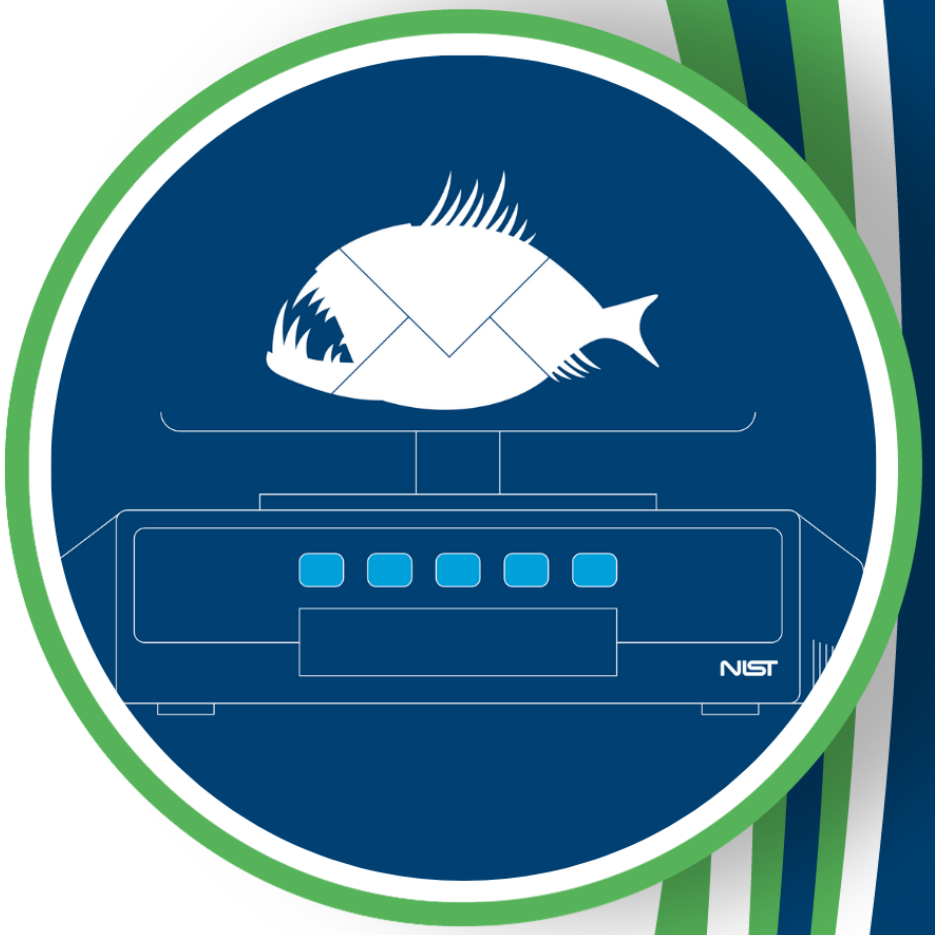
NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY
U.S. DEPARTMENT OF COMMERCE

NIST PHISH SCALE
USER GUIDE

Prepared By:
Shanée Dawkins
Jody Jacobs



NIST Technical Note
NIST TN 2276

NIST Phish Scale User Guide

Shanée Dawkins
Visualization and Usability Group, Information Access Division
Information Technology Laboratory

Jody Jacobs
Visualization and Usability Group, Information Access Division
Information Technology Laboratory

This publication is available free of charge from:
https://doi.org/10.6028/NIST.TN.2276

November 2023

U.S. Department of Commerce
Gina M. Raimondo, Secretary

National Institute of Standards and Technology
Laurie E. Locascio, NIST Director and Under Secretary of Commerce for Standards and Technology

NIST TN 2276

November 2023

NIST Technical Series Policies

Publication History

Approved by the NIST Editorial Review Board on 2023-11-08

How to Cite this NIST Technical Series Publication

Dawkins, S., Jacobs, J. (2023) NIST Phish Scale User Guide. (National Institute of Standards and Technology, Gaithersburg, MD), NIST Series TN 2276. https://doi.org/10.6028/NIST.TN.2276

NIST Author ORCID iDs

Shanée Dawkins: 0000-0002-8114-0608
Jody Jacobs: 0000-0002-6433-884X

Contact Information

human-cybersec@nist.gov

Abstract

Phishing cyber threats impact private and public sectors both in the United States and internationally. Embedded phishing awareness training programs, in which simulated phishing emails are sent to employees, are designed to prepare employees in these organizations to combat real-world phishing scenarios. Cybersecurity and phishing awareness training implementers and practitioners use the results of these programs, in part, to assess the security risk of their organization. The NIST Phish Scale is a method created for these implementers to rate an email's human phishing detection difficulty as part of their cybersecurity awareness and phishing training programs. This User Guide outlines the Phish Scale in its entirety while providing instructional steps on how to apply it to phishing emails. Further, appendices include 1) worksheets to assist training implementers in applying the Phish Scale and 2) detailed information regarding email properties and associated research in the literature.

Keywords

Business Email Compromise (BEC); Cybersecurity; Human-Centered Cybersecurity; Phish Scale; Phishing; Social Engineering; Usable Security, Usable Cybersecurity.

Table of Contents

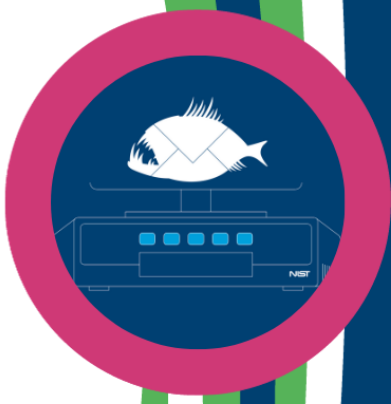
List of Tables

List of Figures

Acknowledgments

NIST would like to acknowledge Michelle Steves, Kristen Greene, Mary Theofanos, and Jennifer Kostick for their efforts in development of the NIST Phish Scale. The authors would also like to thank Fern Barrientos, Susanne Furman, and Kevin Mangold for their contributions in the development of this User Guide. Special thanks to the phishing awareness training program directors who provided feedback on the use of this document in their organizations.

1. Introduction



Phishing is an email-based cybersecurity threat in which cybercriminals attempt to get sensitive information from email recipients. It is a social engineering technique that compels an email recipient to perform an action beneficial to the attacker (e.g., clicking a link to a fraudulent website or downloading a malicious attachment) [17]. Phishing cyber threats impact private and public sectors both in the United States (U.S.) and internationally. It remains one of the top security threats to these organizations, costing companies billions [11]. As technological safeguards and email filters are not guaranteed to block all incoming malicious emails, humans are often an organization's last line of defense against a phishing attack. Therefore, it is imperative that employees in these organizations are prepared in case of a real-world phishing scenario.

Embedded phishing awareness training programs are designed to help combat phishing threats. In these programs, simulated phishing emails are sent to employees in an effort to train them to spot real phishing emails they may receive. The personnel executing these types of programs, referred to in this User Guide as "training implementers," use the results of these programs, in part, to assess the security risk of their organization. These results are usually measured using *click rates* – the number of people who clicked on a potentially malicious link or attachment out of the total number of people sent the simulated phishing email and *reporting rates* – the number of people who report a suspicious email to their organization out of the total number of people sent the simulated phishing email. However, click rates and reporting rates do not provide a complete picture of an organization's phishing risk; they provide a single point of insight – what percentage of people "fell" for the phish. The fact that some phishing emails are more difficult for people to detect than others can be incorporated into the assessment of a simulated phishing training exercise, providing an additional metric in assessing overall cybersecurity risk. The method outlined in this User Guide addresses this concern.

The NIST Phish Scale is a method for training implementers to rate an email's human phishing detection difficulty, evaluating both the properties of a phishing email itself and the characteristics of the email's recipients.

The NIST Phish Scale, hereafter referred to as the Phish Scale, was originally published in a research article in 2019[1] [22], expanded in a research journal in 2020 [23], and further studied in subsequent years [3][4]. These publications go into detail about how the Phish Scale was created using empirical phishing simulation data and the research behind evaluating it with training implementers. This User Guide serves as the first step to bridge the gap from research to practice, outlining the Phish Scale in its entirety while providing instructional steps on how to apply it to phishing emails. A worksheet to assist training implementers in applying the Phish

[1] Prior NIST research found that user context plays a key role in interpreting the click rates [9].

