How to Scale a Phish: An Investigation into the Use of the NIST Phish Scale

Shanée Dawkins [0000-0002-8114-0608], *National Institute of Standards and Technology*
Jody Jacobs [0000-0002-6433-884X], *National Institute of Standards and Technology*

Abstract

Organizations around the world are using the NIST Phish Scale (NPS) in their phishing awareness training programs. As a new metric for measuring human phishing detection difficulty of phishing emails, the use of the NPS by phishing training implementers across different types of organizations has not been formally evaluated. This poster presents the results of a study on the use of the NPS by organizations with established phishing awareness training programs. Initial results suggest that there are areas for improvement, but training implementers perceive the NPS as useful overall and an asset to their organizations' cybersecurity awareness programs.

1. Introduction

Organizations often use embedded phishing awareness training programs to assess their phishing-related security risks. In these programs, phishing exercises are conducted by phishing training implementers, who send simulated phishing emails to employees of their organization to gauge the rate at which employees click or report the phish. However, research has shown that click rates – whether people click or don't click on links and attachments – do not provide a complete picture to understanding staff behaviors [2,3]. Click rates can be perceived as excellent or poor without the appropriate context.

The NIST Phish Scale (NPS) was developed in 2019 to provide context into clicking behaviors so that phishing training implementers know if an email was easy or difficult to detect as a phish [1,5,6]. However, the NPS was developed using data from a single organization's training program. Since its use spread across different types of organizations, its wider use has yet to be evaluated.

The study presented here was designed for multiple purposes, including gauging the accuracy of the NPS metric and measuring its usefulness. The results presented in this poster are on a subset of the data from the conducted study, and focus on the research question, "Are the NPS components (i.e., cues, premise alignment, detection difficulty) easy to understand and useful to an organization?"

USENIX Symposium on Usable Privacy and Security (SOUPS) 2023. August 6 -- 8, 2023, Anaheim, CA USA.

2. NIST Phish Scale

The NPS methodology was based on the results of research to understand staff behaviors, incorporating both observable cues of the phishing email itself, as well as the user context of the email's recipient [3]. Cues are properties of an email that compel a user to click on a fraudulent link or attachment or serve as red flags alerting the user that the email may be a phish. The premise alignment characterizes the relevancy of the email premise for the target audience, based on workplace responsibilities and culture, business practice plausibility, and staff expectations. The output of the NPS is the detection difficulty – a measure of how hard or easy an email is for a human to detect as a phish. Emails with fewer cues and higher premise alignment are harder to detect than emails with many cues and lower premise alignment (see Table 3 in Appendix).

3. Methodology

This study was designed to collect empirical data on how phishing training implementers use the NPS. Phishing training implementers were recruited from a variety of domestic and international organizations across multiple sectors (see Appendix Section 7.1). Since the NPS is intended for use as a metric in phishing awareness training programs, the study was designed so that participants could incorporate its use into their existing and well-established programs.

To participate in the study, the phishing training implementers were provided a user guide with step-by-step instructions on how to apply the NPS to a phishing email in their programs. They were instructed to use this user guide to apply the NPS to a single simulated phishing email being used in their exercises, prior to receiving click rates and other results from the exercise.

Participants were also sent a unique link to an online survey and instructed to complete the survey once their exercise concluded and they had reviewed the click rate results. Both closed- and open- ended questions on the survey were designed to gauge participants' experiences

using the NPS, capturing information about their perceived usefulness of the method.

The study was approved by our institution's Research Protections Office with informed consent required for all participants. To ensure anonymity, each participant was assigned an alphanumeric reference code.

4. Results

The survey asked wide ranging questions regarding the use of the NPS for phishing awareness training exercises. This section reports on a subset of those questions related to the understandability and use of the NPS metric.

4.1. Participant and Organizational Demographics

Individuals responsible for implementing phishing awareness training exercises in their organizations participated in the study. Of the five total participants who completed the study[1], three participants indicated that they were from the U.S. government, one was a U.S. government contractor, and one worked in private industry. Additionally, three participants identified themselves as IT specialists (Cybersecurity/Information Security), one stated they were a program/project manager, and one participant was a training specialist.

