Measuring the Impact of Urgency and Sender Familiarity on Phishing Susceptibility Among Adults

Arjun Singh
Grade 11 | Cybersecurity and Human Behaviour Research
Submitted: May 2026

Abstract

Phishing works not because users are irrational but because it targets the parts of the brain that make decisions before conscious awareness catches up. Despite years of public education and mandatory workplace training, reported phishing incidents continue to rise. This paper argues that the field has been answering the wrong question. The question is not why users lack knowledge about phishing. Most of them do not lack it. The question is why that knowledge fails to protect them at the moment an attack arrives.

Two findings from this study challenge how phishing resilience is currently understood. First, perceived urgency not sender familiarity was the strongest predictor of susceptibility, correlating with click intention at $r = 0.494$ ($p < .001$) across 104 adult participants. Urgency appears to work not by tricking people into thinking a message is legitimate, but by narrowing the time window in which they would bother to check. That is a different problem than deception, and it calls for a different solution.

Second, and more unexpectedly, the participants in this study were far more likely to incorrectly reject a genuine email than to click a phishing one. Only 28.8% and 15.4% correctly accepted the two legitimate control emails, while rejection rates across the six phishing stimuli averaged between 74% and 82%. A false-negative rate nearly six times the false-positive rate. Current training teaches users to be suspicious. It does not, apparently, teach them to be accurate. The paper concludes that phishing resilience needs to be redefined as discrimination ability getting both the rejections and the acceptances right not just as resistance to clicking.

Keywords: phishing susceptibility · urgency cues · sender familiarity · dual-process theory · heuristic-systematic model · false-negative bias · awareness–behaviour gap · cybersecurity training · social engineering · discriminative detection

Introduction

In 2022, the Anti-Phishing Working Group recorded over 4.7 million phishing attacks — the highest yearly total in the organisation's history (APWG, 2023). That number is harder to explain than it looks. Phishing awareness campaigns have existed since the early 2000s. Security training is legally required in many industries. Articles warning people about suspicious emails are everywhere. And yet the numbers keep climbing.

The usual response is more training — more information, more warnings, more simulated emails. There is something reflexive about this: people are still getting fooled, so they must not know enough. But there is

another possibility. Maybe people do know. Maybe the problem is that knowing does not help when an attacker is good at their job.

A user who could define phishing and name the warning signs on Monday might still click a convincing urgent email on Wednesday. Not because they forgot what they learned. But because reading email is largely automatic, and phishing is specifically designed to exploit that automaticity. The systems that determine whether someone clicks are not the same systems that store knowledge about threats.

This study tested that gap across two phases with 176 adult participants: a pilot awareness survey followed by a simulated phishing experiment with eight email stimuli. Urgency was the dominant predictor of susceptibility ($r = 0.494$, $p < .001$), substantially outperforming sender familiarity. More tellingly, participants over-rejected legitimate emails at rates far exceeding their acceptance of phishing ones — a false-negative problem the literature largely ignores.

Background

How Phishing Works

Phishing does not depend on technical sophistication. It depends on being believable at the right moment. Stajano and Wilson (2011) put it plainly: the attack surface is the human decision-making process, not the network. Anderson and Moore (2006) add an economic dimension — verifying every email is costly; ignoring most is cheap and usually correct. Attackers exploit that default by making inaction appear catastrophically expensive. The attacker does not need to fool a careful evaluator. They just need to prevent careful evaluation from happening.

The Cognitive Framework

Kahneman's (2011) dual-process theory is the standard theoretical tool here. System 1 is fast and automatic; System 2 is slow and deliberate. Phishing is built to activate System 1 and block System 2. Chaiken's (1980) Heuristic-Systematic Model specifies the mechanism: when processing capacity or motivation falls, people rely on simple rules. Urgency reduces capacity by manufacturing time pressure. Familiarity reduces motivation by activating the trust heuristic. Both push users away from the careful evaluation that threat detection requires.