Scale is in Appendix A. Detailed information regarding email properties and associated research in the literature is in Appendix B.

1.1. Using this User Guide

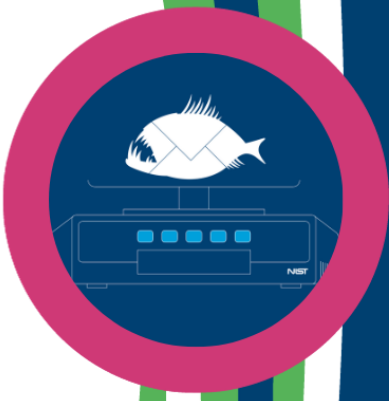
This User Guide is intended for use by practitioners – phishing awareness training implementers, cybersecurity awareness training professionals and other computer security professionals responsible for conducting phishing training exercises (e.g., designing, executing and/or analyzing data). The human phishing detection difficulty rating resulting from application of the Phish Scale helps phishing awareness training implementers in two primary ways:

1. By providing context regarding training message click rates and reporting rates for a target audience, and
2. By providing a way to characterize actual phishing threats so the training implementer can reduce the organization's security risk by tailoring training to the types of threats their organization is facing [22].

While this guide was developed using empirical data [22], when considering use of the Phish Scale, training implementers should customize the method to fit their organization's current environment and employee population. Additionally, when applying the Phish Scale to an email, it is important that the training implementer is consistent in their evaluation to ensure the effectiveness of the human phishing detection difficulty comparison across emails.

Phishing is just one aspect of a training implementer's overall cybersecurity program. Considering an organization's cybersecurity awareness and training program as a whole, the Phish Scale is an additional metric that training implementers can use to reduce their organization's security risk while still meeting their organization's mission and risk tolerance.

2. The NIST Phish Scale



The Phish Scale has two main components used collectively to determine human phishing detection difficulty [22]:

1. A scoring system for observable characteristics of the phishing email itself
2. A scoring system for alignment of the phishing email premise with respect to a target audience

The first component is measured by assessing the visual indicators - cues - present in the email which may alert email recipients when spotting a phish, such as the number of cues, nature of the cues, and repetition of cues. The second component – premise alignment – is based on current events, the environment of an organization, and the recipient's roles and responsibilities. Both components are first measured, then interpreted collectively, resulting in an overall human detection difficulty rating for a phishing email. Sections 2.1 through 2.3 detail these components with instructions on how to determine overall human phishing detection difficulty for a phishing email.

2.1. Email Cues

The first component of the Phish Scale is a scoring system for observable characteristics of the phishing email itself, referred to as email *cues* [22].

Cues are the properties of an email that either compel a user to click on a fraudulent link or attachment or alert the user that the email may be a phish. A lower number of cues in a phishing email indicates an email that is more difficult for someone to detect as a phish; a higher number of cues indicates easier detection.

The cues in an email provide an objective measure of the email itself; the number of cues present in an email is categorized in this component of the Phish Scale. This cue category along with an email's premise alignment category, are used to determine detection difficulty. When categorizing the number of cues in a phishing email, it is important to first understand the types of cues that may be present in a phishing email and where they occur.

2.1.1. List of Cues

Phishing email cues are categorized into five types [22]:

1. Errors – relating to spelling and grammar errors and inconsistencies contained in the message;

2. Technical indicators – pertaining to email addresses, hyperlinks and attachments;
3. Visual presentation indicators – relating to branding, logos, design and formatting;
4. Language and content – such as a generic greeting and lack of signer details, use of time pressure and threatening language; and
5. Common tactics – use of humanitarian appeals, "too good to be true" offers, time-limited offers, poses as a friend, colleague, or authority figure, and so on.

Each cue type has associated cues, 23 in total, listed in Table 1. Further insight into the cues and what to look for in an email can be found in Appendix B.

Table 1. List of Cues by Type

Cue Type	Cue Name
Error	Spelling and grammar irregularities
	Inconsistency
Technical indicator	Attachment type
	Sender display name and email address
	URL hyperlinking
	Domain spoofing
Visual presentation indicator	No/minimal branding and logos
	Logo imitation or out-of-date branding/logos
	Unprofessional looking design or formatting
	Security indicators and icons
Language and content	Legal language/copyright info/disclaimers
	Distracting detail
	Requests for sensitive information
	Sense of urgency
	Threatening language
	Generic greeting
	Lack of signer details
Common tactic	Humanitarian appeals
	Too good to be true offers
	You're special
	Limited time offer
	Mimics a work or business process
	Poses as friend, colleague, supervisor, authority figure

2.1.2. Identifying Cues

When analyzing an email, an understanding of how to properly identify its cues is essential. While Sec. 2.1.1 specifies the characteristics of each cue, this section dissects the anatomy of an email and highlights where different types of cues are typically found. Figure 1 illustrates a sample phishing email. In an overall email, there are four major components:

- Header – includes *From*, *Sent*, and *To* information
- Subject
- Attachment – if present, one or more downloadable files
- Message – including the Salutation (greeting), Body (primary content), Closing (signature), and Postscript (disclaimers)

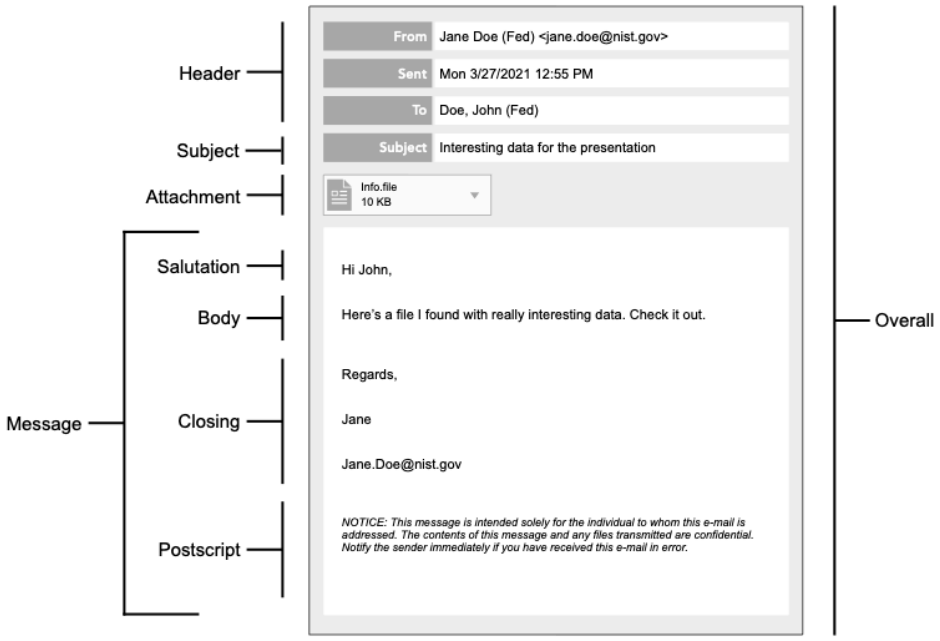


Figure 1. Phishing Email Template Example

"Error" and "Technical indicator" cue types can be found anywhere in an overall email, depending on the cue. "Visual presentation indicator" cue types are related to how an email is displayed, and therefore are typically found in an email's message. "Language and content" and "Common tactic" cue types are more about the premise of the email and are located in the subject or message of an email. See Appendix B for the typical location for individual cues[2].