Study participants were asked to provide background and organizational experience with the NPS and phishing awareness training exercises. Four study participants indicated that they had 1–5 years of phishing awareness training experience; one participant had 6–10 years. Study participants also indicated that their organizations conducted phishing awareness training exercises with simulated phishing emails over several years: two organizations for 3–4 years; one organization for 5–6 years; one participant for more than 6 years. Two organizations indicated that they have been using the NPS for more than 1 year, while another had only used it for 6–12 months.

4.2. Phishing exercises

Study participants were asked questions about the simulated phishing exercise their organization conducted and the results of the exercise. Participants stated that their exercise click rates ranged from 2.00% to 8.00%. Detection difficulty ratings of the email used for the exercise ranged from *least to moderately difficult* to *very difficult*. None of the emails were categorized as *least difficult,* suggesting that, as in previous studies [6], the simulated phishing emails used for training purposes are intended to train individuals on more difficult phishes to detect.

The NPS has proven to be effective when used with a target audience consisting of individuals with similar roles and responsibilities – typically a small group of people [6]. To gauge the perceived effectiveness of the NPS in this study, participants were asked to provide the approximate size of their target audience and whether they felt their click rates aligned with their NPS detection difficulty as expected. Participants with smaller target audiences, less than 5,000 people, indicated that their click rates and detection difficulty rating aligned as expected. Those who sent the simulated phishing email to a larger target audience, between 10,000 and 29,000 people, had mixed results. One participant said that their organization's click rate (3.40%) did not align with their detection difficulty rating (*very difficult*) as they expected. Another participant indicated that their organization's click rate (7.43%) did align with their detection difficulty rating (*very difficult*) as they expected.

4.3 Applying the NPS

Participants were asked how easy or difficult it was to apply seven aspects of the NPS to their phishing exercise email (see Table 4 in Appendix). Of these seven items, five were *easy* or *very easy* to apply by at least 75% of participants. The two items that participants had more difficulty applying were related to the NPS premise alignment (40% indicated *difficult* or *very difficult*).

In their open-ended responses, participants indicated a lack of clarity in understanding two elements of the premise alignment – Element 1: Mimics workplace process or practice and Element 2: Has workplace relevance. Participants stated that these two elements "could be misunderstood to be similar or the same thing" (I301). Another further suggested that the first two elements of the premise alignment "could be differentiated a little more clearly" (F102). Some participants stated that they had some difficulties calculating and scoring the premise alignment elements. Participant F108 stated that it is, "hard to tell how the [numerical scores for the] first three categories differed substantially from each other."

Despite the challenges experienced in applying the premise alignment component of the NPS, participants expressed an overall positive attitude towards its use: 80% of participants indicated the NPS was *useful* or *very useful*. Likewise, 80% of participants indicated that the NPS was appropriate for use in their organizations. Overall, participants felt that the NPS helps their organization contextualize click rates and contributes to their overarching phishing awareness training program.

5. Disclaimer

Any mention of commercial products or companies is for information only and does not imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the products are necessarily the best available for the purpose.

[1] Testing with a small sample can be effective [4].