Vishwanath, Harrison, and Ng (2016) showed empirically that habitual, automatic email processing is the strongest individual predictor of susceptibility. The uncomfortable implication: more email experience does not make users safer. It makes them more automatic, and therefore more exposed.

Urgency, Familiarity, and What They Actually Do

Urgency exploits loss aversion (Kahneman & Tversky, 1979). Framing inaction as immediately costly — "your account will be suspended," "verify within 24 hours" — triggers impulsive responses. Tian, Jensen, and Durcikova (2023) showed this effect compounds in busy work environments where cognitive load is already high.

Sender familiarity exploits the inference that known senders are safe. Jagatic, Johnson, Jakobsson, and Menczer (2007) demonstrated that social phishing — impersonating known contacts — achieved click rates roughly four times higher than unfamiliar-sender attacks. The mechanism is source bypass: users never evaluate the content because the sender has already been implicitly approved.

The Awareness Gap and the False-Negative Problem

The gap between knowing about phishing and resisting it is one of the most replicated findings in security research (Sheng et al., 2010; Egelman & Peer, 2015; Workman, 2007). The counterargument here is that the gap is not primarily informational. When someone encounters a high-urgency email, the operative question is not "what do I know about phishing?" It is "does this feel urgent enough to act now?" Urgency answers that question in ways that bypass evaluation regardless of what the person knows.

The complementary problem — over-rejecting legitimate email — rarely appears in the literature as a primary concern. Blythe, Camp, and Adjerid (2015) noted the operational costs: missed communications, broken workflows, degraded trust. Wash, Rader, and Lee (2018) documented these in enterprise settings. A user who refuses to engage with legitimate institutional email is not resilient — they are dysfunctional in a different direction. This study measured both failure modes directly.

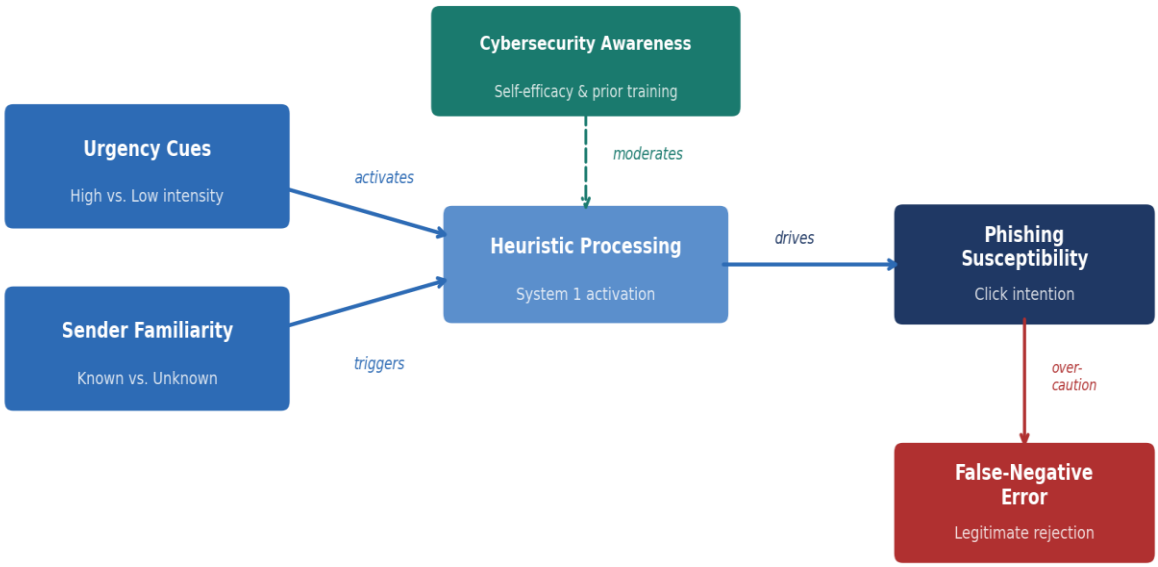
Proof of Concept: Methodology, Data, and Analysis

Overview

The study ran in two phases. Phase 1 was a pilot awareness survey (N = 72) administered to establish baseline profiles how much participants knew about phishing, what training they had received, and whether they had encountered attacks before. Phase 2 was a simulated phishing experiment (N = 104) in which participants evaluated eight emails and rated their likelihood of clicking each one.