2.1.3. Categorizing the Number of Cues

Cue categorization depends on the number of observable cues in a phishing email. A phishing email should be closely examined to locate and identify the cues present. To do this, use the criteria in Table 2, counting each instance of the cues present in a phishing email. For example, if

[2] Due to varying email clients, user email customization, and the device on which users will view their email, the location of cues may vary.

a phishing email has three spelling errors and two grammatical errors, five “Spelling and grammar irregularities” are tallied towards the total number of cues. Note that some email characteristics can be identified and counted as multiple cues. For example, an email with a hyperlink displayed as “www.niist.gov” (as opposed to the real domain www.nist.gov) that has an underlying Uniform Resource Locator (URL) to “commerce.gov” would be counted as both “Spelling and grammar irregularities” and “URL hyperlinking” cues. Score the phishing email by tallying the number of cues present in the email and summing the total number of cues. The form included in Appendix A can be used as a worksheet to assist you with counting cues.

Table 2. Criteria for Counting Cues

Cue Type	Cue Name	Criteria for Counting
Error	Spelling and grammar irregularities	Does the message contain Inaccurate spelling or grammar use, including mismatched plurality?
	Inconsistency	Are there inconsistencies contained in the email message?
Technical indicator	Attachment type	Is there a potentially dangerous attachment?
	Sender display name and email address	Does a display name hide the real sender or reply-to email addresses?
	URL hyperlinking	Is there text that hides the true URL behind the text?
	Domain spoofing	Is a domain name used in addresses or links plausibly similar to a legitimate entity's domain?
Visual presentation indicator	No/minimal branding and logos	Are appropriately branded labeling, symbols, or insignias missing?
	Logo imitation or out-of-date branding/logos	Do any branding elements appear to be an imitation or out-of-date?
	Unprofessional looking design or formatting	Does the design and formatting violate any conventional professional practices? Do the design elements appear to be unprofessionally generated?
	Security indicators and icons	Are any markers, images, or logos that imply the security of the email present?
Language and content	Legal language/copyright info/disclaimers	Does the message contain any legal-type language such as copyright information, disclaimers, or tax information?
	Distracting detail	Does the email contain details that are superfluous or unrelated to the email's main premise?
	Requests for sensitive information	Does the message contain a request for any sensitive information, including personally identifying information or credentials?
	Sense of urgency	Does the message contain time pressure to get users to quickly comply with the request, including implied pressure?
	Threatening language	Does the message contain a threat, including an implied threat, such as legal ramifications for inaction?
	Generic greeting	Does the message lack a greeting or lack personalization in the message?
	Lack of signer details	Does the message lack detail about the sender, such as contact information?
Common tactic	Humanitarian appeals	Does the message make an appeal to help others in need?
	Too good to be true offers	Does the message offer anything that is too good to be true, such having won a contest, lottery, free vacation and so on?
	You're special	Does the email offer anything just for you, such as a valentine e-card from a secret admirer?
	Limited time offer	Does the email offer anything that won't last long or for a finite length of time?
	Mimics a work or business process	Does the message appear to be a work or business-related process, such as a new voicemail, package delivery, order confirmation, notice of invoice?
	Poses as friend, colleague, supervisor, authority figure	Does the message appear to be from a friend, colleague, boss or other authority entity?

The Phish Scale has three categories that map to the total number of cues score:

- Few – the phishing email has a lower number of cues with fewer opportunities to identify the email as a phish
- Some – the phishing email has a moderate number of cues
- Many – the phishing email has a higher number of cues, with more opportunities to identify the email as a phish

Use the mapping shown in Table 3 to determine a phishing email's cue category based on the total number of cues score for that email.

Table 3. 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 few, some, or many cue category carries forward into Sec. 2.3, Determining Detection Difficulty, along with the premise alignment which is covered comprehensively next in Sec. 2.2.

2.2. Premise Alignment

The second component of the Phish Scale focuses on the relationship between user context and the phishing email message, referred to as the *premise alignment* [22]. When the relationship is strong, this is close to what is often called spear phishing.

Premise alignment is a measure of how closely an email matches the work roles or responsibilities of an email's recipient or organization. The stronger an email's premise alignment, the more difficult it is to detect as a phish. Inversely, the weaker an email's premise alignment, the easier it is to detect as a phish.

Evaluating a phishing email's premise alignment is a process of characterizing the relevance of the email message premise to a *target audience*. This *target audience* can be centered on one of

the various levels within your organization (e.g., divisions, departments, groups, teams) to contextualize click rates and their direct relationship to specific departments or employees. The premise alignment cannot be evaluated without knowledge of the *target audience's* context of work with respect to the premise of the phishing email's message. As such, measuring a phishing email's premise alignment should be performed by an individual with knowledge of the *target audience's* work culture and responsibilities.

Measuring a phishing email's premise alignment begins with assigning a numerical value – the applicability score – to five individual premise alignment elements. The premise alignment rating is then calculated using these applicability scores. This final premise alignment rating is then mapped to a strong, medium, or weak premise alignment category, which, considered with the cues category, is used to determine an email's detection difficulty. An understanding of the five elements that are the basis for this process is needed before evaluating a phishing email. A description of these elements and how to use them to calculate the premise alignment rating follows in the sections below. The worksheet form included in Appendix A can be used to assist training implementers in measuring an email's premise alignment.

2.2.1. Premise Alignment Elements

Each of the five premise alignment elements are described in detail below.

Element 1 – Mimics a Workplace Process or Practice

This element reflects the relevance of the email's premise to a process or practice of the target audience. Any processes or functions that typically happen in your organization should be considered for this element.

For example, if the target audience typically receives official organization chat notifications via an app, an email that notifies the recipient of a missed chat message would have a lower applicability score for this element. However, if email is the typical mechanism for chat notifications, that email would have a higher applicability score for this element.

Element 2 – Has Workplace Relevance

This element reflects the relevance of the premise to the work of the target audience – including their roles and responsibilities. It is key to have knowledge of your target audience's job duties and job functions, in order to appropriately evaluate this element.

For example, if the target audience is the finance department and an email has a premise of a late or missed payment, that email would have a higher applicability score for this element. Another consideration for this element is the sender's domain. If the domain name of the email sender is the same or similar to your organization's domain, (e.g., john.doe@nist.gov is receiving an email from jane.doe@nist.gov), the email would have a higher applicability score for this element. Although many organization's security policies prohibit or discourage the use of personal email for business purposes, it is still a common enough practice that senders with familiar names and public domains should be considered for this element. If the target audience is familiar with an employee at your organization, e.g., an executive named John Doe, then an email from "john.doe@gmail.com" would have a higher applicability score.

Also, consider if the email is relevant to the work and responsibilities of all or a subset of the target audience. For example, if the premise of the email has a higher contextual alignment, but only to a small portion of the target audience, then the workplace relevance element should be assigned a mid-range applicability score.

Element 3 – Aligns with other situations or events, including external to the workplace

This element is based on timing of when the phishing email is received by the target audience. It reflects the alignment of the premise to internal and external situations or events directly or indirectly affecting your organization. Examples of calendar-related external events in the U.S. are Christmas, New Year's Day, and Memorial Day; examples of internal events are a new organization director/president/leader hired, the opening of a new branch office, and the start or end of a fiscal year. If the event is current and relevant for the target audience, then the applicability score for this element should be higher (e.g., a phishing email sent around February 14, related to Valentine's Day). Conversely, if an event is not current or relevant to the intended audience, then the premise should have a lower applicability score (e.g., a phishing email about an office winter holiday party sent mid-summer).

Element 4 – Engenders concern over consequences for NOT clicking

This element reflects potentially harmful ramifications if no action is taken, raising the likelihood of the phishing email recipient clicking on fraudulent links or attachments. Certain email premises elicit this reaction in recipients more than others. For example, a phishing email acting on a user's fear of missing out (e.g., a missed message, an informational notice) may not produce the same response as an email with a more serious allegation (e.g., potential leak of protected health information, exposure of personal information, ransomware). The former would yield a lower applicability score for this element than the latter.

Element 5 – Has been the subject of targeted training, specific warnings, or other exposure

This element reflects the effects of training on the target audience, including, for example, phishing-related organizational training on recognizing and reporting phishing emails. Ideally, employees who have had exposure to some form of IT security training related to phishing would be more judicious in identifying an email as a phish (a higher applicability score for this element) versus those who have not received training (a lower applicability score for this element). Phishing training is not constrained to components within formal IT security training courses; training refers to any awareness materials or guidance to which the target audience has been exposed. Training may refer to:

- formal IT cybersecurity awareness and training programs [5][7][16][19][25];
- educational materials or seminars on how to identify a phishing email; or
- organizational emails alerting employees to be on the lookout for phishing attempts or warning against specific types of phishing attacks.

