Phishing Simulation

Last Updated: 27 April 2020

How to Setup Phishing Simulations for your Clinic or Lab.

Introduction to Phishing

Phishing is on the rise in both severity and amount of attacks. According to the 2019 Data Breach Investigations Report conducted by Verizon, phishing is the top threat action used in 32% of all successful data breaches, with attacks involving social engineering and malware [1].

While there is a wide variety of how to do phishing guides, ranging from advocacy groups such as EFF to commercial products, available on the web, we were unable to find any publicly accessible phishing policies. We can draw some insights from research that has been conducted in the space of evaluating the effectiveness of phishing training as well as the CPHS IRB process as a framework for ethical considerations [2, 3].

Citizen Clinic Phishing Policy

Existing Phishing Policies

UC Berkeley has an Information Security and Policy group that reportedly has an existing phishing policy and does conduct phishing training for the university. We have reached out to security@berkeley.edu, but have yet to receive a response.

Their resources can be found at:

https://security.berkeley.edu/resources/phishing

And they have a phishing quiz that is currently in use:

https://phishingquiz.withgoogle.com

For conducting phishing training at UC Berkeley, one is supposed to contact security@berkeley.edu for questions and review of their procedures.

IRB for Phishing Policy

Though much of the IRB policies are more research specific, we can leverage its frameworks for thinking about how to protect the rights and welfare of human subjects. The section on Assessment of Risks and Benefits as laid out in the Belmont Report [4] on which ethical research guidelines are based is especially pertinent.

The basic ethical principles the report asks us to follow are:

1. Respect for persons - individuals should be treated as autonomous agents and that persons with diminished autonomy (e.g. underserved populations) are entitled to protection
2. Beneficence - actions should (1) do not harm and (2) maximize possible benefits and minimize possible harms
3. Justice - consider who ought to receive the benefits and bear its burdens (this is more relevant for research, such as in choosing what population to involve)

Of the three, the most important in our assessment is the principle of beneficence in assessing the risk versus benefit of a good policy. The report uses the following definitions:

- “Benefit” as non-probabilistic positive value related to health or welfare; this may include contribution to generalizable knowledge and direct benefit to participant(s)
- “Risk” as possibility that harm may occur; including both chance (probability of experiencing harm) and severity (magnitude of such harm)

We recommend when developing policies to have at least one author or reviewer having taken the Group 2 Human Subjects Training: Social and

Behavioral Research Investigators Course, available to all UC Berkeley affiliates through CPHS and CITI, available at https://cphs.berkeley.edu/training.html.

Relevant sections from the CPHS training are as follows:

- Assessing risk and privacy
    - Assessing Risk (ID 503)
    - Informed Consent (ID 504)
    - Privacy and Confidentiality (ID 505)
- Given clinic work with vulnerable populations
    - Vulnerable Subjects (ID 483)
    - Unanticipated Problems and Reporting Requirements in Social and Behavioral Research (ID 14928)
- Given work with groups in other nations
    - Cultural Competence in Research (ID 15166)
    - International Research (ID 509)

Considerations for Constructing a Policy

When forming a phishing policy, we would recommend considering the following questions:

- What benefits would your phishing training and simulation bring to your organization?
- What is the goal of the phishing training or simulation?
    - For my organization?
        - To assess the risk of phishing to my organization
        - To evaluate the effectiveness of phishing trainings (e.g. they do not click 1 week, 1 month, or 6 months in the future)
    - For the participants?
        - To introduce and inform participants about phishing
        - To remind participants about phishing

- When should participants be informed (e.g. ahead of time,
- What should be the size of my phishing training (e.g. select individuals, all individuals, by team, by department, across the organization)?
- Where should training material be placed (e.g. in a live workshop ahead of time, in the phishing email, in training afterwards)?
- When do you inform participants they may be sent
- What emotional duress or stress might your training materials (emails or presentation) cause your participants?
- How long should phishing simulations last?

A policy for a phishing simulation should include the following:

- Goal of outcome
- Target group definition
- Set duration
- Plan for disclosure to participants
    - Ahead of time: consent getting, alert (potentially with opt out)
    - Afterwards: debrief, share outcomes, additional training
- Review of phishing materials that evaluates potential harms (e.g. emotional duress or stress) against actual benefits (for organization or participant)

If our goal is to conduct generalizable research, we should additionally consider:

- Who should be included in the sample? Who should we exclude (e.g. undue burden)?
- Can this take place in a lab or field setting?
- What is my outcome variable?
- Should I get informed consent ahead of time?
- How should I debrief the participants?

Table 1, shown below, is from a 2018 CHI paper provides an overview of previous research design for phishing. Since they were interested in understanding how people who respond to different kinds of training material

when shown after they have clicked on the phishing link, they did not get informed consent ahead of time. In this situation the debrief becomes very important as is having an IRB already in place [2].

