020.