If the phishing email recipient has had a considerable amount of training in the detection of phishing emails, this element would have a higher applicability score. Likewise, the applicability score would be high if the organization has had a robust phishing awareness training program in place for an extended period of time.

2.2.2. Scoring the Premise Alignment Elements

To calculate the premise alignment rating, first assign each of the five premise alignment elements an even numerical value between zero and eight – the applicability score (see Table 4). An element with an applicability score of zero indicates that there is a complete mismatch in relevancy of the element to the target audience. Inversely, a high applicability score of eight indicates that the element is very applicable to the target audience.

Table 4. Premise Alignment Applicability Scale

Applicability Scale[3]	Applicability Score
Extreme applicability, alignment, or relevancy	8
Significant applicability, alignment, or relevancy	6
Moderate applicability, alignment, or relevancy	4
Low applicability, alignment, or relevancy	2
Not applicable, no alignment, or no relevancy	0

Table 5 provides criteria for the five premise alignment elements. Use these criteria, along with the applicability scale, to determine the applicability score for each element.

Table 5. Premise Alignment Elements Scoring Criteria

Premise Alignment Elements	Scoring Criteria
1: Mimics a workplace process or practice	Does this element attempt to capture premise alignment with workplace process or practice for the target audience?
2: Has workplace relevance	Does this element attempt to reflect pertinence of the premise for the target audience?
3: Aligns with other situations or events, including external to the workplace	Does this element align to other situations or events, even those external to the workplace, lending an air of familiarity to the message?
4: Engenders concern over consequences for NOT clicking	Does this element reflect potentially harmful ramifications for not clicking raise the likelihood to clicking?
5: Has been the subject of targeted training, specific warnings, or other exposure	Does this element reflect targeted training effects that would lead to premise detection? Care must be taken to appropriately incorporate the training or warning specificity, as transfer of learning is quite difficult.

[3] Steves, et. al. used the term “anchors” in the most recent phish scale publication [23]. For this handbook, the term “applicability scale” is used for clarity.

The applicability score for each premise alignment element is used to calculate the final premise alignment rating.

2.2.3. Categorizing the Premise Alignment

The applicability scores from the previous section are used to calculate the final premise alignment rating: sum the applicability scores for elements one through four, then subtract the applicability score for element 5 from the total. Element five pertains to training and helps with detection; therefore, the numerical value assigned to this element is subtracted from the total sum. Equation 1 below shows the calculation needed for the premise alignment.

$$Premise\ Alignment\ Rating = Sum(Element1\ through\ Element4) - Element5 \quad (1)$$

The highest possible premise alignment rating is 32, indicating that a phishing email message matches up with the target audience and the target audience has not had any related training nor received any prior alert or warning about an upcoming phishing exercise. The lowest possible premise alignment rating is -8, indicating that the phishing email is a complete mismatch with the target audience and they have received prior phishing-related training, alerts or warnings.

Once the final premise alignment rating is calculated, it can be mapped to one of the three premise alignment categories[4]:

- Strong – the alignment of the phishing email's premise to the target audience is high, making the email difficult to detect as a phish
- Medium – the alignment of the phishing email's premise to the target audience is moderate
- Weak – the alignment of the phishing email's premise to the target audience is low, making the email less difficult to detect as a phish

Use the mapping depicted in Table 6 to determine the premise alignment category for the phishing email.

Table 6. Phishing Email Premise Alignment Category Mapping

Premise Alignment Rating	Premise Alignment Category
10 and below	Weak
11 – 17	Medium
18 and higher	Strong

The weak, medium, or strong premise alignment category carries forward into Sec. 2.3, along with the cues category from Sec. 2.1.3.

[4] In Steves, et. al publication [23], the categories used for premise alignment were high, medium, and low. For this User Guide, the terms strong, medium, and weak will be used. While the categories are nominally different, the meaning and categorization mapping remain the same.

2.3. Determining Detection Difficulty

The final step in applying the Phish Scale is to determine the overall detection difficulty of an email. The previously determined cues (see Sec. 2.1.3) and premise alignment (see Sec. 2.2.3) categories for a phishing email are analyzed collectively to determine the phishing email's detection difficulty, as shown in Table 7.

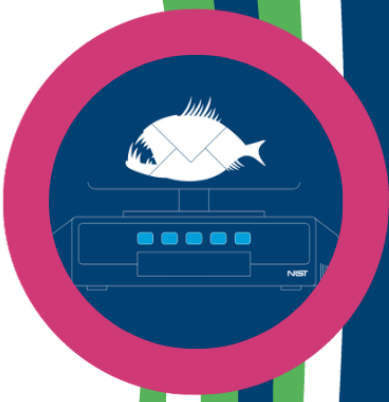
Table 7. The Phish Scale - Detection Difficulty

Cues Category	Premise Alignment Category	Detection Difficulty
Few (more difficult)	Strong	Very difficult
	Medium	Very difficult
	Weak	Moderately difficult
Some	Strong	Very difficult
	Medium	Moderately difficult
	Weak	Moderately to Least difficult
Many (less difficult)	Strong	Moderately difficult
	Medium	Moderately difficult
	Weak	Least difficult

Emails with few cues and a strong premise alignment are more difficult for a human to detect as a phish than those with more cues and a weak premise alignment.

For example, phishing emails categorized as having "Few" cues and "Medium" premise alignment have a "Very difficult" detection difficulty rating. Or, phishing emails categorized as having "Some" cues and a "Medium" premise alignment have a detection difficulty rating of "Moderately difficult".

3. Interpreting Results



Using the Phish Scale to understand the detection difficulty of a phishing email helps phishing awareness training implementers in two primary ways. First, the Phish Scale provides context regarding training message click rates and reporting rates for a target audience. For example, phishing emails that are Very Difficult to detect might understandably result in high click rates when used in a simulated phishing exercise; Least Difficult emails might likely result in lower click rates. However, when a phishing exercise yields unexpected results (e.g., a Least Difficult email that results in high click rates), it may indicate that modified or additional training is needed for the target audience.

Second, the Phish Scale provides a way to characterize actual phishing threats so training implementers can reduce their organization's security risk by tailoring training to the types of threats their organization faces while still maintaining a resilient security posture. One benefit of a strong and resilient security posture is safeguarding internal and external trust. A robust phishing program should not be a stagnant "check the box" type of exercise, but rather an evolving part of a mature cybersecurity awareness and training program to provide mature, metrics-driven results. Organizations need to tailor their cybersecurity and awareness training program to their unique environment and employees' needs and requirements while still meeting their organization's mission and risk tolerance. The level of security an organization implements should be commensurate with its risk and organizational purpose and operations. In other words, the higher the risk the organization encounters, the higher the level of security the organization should implement.

Lastly, an organization's cybersecurity awareness and training program is not a silver bullet cure-all. An organization needs a multi-pronged approach, considering technology, processes, and people, to identify, react, and report suspicious phishes. When applied to both a user's home and work environment, these tactics can give users concrete skills and knowledge to better prepare them to defend against potential phishing attempts, protecting both the user and their organization.

References

[1] Alazab, Mamoun and Broadhurst, Roderic, Spam and Criminal Activity (2016). Trends and Issues in Crime and Criminal Justice (Australian Institute of Criminology), No 52, RegNet Research Paper No. 2014/44. DOI: 10.2139/ssrn.2467423

[2] Alutaybi, A., Arden-Close, E., McAlaney, J., Stefanidis, A., Phalp, K., and Ali, R. (2019, October). How Can Social Networks Design Trigger Fear of Missing Out?. In Proc. of 2019 IEEE International Conference on Systems, Man and Cybernetics (SMC) (pp. 3758-3765). IEEE.