Keeping the two phases separate was intentional. The pilot was not designed to predict who would perform well in the experiment; the samples were anonymous and unlinked. The goal was to show, within the same research context, that a sample with high declared awareness could still show substantial real susceptibility. Figure 1 maps the theoretical model connecting the study's variables.

Figure 1. Conceptual Framework of Phishing Susceptibility Determinants



Theoretical Framework: Dual-Process Theory × HSM

Figure 1. Conceptual framework. Urgency cues and sender familiarity operate as independent variables; cybersecurity awareness moderates their effect on heuristic processing. Heuristic processing drives both susceptibility (click intention) and over-caution (false-negative error).

Phase 1: Pilot Survey (N = 72)

The pilot sample (N = 72) was recruited through convenience and snowball sampling via personal and professional networks. The sample skewed toward established adults: 48.6% were aged 41 and above, and 70.9% reported frequent or very frequent internet use. Training backgrounds were evenly distributed: 50% formal training, 34.7% self-taught, 15.3% none. Figure 2 presents the sample profile alongside the awareness behaviour gap data.

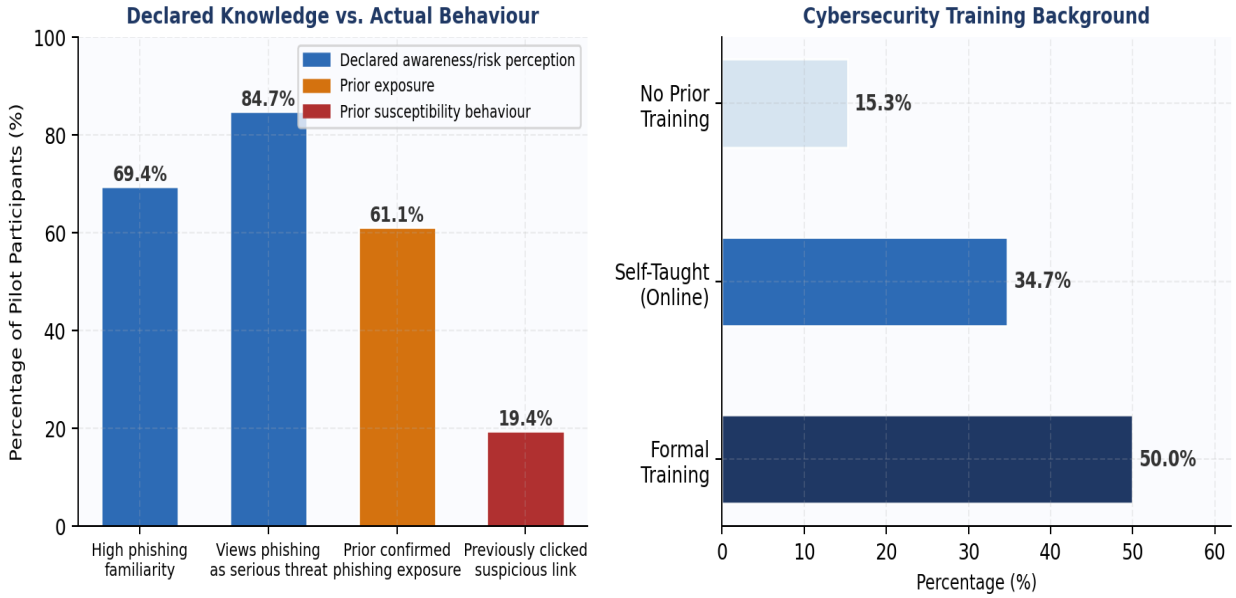


Figure 2. Pilot Study (N = 72): The Awareness-Behaviour Gap and Sample Training Profile

Figure 2. Pilot Study (N = 72): Left panel declared phishing awareness and prior susceptibility behaviour, illustrating the awareness–behaviour gap. Right panel cybersecurity training background distribution.