Authors	Publication	Location	N	Click	Outcome	Repeat
Ferguson	EDUCAUSE 2005	Field	512	80%	Click	None
Wu et al.	CHI 2006	Lab	30	52%	Click	None
Jagatic et al.	CACM 2007	Field	921	72% / 16%	Info	None
Kumaraguru et al.	CHI 2007	Lab	30	90%	Click	None
Kumaraguru et al.	eCrime 2007	Lab	42	90%	Click	7 days
Kumaraguru et al.	eCrime 2008	Field	311	42%/39%	Both	2 and 7 days
Kumaraguru et al.	SOUPS 2009	Field	515	52%/51%/45%	Both	28 days
Caputo et al.	IEEE S&P Mag, 2014	Field	1,359	60%	Click	7 months
Wash et al	CHI 2018	Field	1,945	11.7%	Click	2,7 and 42 days

Table 1. Summary of previous research about Phishing clicking and training. Field studies were conducted with unsuspecting subjects, and thus likely represent more accurate estimates of click rates. Click rates in this table are before subjects received training (if the study included training); if multiple rates are listed, the original paper reported separate rates for different conditions. Most studies considered a 'click' to be falling for the scam, but a few considered entering personal information ('info') into the subsequent webpage.

Also interesting is that phishing studies [2, 3] referred to timeframes for phishing as short (2-days), medium (7-days) and long (30-days). Outcome variables refer to if the user clicked on the phishing link (click), if they shared personal information (info), or click and information (both).

Please note that any phishing done for generalizable research purposes would need to submit a protocol for IRB review. You can find more information at https://cphs.berkeley.edu/.

Limitations and Future Work

Due to time constraints, there is much more work that can be done, including but not limited to conducting expert interviews, getting existing phishing policies currently in practice, and getting feedback from practitioners and past phishing simulation participants

Further questions to consider:

- Is the risk versus benefit assessment approach of the IRB a good approach to consider? Is there another framework that can be used?
- How can we assess what is acceptable "harm" or "risk" in a phishing policy?
- What are some normative constraints to consider (e.g. culture of organization, size of campaign, frequency of phishing attempts)?

- Who are all of the stakeholders who should be considered with regard to a phishing policy (e.g. managers, employees, HR, legal team)?

References

1. 2019 Data Breach Investigations Report. Verizon, 2019. https://enterprise.verizon.com/resources/reports/dbir/
2. Rick Wash and Molly M. Cooper. 2018. Who Provides Phishing Training?: Facts, Stories, and People Like Me. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18). ACM, New York, NY, USA, Paper 492, 12 pages. DOI: https://doi.org/10.1145/3173574.3174066
3. Kumaraguru, Ponnurangam & Sheng, Steve & Acquisti, Alessandro & Cranor, Lorrie & Hong, Jason. (2008). Lessons From a Real World Evaluation of Anti-Phishing Training. eCrime Researchers Summit, eCrime 2008. 1 - 12. 10.1109/ECRIME.2008.4696970.
4. National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. The Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research. [Bethesda, Md.]: The Commission, 1978. https://www.citiprogram.org/citidocuments/_001pic/1127_the_belmont_report.pdf

Sample Policies

Internal

	Attribute
Goal	To conduct a risk assessment of current employee phishing behaviors
Type	Phishing simulation
Target Group	All employees of the organization
Content	Targets will be sent emails that look similar to email already sent to the organization with indicators such as, incorrect sender email addresses, poor formatting. Content will not utilize spear-phishing (using personalized information of target) and will be generalized to the organization Content will utilize social engineering strategies such as urgency while avoiding threatening language.
Duration	Around 40 days. No longer than 90 days
Repeatability	Yes
Outcome	If participants click on link in phishing email; if participants disclose information to sender of phishing email
Disclosure	Participants were given no prior training, one email was sent to inform them that the organization would be conducting a phishing simulation. At the end of the simulation, participants will be sent an email allowing them to review their own behavior.
Data & Privacy	For the duration of the simulation information will be collected on each individual regarding if they clicked and/or shared information due to phishing. At the conclusion of the duration each participant will be sent a report regarding their behaviors. The overall results will be summed together in a final report for executive review. Individual performance in the phishing simulation will be retained for future follow-ups. If no future-follow-up is planned, it is advised that the non-summative data be deleted.
Risk	Emotional duress or stress due to content of phishing email. Stress due to disclosure of individual reports. Data breach of individual identifiers along with phishing performance data