[3] 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. DOI: 10.1007/978-3-030-78642-7_52

[4] Dawkins, S., & Jacobs, J. (2023). How to Scale a Phish: An Investigation into the Use of the NIST Phish Scale. Proceedings of the Nineteenth Symposium on Usable Privacy and Security, Anaheim, CA, US.

[5] deZafra, D., Pitcher, S., Tressler, J., Ippolito, J. (1998). Information Technology Security Training Requirements: a Role- and Performance-Based Model. (National Institute of Standards and Technology, Gaithersburg, MD), NIST Special Publication (SP) 800-16. DOI: 10.6028/NIST.SP.800-16

[6] Downs, J.S., Holbrook, M., and Cranor, L.F., (2006, July). Decision strategies and susceptibility to phishing. In Proc. of the Second Symposium on Usable Privacy and Security (SOUPS '06), ACM, 2006, pp. 79–90.

[7] Federal Information Security Modernization Act of 2014, Pub. L. No. 113-283 (2014). https://www.govinfo.gov/app/details/PLAW-113publ283 (Last retrieved October 2023).

[8] Grazioli, S. (2004, March).Where Did They Go Wrong? An Analysis of the Failure of Knowledgeable Internet Consumers to Detect Deception Over the Internet. Group Decision and Negotiation 13, 149–172 (2004). DOI: 10.1023/B:GRUP.0000021839.04093.5d

[9] Greene, K. K., Steves, M., Theofanos, M., and Kostick, J. (2018). User context: an explanatory variable in phishing susceptibility. In Proc. of 2018 Workshop Usable Security (USEC) at the Network and Distributed Systems Security (NDSS) Symposium.

[10] Hadnagy, C. and Fincher, M. (2015). Phishing Dark Waters: The Offensive and Defensive Sides of Malicious Emails. Wiley & Sons.

[11] Internet Crime Complaint Center (IC3), Federal Bureau of Investigation. (2023). 2022 Internet Crime Report. https://www.ic3.gov/Media/PDF/AnnualReport/2022_IC3Report.pdf (Last retrieved October 2023).

[12] Jakobsson, M. (2007). The human factor in phishing. Privacy & Security of Consumer Information.

[13] Karakasiliotis, A., Furnell, S.M., and Papadaki, M. (2006). Assessing end-user awareness of social engineering and phishing. In Proceedings of 7th Australian Information Warfare and Security Conference, Edith Cowan University, Perth Western Australia, December 4-5, 2006. DOI: 10.4225/75/57a80e47aa0cb

[14] McAlaney J. and Hills PJ. Understanding Phishing Email Processing and Perceived Trustworthiness Through Eye Tracking. Front Psychol. 2020 Jul 28;11:1756. DOI: 10.3389/fpsyg.2020.01756. PMID: 32849040; PMCID: PMC7399207.

[15] Molinaro, K. A. (2019). Understanding the phish: Using judgment analysis to evaluate the human judgment of phishing emails (Doctoral dissertation, State University of New York at Buffalo).

[16] National Institute of Standards and Technology (2020) Security and Privacy Controls for Federal Information Systems and Organizations. Special Publication 800-53, Rev 5. DOI: 10.6028/NIST.SP.800-53r5

[17] Nieles, M., Dempsey, K., and Pillitteri, V.Y. (2017). An Introduction to Information Security. (National Institute of Standards and Technology, Gaithersburg, MD), NIST Special Publication (SP) 800-12, Rev 1. DOI: 10.6028/NIST.SP.800-12r1

[18] O'Donnell, L. (2019, May). ThreatList: Top 5 Most Dangerous Attachment Types. https://threatpost.com/threatlist-top-5-most-dangerous-attachment-types/144635/ (Last retrieved October 2023)

[19] Office of Management and Budget Circular A-130, Managing Information as a Strategic Resource, July 2016. https://obamawhitehouse.archives.gov/sites/default/files/omb/assets/OMB/circulars/a130/a130revised.pdf (Last retrieved October 2023).

[20] Parsons, K., Butavicius, M., Pattinson, M., Calic, D., Mccormac, A., Jerram, C. (2016). Do Users Focus on the Correct Cues to Differentiate Between Phishing and Genuine Emails?

[21] Parsons K., McCormac A., Pattinson M., Butavicius M., Jerram C. (2013). Phishing for the Truth: A Scenario-Based Experiment of Users' Behavioural Response to Emails. In: Security and Privacy Protection in Information Processing Systems (SEC 2013). IFIP Advances in Information and Communication Technology, vol 405. Springer, Berlin, Heidelberg. DOI: 10.1007/978-3-642-39218-4_27

[22] Steves, M., Greene, K., and Theofanos, M., (2019). A phish scale: rating human phishing message detection difficulty. In proceedings of the 2019 Workshop on Usable Security. San Diego, CA. DOI: 10.14722/usec.2019.23028

[23] Steves, M., Greene, K., and Theofanos, M., (2020). Categorizing human phishing difficulty: A Phish Scale. Journal of Cybersecurity, 6(1), tyaa009. DOI: 10.1093/cybsec/tyaa009

[24] Tsow, A. and Jakobsson, M. (2007). Deceit and Deception: A Large User Study of Phishing.

[25] Wilson, M. and Hash, J., (2003). Building an Information Technology Security Awareness and Training Program. (National Institute of Standards and Technology, Gaithersburg, MD), NIST Special Publication (SP) 800-50. DOI: 10.6028/NIST.SP.800-50

[26] Wright, R. T., Jensen, M. L., Thatcher, J. B., Dinger, M., and Marett, K. (2014) Research Note—Influence Techniques in Phishing Attacks: An Examination of Vulnerability and Resistance. Information Systems Research 25(2):385-400. DOI: 10.1287/isre.2014.0522

Appendix A: NIST Phish Scale Worksheet

This fillable worksheet can be used when applying the NIST Phish Scale to a phishing email. The worksheet has three parts: Cues (Sec. A.1), Premise Alignment (Sec. A.2), and Detection Difficulty (Sec. A.3).

A.1. Email Cues

The form below can be used to count the number of cues of a phishing email. The first part of the form consists of questions about the email with yes/no responses. Responses to the questions in the second part of the form should be the total number of instances of the corresponding cue found in the email. This total, added to the number of "yes" answers in the first part of the form, results in a final total of observed cues for that phishing email. This final total is then used to categorize the phishing cues.

Part 1: Answer "yes" or "no" to the following questions:

Technical Indicators

Is the sender's name unrelated to the sender's email address, including "reply-to" address?

Is a domain name used in the sender's email address plausibly similar to a recognizable entity's domain?

Visual Presentation Indicators

Are appropriate branding elements (text or logos) missing?

Do the design and formatting of the email appear unprofessional?

Language and Content

Is the email missing a generic greeting, such as a formal or informal salutation?

Is the email missing personalization?

Is the message missing detail about the sender, such as sender or contact information?

Common Tactics

Does the message appear to be a work or business-related process?

Does the message appear to be from a friend, colleague, boss, other authority entity, or other reputable authority entity?

Total number of "yes" responses: _0_____

Part 2: Tally the total number of times the following appear in the email:

Errors

How many spelling errors are in the email? _____

How many grammar errors are in the email, including mismatched plurality? _____

How many inconsistencies are in the email? _____

Technical Indicators

How many potentially dangerous attachments are included? _____

How many times does text hide the true URL in a hyperlink? _____

How many links have a domain name plausibly similar to a to a recognizable entity's domain? _____

Visual Presentation Indicators

How many branding elements (text or logos) appear to be an imitation? _____

How many branding elements (text or logos) appear to be out-of-date? _____

How many inappropriate security indicators or security icons are in the email? _____

Language and Content

How many times is legal language used in the message, such as copyright information, disclaimers, or tax information? _____

How many detailed aspects that are not central to the content are in the message? _____

How many requests for sensitive information are in the email, including personally identifying information or credentials? _____

How many times does the email express time pressure, including implied? _____

How many threats are included in the message, including implied threats? _____

Common Tactics

How many appeals does the email make to help others? _____

How many times does the email offer something that is too good to be true, such as having won a contest, lottery, free vacation and so on? _____

Does the email offer anything personalized and unexpected just for you? _____

How many times does the email offer something for a limited time? _____

Sum of tallied cues: 0

Total cue count from Part 1 ("yes" responses) and Part 2 (tallied cues): 0

The *total cue count* is mapped to the appropriate category in Phish Scale Worksheet Table 1.