The divergence between declared awareness and actual behavioural history is immediate and stark. Across the pilot sample, 69.4% reported high phishing familiarity and 84.7% rated it a serious threat profiles consistent with a well-informed, digitally active population. Yet 19.4% had previously clicked a suspicious link, and an additional 20.8% could not confirm they had not. Over 40% of the high-awareness subgroup showed evidence of prior susceptibility. This is the awareness behaviour gap as a measurable empirical phenomenon: knowledge and experience cannot protect against attacks that bypass the cognitive systems through which knowledge is retrieved.

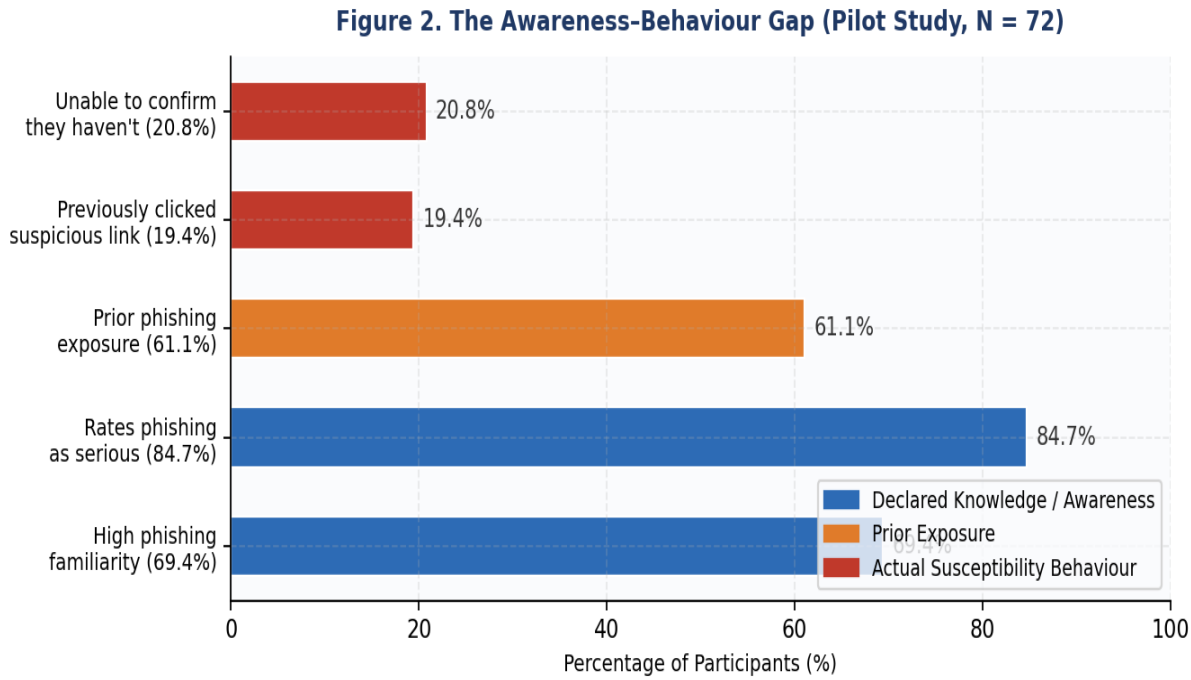


Figure 2. The Awareness–Behaviour Gap: Declared knowledge vs. actual susceptibility behaviour (Pilot Study, N = 72).