6. References

1. Barrientos, F., Jacobs, J., and Dawkins, S., Scaling the Phish: Advancing the NIST Phish Scale. In Proceedings of HCII 2021 (23rd International Conference on Human-Computer Interaction). July 24 – July 29, 2021. https://doi.org/10.1007/978-3-030-78642-7_52 (Accessed February, 2023)
2. Greene, Kristen & Steves, Michelle & Theofanos, Mary. (2018). No Phishing beyond This Point. Computer. 51. 86-89. https://doi.org/10.1109/MC.2018.2701632 (Accessed February, 2023)
3. Greene, K. K., Steves, M., Theofanos, M., & Kostick, J.: User context: an explanatory variable in phishing susceptibility. In Proc. 2018 Workshop Usable Security (USEC) at the Network and Distributed Systems Security (NDSS) Symposium. (2018).
4. Neilsen, J., How Many Test Users in a Usability Study? (2012). https://www.nngroup.com/articles/how-many-test-users/ (Accessed June 2023)
5. Steves, M. P., Greene, K. K., & Theofanos, M. F.: A phish scale: rating human phishing message detection difficulty. Proceedings 2019 Workshop on Usable Security. Workshop on Usable Security, San Diego, CA. (2019). https://doi.org/10.14722/usec.2019.23028
6. Steves, M., Greene, K., & Theofanos, M.: Categorizing human phishing difficulty: A Phish Scale. Journal of Cybersecurity, 6(1), tyaa009. (2020). https://doi.org/10.1093/cybsec/tyaa009

7. Appendix

7.1 Recruitment & Sampling

The specialized nature of the users of the NPS necessitated the use of a convenience sample for this study; participants recruited for this study were known to the researchers and regularly conducted simulated phishing exercises as a part of their organization's embedded phishing awareness training programs. Participants recruited included federal and non-federal chief information security officers (CISOs) and phishing training implementers, both in the U.S. and international organizations.

Inclusion Criteria:

1. Participants were required to be 18 years of age or older.
2. Participants' job responsibilities included conducting their organization's embedded phishing awareness training exercises.
3. Participants conducted at least two phishing awareness training exercises per year.

Exclusion Criteria:

1. Individuals who held cybersecurity awareness management positions, or similar, were excluded if they were not directly involved in conducting their organization's embedded phishing awareness training exercises.

The study ran for six months in 2022. This allowed participation from organizations that conduct their phishing awareness training exercises at various time intervals (e.g., quarterly, semi-annually).

7.2 NPS Components

The NPS cues component consists of 23 cues, grouped into five types. Each instance of a cue in a phishing email is tallied and categorized according to Table 1.

Table 1. Phishing Email Cue Category Mapping

Total Cue Count	Cue Category
1 – 8 cues	Few (more difficult)
9 – 14 cues	Some
15 or more cues	Many (less difficult)

The NPS premise alignment component consists of five elements aimed to collect context about the target audience of the phishing email:

1. Mimics a workplace process or practice.
2. Has workplace relevance.
3. Aligns with other situations or events, including external to the workplace.
4. Engenders concern over consequences for NOT clicking.
5. Has been the subject of targeted training, specific warnings, or other exposure.

These elements are assessed using a five-point applicability scale from 0 to 8, increasing in severity by units of 2. These scores are used in an equation that results in a rating for the premise alignment component, which is then categorized according to the mapping in Table 2.

Table 2. Phishing Email Premise Alignment Category Mapping

Premise Alignment Rating	Premise Alignment Category
10 and below	Low
11 – 17	Medium
18 and higher	High

The cues category and the premise alignment category are both considered when determining the overall detection difficulty of a phishing email (see Table 3).

Table 3. The Phish Scale - Detection Difficulty

Number of Cues	Premise Alignment Rating	Detection Difficulty
Few (more difficult)	High	Very difficult
	Medium	Very difficult
	Low	Moderately difficult
Some	High	Very difficult
	Medium	Moderately difficult
	Low	Moderately to Least difficult
Many (less difficult)	High	Moderately difficult
	Medium	Moderately difficult
	Low	Least difficult

7.3 Survey Data

This section provides a subset of the survey data presented in this paper.

Table 4. How easy or difficult was your experience applying the NPS to a simulated phishing email?

	Very difficult	Difficult	Easy	Very Easy
Counting cues	0	1	2	2
Categorizing the cues	1	0	3	1
Applying the cues component overall	0	1	1	2
Rating premise alignment elements	0	2	2	1
Categorizing the premise alignment	1	1	2	1
Applying the premise alignment component overall	1	0	3	1
Determining overall detection difficulty	1	0	3	1