Phish Scale Worksheet Table 1. 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)

Cue Category: _____

A.2. Premise Alignment

This worksheet can be used to calculate the premise alignment.

For each element below, assign an applicability score according to the applicability scale in Phish Scale Worksheet Table 2.

1) ***Mimics a workplace process or practice***

   How applicable is the email to workplace processes or practices for the target audience?

2) ***Has workplace relevance***

   How pertinent is the email's premise to the roles and responsibilities of the target audience?

3) ***Aligns with other situations or events, including external to the workplace***

   How well does the email align to other situations or events, even those external to the workplace?

4) ***Engenders concern over consequences for NOT clicking***

   How applicable is the email to concerns over potentially harmful ramifications for *not* clicking the links or attachments?

5) ***Has been the subject of targeted training, specific warnings, or other exposure***

   How applicable is the email's reflection of targeted training effects that would lead to premise detection? Care must be taken to appropriately incorporate the training or warning specificity, as transfer of learning is quite difficult.

Phish Scale Worksheet Table 2. Applicability Scale

Applicability Scale	Applicability Score
Extreme applicability, alignment, or relevancy	8
Significant applicability, alignment, or relevancy	6
Moderate applicability, alignment, or relevancy	4
Low applicability, alignment, or relevancy	2
Not applicable, no alignment, or no relevancy	0

The sum of the applicability scores for premise alignment elements 1 through 4, minus the applicability score for premise alignment element 5 is your premise alignment rating.

Premise Alignment Rating: _____

The premise alignment rating is mapped to the appropriate category in Phish Scale Worksheet Table 3.

Phish Scale Worksheet Table 3. Premise Alignment Category Mapping

Premise Alignment Rating	Premise Alignment Category
10 and below	Weak
11 – 17	Medium
18 and higher	Strong

Premise Alignment Category: _____

A.3. Detection Difficulty

The cue category and premise alignment category are used to determine the detection difficulty according to Phish Scale Worksheet Table 4. This final detection difficulty rating should be utilized to contextualize click rates and reporting rates in phishing awareness training exercises.

Phish Scale Worksheet Table 4. The Phish Scale - Detection Difficulty

Cues Category	Premise Alignment Category	Detection Difficulty
Few (more difficult)	Strong	Very difficult
	Medium	Very difficult
	Weak	Moderately difficult
Some	Strong	Very difficult
	Medium	Moderately difficult
	Weak	Moderately to Least difficult
Many (less difficult)	Strong	Moderately difficult
	Medium	Moderately difficult
	Weak	Least difficult

Overall Detection Difficulty Rating: _____

Appendix B: Detailed Cues Descriptions

This appendix provides more information regarding cues and cue types, including examples of cues, where to find the cue in an email message, and references in literature.

B.1. Error Cues

B.1.1. Spelling and grammar irregularities

Typically found in any part of the email overall

Example shown in Appendix B Figure 1

Note any grammatical errors, spelling errors, punctuation errors, or mismatched plurality in the headers, body, and subject line of the email (e.g., using the word "compliment" instead of "complement"). Mismatched plurality occurs if the email body uses plural pronouns (e.g., "we") but the signature line indicates a singular person (e.g., from an individual "John Doe"), or vice versa.

References: Parsons, et al., 2016 [20]; Karakasiliotis, et al., 2006 [13]

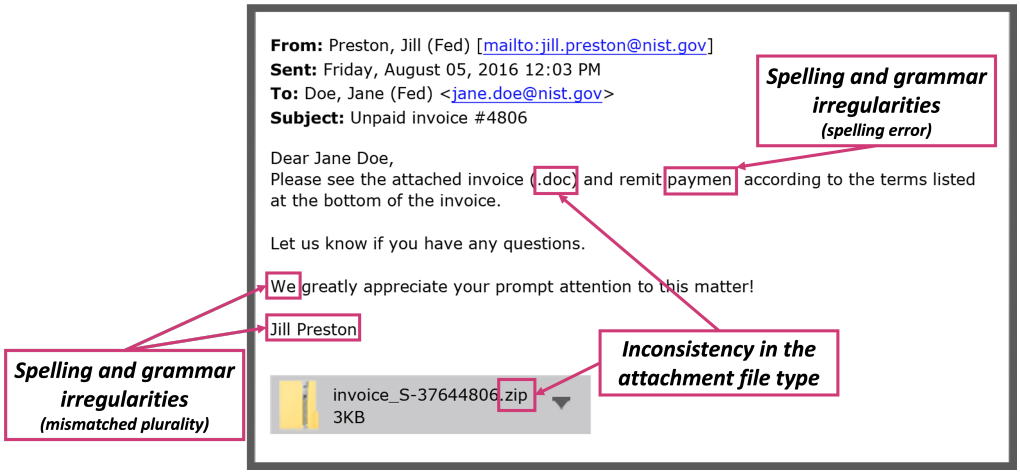
B.1.2. Inconsistency

Typically found in any part of the email overall

Example shown in Appendix B Figure 1

Inconsistent cues are items which would seem off or unexpected in a legitimate email but are common in phishing emails. Inconsistencies in the email message can include a mismatch in the type of attachment sent and mentioned in the body of the email or a signature in the body of the email that does not match the sender in the 'from' line.

References: Grazioli, 2004 [8]



Appendix B Figure 1. Sample email for Error type cues

B.2. Technical Indicator Cues

B.2.1. Attachment type

Typically found in the attachment part of the email

Example shown in Appendix B Figure 2

Any attachment type would warrant the inclusion of this cue in the cue count, including images, (e.g., .jpeg), PDFs (e.g., .pdf), executables (e.g., .exe), and compressed files (e.g., .zip).

References: Alazab and Broadhurst, 2016 [1]; O'Donnell, 2019 [18]

B.2.2. Sender display name and email address

Typically found in the header of the email

Example shown in Appendix B Figure 2

Count this cue if the sender display name does not match up with the "From" and/or the "Reply-to" address. A spoofed display name may say "IT Helpdesk", but the reply-to address may be "accounts-payable@gmail.com." Be sure that the 'from' line in the header is consistent across these two elements.

References: Parsons, et. al., 2016 [20]; Karakasiliotis, et al., 2006 [13]; Molinaro, 2019 [15]

B.2.3. URL hyperlinking

Typically found in the message body or message postscript of the email

Example shown in Appendix B Figure 2

This cue occurs when a hyperlink hides the true URL behind text, formatted either as plain text or a different link. An example would be a hyperlink to www.nist.gov that is incorrectly displayed as "Department of Defense" in the text.

References: Parsons, et al., 2016 [20]; Tsow and Jakobsson, 2007 [24]; Karakasiliotis, et al., 2006 [13]