Phase 2: Simulated Phishing Experiment (N = 104)

The main experiment (N = 104) presented eight email stimuli: six simulated phishing emails varying in urgency level (high/low) and sender type (familiar/unfamiliar), and two genuine legitimate control emails (E4, E6). Participants rated click intention (1 = Definitely would not click; 5 = Definitely would click), perceived legitimacy, and perceived urgency for each stimulus. All phishing link elements were non-functional. Participants were informed they would evaluate a range of potentially suspicious emails a priming condition that, as discussed in the Limitations section, is important to consider when interpreting results.

Table 1. Per-Email Stimulus Characteristics, Mean Click Intention, and Correct Classification Rates

Email	Type	Urgency	Sender	Click M (SD)	Correct %
E1	Phishing	Low	Unfamiliar	1.59 (1.09)	81.7% rejected
E2	Phishing	Low	Unfamiliar	1.66 (1.02)	79.8% rejected
E3	Phishing	High	Familiar	1.78 (1.15)	77.9% rejected
E4	Legitimate			2.37 (1.46)	28.8% accepted ⚠
E5	Phishing	Low	Unfamiliar	1.76 (1.07)	74.0% rejected
E6	Legitimate			1.98 (1.28)	15.4% accepted ⚠
E7	Phishing	High	Familiar	1.88 (1.16)	76.0% rejected
E8	Phishing	Low	Unfamiliar	1.62 (0.98)	79.8% rejected

Note. 1 = Definitely would not click; 5 = Definitely would click. ⚠indicates correct-acceptance rates below 30% for legitimate stimuli the primary false-negative finding. Phishing correct % = participants who rated

1–2 (correctly rejecting). Legitimate correct % = participants who rated 4–5 (correctly accepting). SD = standard deviation.

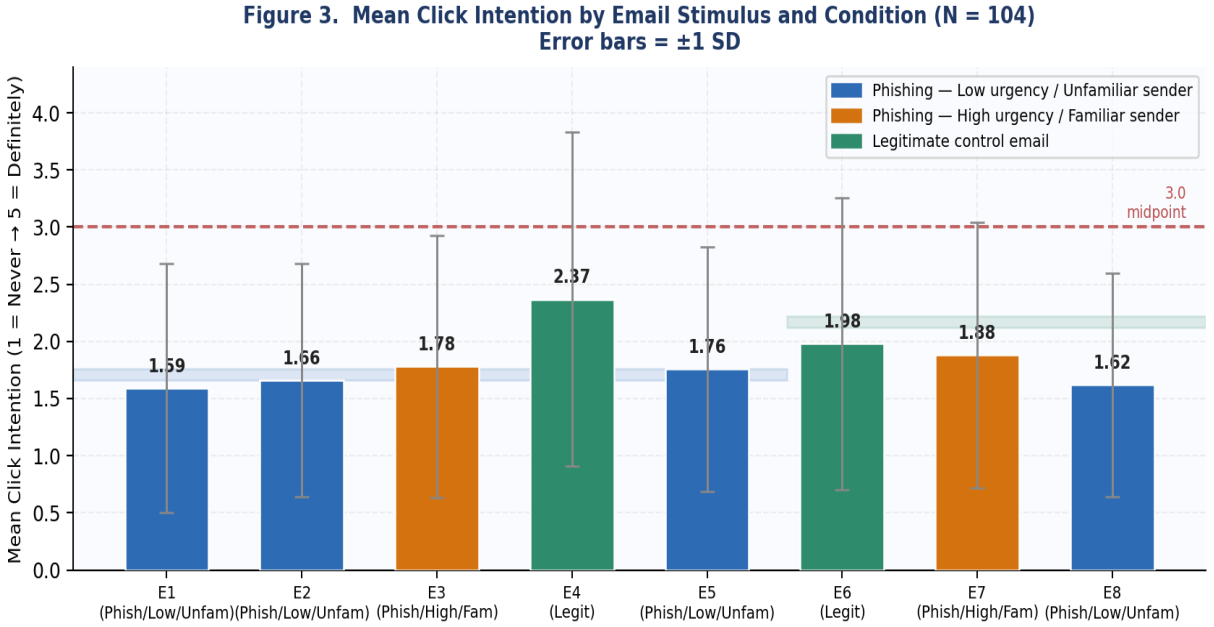


Figure 3. Mean click intention per email stimulus with ±1 SD error bars (N = 104). Orange bars indicate high-urgency familiar-sender phishing; green bars indicate legitimate control emails. The red dashed line marks the scale midpoint (3.0). All emails including legitimate controls fall well below midpoint, indicating pervasive over-caution.