External

	Attribute
Goal	To raise awareness and introduce phishing as a concept to members of a client organization
Type	Phishing simulation coupled with training
Target Group	All employees of the client organization
Content	Training session will take place first, introducing employees to the concept of phishing, provide examples, and recommend that they forward suspicious emails to a given email in IT. Targets will be sent emails that look similar to email already sent to the organization with indicators such as, incorrect sender email addresses, poor formatting. Upon clicking, targets will be informed that they have been phished, directed to an info page about phishing, and given a quiz (optional) to take about phishing. Content will utilize social engineering strategies such as urgency and may include threatening but non-violent language.
Duration	30 days
Repeatability	Yes
Outcome	If participants click on link in phishing email; if participants forward suspicious email to IT
Disclosure	Participants are given prior training and are informed that they will be targeted with phishing emails over the course of the next month. They are asked to be on the look out and given the option to opt out.
Data & Privacy	For the duration of the simulation information will be collected of clicks and forwards. This information will not be collected in a personally identifiable way. At the conclusion, the overall results will be summed together in a final report for executive review and the collected information deleted.
Risk	Emotional duress or stress due to content of phishing email. Stress from being asked to look out for phishing emails
Benefit	To inform participants about phishing, raise awareness, and encourage

Research

	Attribute
Goal	To understand the impact of behavior change if click content of phishing simulation email is (a) a quiz about phishing, (b) information about phishing, (c) no information is provided about phishing and target is sent via link to where they believe they were going
Type	Phishing simulation
Target Group	Selected targets in an organization randomly selected and distributed over race and gender
Content	Targets will be sent emails that look similar to email already sent to the organization with indicators such as, incorrect sender email addresses, poor formatting. Content will not utilize spear-phishing (using personalized information of target) and will be generalized to the organization Content will utilize social engineering strategies such as urgency while avoiding threatening language.
Duration	At 2, 7, and 42 days
Repeatability	Yes (3 times)
Outcome	If participants click on link in phishing email; if participants disclose information to sender of phishing email
Disclosure	Participants will be given no prior training. Informed consent will not be gotten ahead of time. Deception is necessary because awareness of phishing could influence participant's natural responses. Participants will be debriefed after the duration of the experiment has elapsed and informed of the deception.
Data & Privacy	For the duration of the simulation information will be collected on each individual regarding if they clicked and/or shared information due to phishing. Data will be saved in a way to anonymize the participants. The overall results will be summed together in a final report for executive review. All collected data except for the summative analysis will be deleted at completion.

Sample End of Exercise Notification

Hi *NAME(S)*,

This notification is for your awareness that, as of *TIME* today, *DATE*, the phishing attack simulation Citizen Clinic conducted for the *NAME OF ORG* has concluded.

In my opinion, the results will successfully highlight practices to sustain and improve upon both at the individual and organizational level. There are some things that *NAME OF ORG* did well and areas that *NAME OF ORG* will need to improve upon to increase their digital security. We will compile a report of our findings to share with leadership in private and with *MODIFYasNEEDED* as previously discussed.

We will discuss areas for improvement and procedures for recovery & response. Do advise participants that we took steps to mitigate risks such as no passwords being collected / stored and requesting account credentials for accounts that already have multi-factor authentication enabled or presumably would be "virtual identities." Any information sent to the two "attacker" accounts is also protected by strong authentication (Yubikeys).

Security is a team effort so any information disclosed should not seem like personal failures, but a combination of weak links in a chain. When a participant realizes they did disclose information to us, they should also realize that their disclosure was part of a larger system where ultimately the odds (and human nature) are stacked up against them. Participants should not compare themselves with others - neither at the individual or organizational level. Instead, we seek improvement that is relative to the current state - the goal is that participants are themselves better prepared to handle similar threats today or tomorrow compared to where they were yesterday.

That being said:

1) Participants should know that they will not be receiving any more simulated phishing emails from us via our attacker personas (*INSERT ATTACKER EMAILS*). It is possible that prior emails may be bumped in their inbox due to your email service's reminder feature. Any forms or "malicious" links have been disabled. Any messages sent to those accounts will not be returned.

2) The two attacker persona accounts (*INSERT ATTACKER EMAILS*) have already been reported and confirmed. We will detail the attacks in our brief, but for participants' peace of mind, there is no need to take any immediate recovery steps, although changing one's password, reviewing sign-in activity, and ensuring multi-factor authentication is enabled can provide some relief to personal feelings of discomfort.

3) The *NAME OF ORG* should resume their security and incident response policies & practices as usual. We're not in a posture to continually evaluate suspicious emails, but we're still, of course, happy to help as we can. *XYZ* should, in most cases, be your primary resource beyond internal information sharing and precautions.

Let me know if you have any questions. Thanks for participating!

Note: Additional confirmation of end of exercise to be sent via trusted channel (eg. Signal instead of email)

Last update: April 27, 2020