B.2.4. Domain spoofing

Typically found in the header, message body, or message postscript of the email

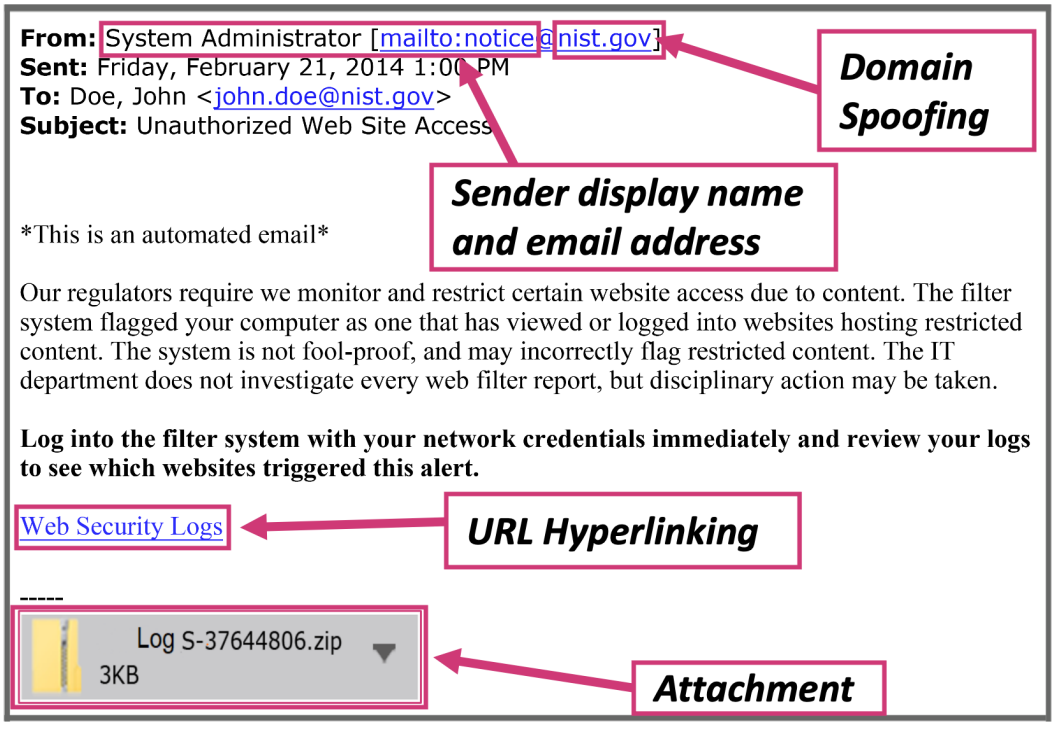
Example shown in Appendix B Figure 2

This cue is where a domain looks to be from a website well-known to the target audience. To count this cue, the domain name should not just be plausible; it should be recognizable to the target audience as a familiar and legitimate domain. The domain can appear to be legitimate or can closely resemble a legitimate domain. This cue can be counted in multiple ways, including:

- If the header includes a "From" email address with a domain that looks reputable. For example, if the sender was "jane.doe@nist.gov."
- If the displayed text for a hyperlink in the body of the email has domain that looks reputable. For example, "Please go to https://www.ni-st.gov for more information."
- If the URL for a hyperlink in the body of the email has a domain that looks reputable. For example, if the URL displayed when hovering over text is "https://www.nist.com."

Note: Do not count this cue if the URL and displayed text match, and link to a legitimate website (e.g., www.nist.gov).

References: Karakasiliotis, et al., 2006 [13]; Tsow and Jakobsson, 2007 [24]



Appendix B Figure 2. Sample email for Technical Indicator type cues

B.3. Visual Presentation Indicator Cues

B.3.1. No/minimal branding and logos

Typically found in the subject or message of the email

Example shown in Appendix B Figure 3

Count this cue if any appropriate branding appears to be missing. This can include missing logos, banners, text, and trademark fonts. If an email would typically include branding or logos (e.g., from an outside vendor), count this cue if they are missing; however, if a typical email would not include branding (e.g., from a coworker), then do not count this cue.

References: Karakasiliotis, et al., 2006 [13]; Tsow and Jakobsson, 2007 [24]

B.3.2. Logo imitation or out of date branding/logos

Typically found in the message of the email

Example shown in Appendix B Figure 3

Count this cue if there appear to be logos, banners or fonts which seem to be imitations of legitimate brands/logos or not up to date.

References: Karakasiliotis, et al., 2006 [13]; Greene et al., 2018 [9]

B.3.3. Unprofessional looking design or formatting

Typically found in the message of the email

Example shown in Appendix B Figure 3

Count this cue if the body of an email appears unprofessional, e.g., line breaks mid-sentence, inappropriate highlighting, abnormally formatted text or headings.

References: Karakasiliotis, et al., 2006 [13]; Jakobsson, 2007 [12]; Tsow and Jakobsson, 2007 [24]

B.3.4. Security indicators and icons

Typically found in the message of the email

Example shown in Appendix B Figure 3

Count if a padlock icon, security endorsement, “digitally signed” text, etc. is seen within the body of a phishing email.

References: Tsow and Jakobsson, 2007 [24]; Downs et al., 2006 [6]

Appendix B Figure 3. Sample email for Visual Presentation Indicator type cues

B.4. Language and Content Cues

B.4.1. Legal language/copyright info/disclaimers

Typically found in the message postscript of the email

Example shown in Appendix B Figure 4

Count this cue if the email includes text similar to a disclaimer, including security warnings or trademark, copyright, or other legal information.

Note: Not all fine print should be considered a disclaimer.

References: Jakobsson, 2007 [12]; Tsow and Jakobsson, 2007 [24]

B.4.2. Distracting detail

Typically found in the message body of the email

Example shown in Appendix B Figure 4

Any content not central to the purpose of the email can be counted as a distracting detail.

References: Greene, et al., 2018 [9]

B.4.3. Requests for sensitive information

Typically found in the subject or message body of the email

Example shown in Appendix B Figure 4

Count this cue if the email requests information such as personally identifiable information (PII) or anything directly tied to an individual's or organization's finances or identity (e.g., login credentials, social security number).

References: Downs et al., 2006 [6]

B.4.4. Sense of urgency

Typically found in the subject or message body of the email

Example shown in Appendix B Figure 4

The subject line or body of the email has wording which suggests urgency to try to get users to quickly comply. This cue includes explicit deadlines to act and more implicit phrasing (e.g., "immediately").

References: Parsons, et al., 2013 [21]; McAlaney and Hills, 2020 [14]; Wright, et al., 2014 [26]

B.4.5. Threatening language

Typically found in the subject or message body of the email

Example shown in Appendix B Figure 4

Count this cue if the email makes a threat due to the recipient's inaction. Threats could be personal, professional, legal, etc.

References: Karakasiliotis, et al., 2006 [13]

B.4.6. Generic greeting

Typically found in the subject or message salutation of the email

Example shown in Appendix B Figure 4

The email lacks personalization (e.g., a specific recipient name) or any sort of greeting.

References: Parsons, et al., 2016 [20]; Jakobsson, 2007 [12]; Karakasiliotis, et al., 2006 [13]; Downs, et al., 2006 [6]

B.4.7. Lack of signer details

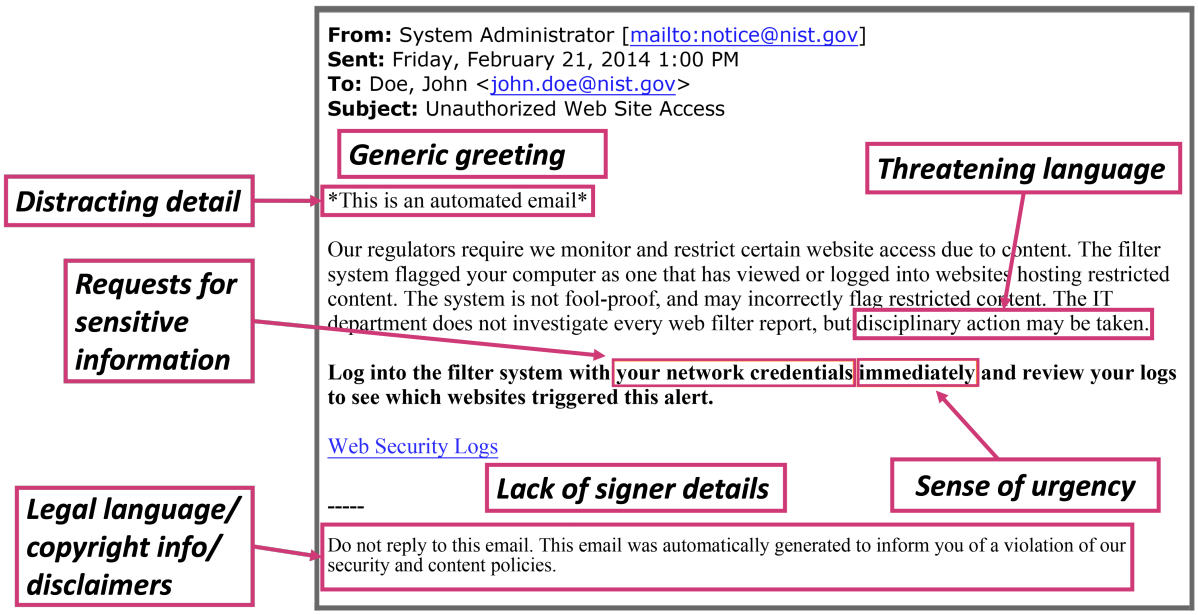
Typically found in the message closing of the email

Example shown in Appendix B Figure 4

Count this cue if any of the following are missing:

- An individual's signature in the body of the email.
- Contact information in the body of the email. This could be the sender's title, phone number, fax, email or business address.

References: Jakobsson, 2007 [12]



Appendix B Figure 4. Sample email for Language and Content type cues

B.5. Common Tactic Cues

B.5.1. Humanitarian appeals

Typically found in the subject or message body of the email

Example shown in Appendix B Figure 5

Count this cue if the email focuses on a recipient's desire to be helpful, including if attempts are made to garner support for a charitable cause or similar humanitarian efforts.

References: Karakasiliotis, et al., 2006 [13]; Alutaybi et al., 2019 [2]; Hadnagy and Fincher, 2015 [10]

B.5.2. Too good to be true offers

Typically found in the subject or body of the email

Example shown in Appendix B Figure 5

The email offers an unexpected or unbelievable prize to the recipient, such as contest winnings or other unlikely monetary and/or material offerings.

References: Jakobsson, 2007 [12]; Grazioli, 2004 [8]

B.5.3. You're special

Typically found in the subject or message body of the email

Example shown in Appendix B Figure 5

The email has wording which suggests that something special is unexpected and offered only to the recipient (e.g., a valentine e-card, a special birthday coupon). This cue is not just for use when a recipient receives an email; it should only be counted when something special is offered.

References: Hadnagy and Fincher, 2015 [10]

B.5.4. Limited time offer

Typically found in the subject or message body of the email

Example shown in Appendix B Figure 5

Similar to the "sense of urgency" cue, this cue should be counted if the email includes an explicit deadline to act. However, this cue should not be counted as a typical time-pressure cue; it should only be counted if the content of the email offers something to the recipient.

References: Steves, et al., 2019

B.5.5. Mimics a work or business process

Typically found in the subject or message of the email

The premise of the emails in Appendix B Figures 1 and 2 are examples of this cue.

Count this cue if the email is related to any workplace practices or business operations of the organization, (e.g., voicemail notification, package delivery).

References: Steves, et al., 2019

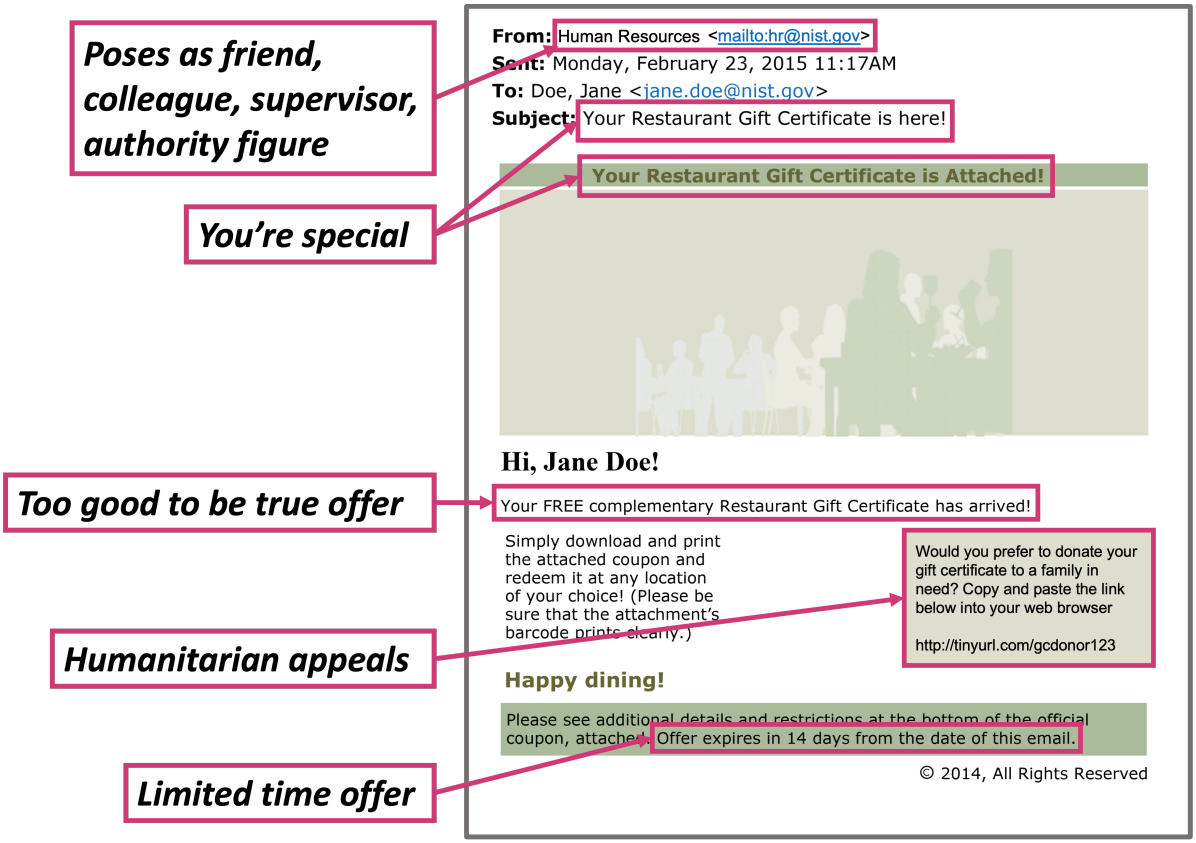
B.5.6. Poses as friend, colleague, supervisor, authority figure

Typically found in the header or message closing of the email

Example shown in Appendix B Figure 5

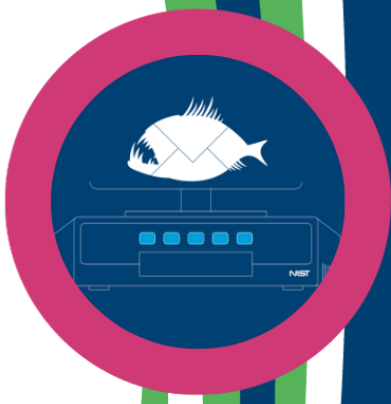
This cue can be found anywhere in an email and occurs whenever the sender appears as someone the recipient can trust (e.g., boss, legitimate organization, family and friends).

References: Karakasiliotis, et al., 2006 [13]



Appendix B Figure 5. Sample email for Common Tactic type cues

Appendix C: Glossary



Click Rate
The ratio of the number of people who clicked on a simulated phishing email's potentially malicious link or attachment to the total number of people sent the simulated phishing email.

Detection Difficulty
The human phishing detection difficulty is the result of applying the NIST Phish Scale to an email; it is a measurement of how easy or difficult the email is for someone to detect as a phish.

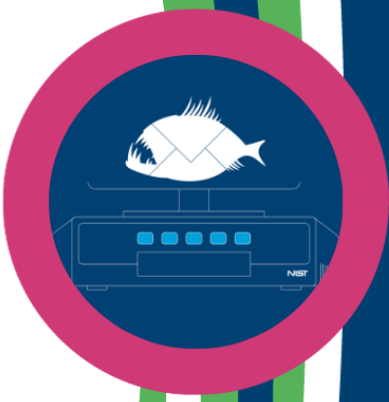
Cue
The observable characteristics of an email that either compel a user to click on a fraudulent link or attachment or alert the user that the email may be a phish.

NIST Phish Scale (Phish Scale)
A method for rating the human detection difficulty of a phishing email.

Premise Alignment
The applicability of a phishing email's premise to a target audience.

Target Audience
A population of individuals that have similar work culture or responsibilities who are sent a simulated phishing email.

Training Implementer
Phishing awareness training implementers are the practitioners who use the NIST Phish Scale and typically are cybersecurity awareness training professionals or other computer security professionals responsible for conducting phishing training exercises.


NIST