Statistical Analysis and Hypothesis Testing

A paired-samples t-test confirmed participants distinguished email types: phishing (M = 1.71) vs. legitimate (M = 2.17), t(103) = −4.70, p < .001, d = −0.46. H1 was supported: urgency correlated strongly with click intention (r = 0.494, p < .001). H2 received partial support: familiar-sender emails showed higher means (M = 1.83 vs. M = 1.66) but the difference was not statistically significant (t = 1.13, p = .259). H3 was supported: post-experiment confidence negatively correlated with click intention (r = −0.332, p < .001). H4 was confirmed: urgency was the stronger predictor. Figure 4 summarises all hypothesis outcomes.

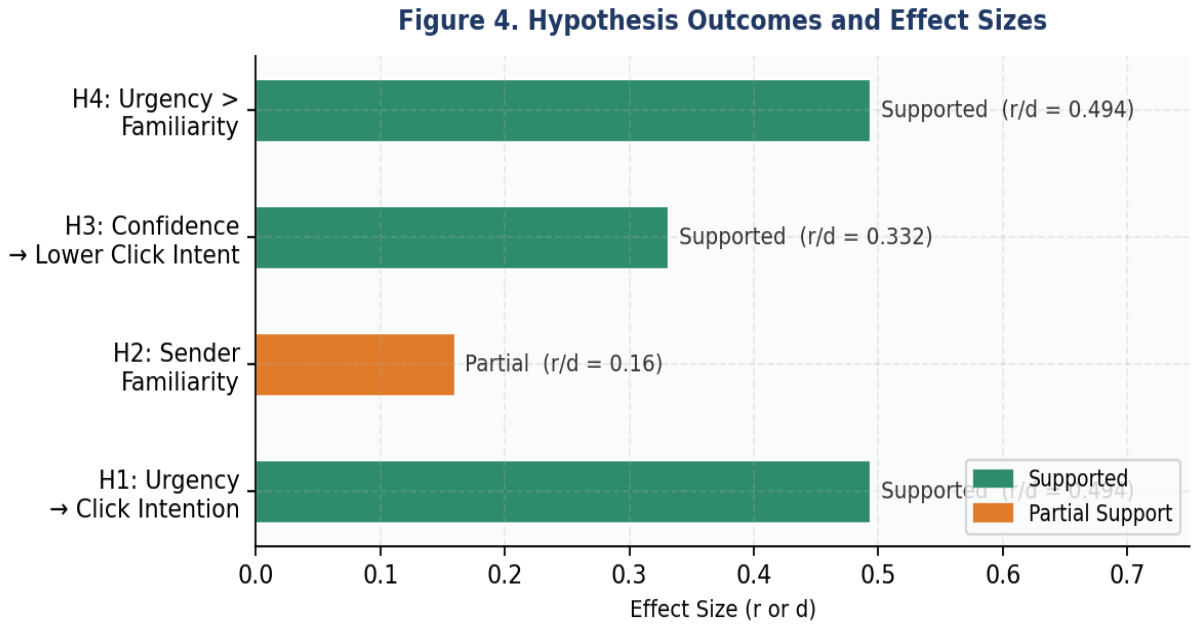


Figure 4. Summary of Hypothesis Outcomes and Effect Sizes. Green = Supported; Orange = Partial Support.

Hypothesis	Test Used	Result	Statistic	Verdict
H1: Urgency → Click Intent	Pearson r	r = 0.494, p < .001	Medium–strong	Supported
H2: Familiarity → Click Intent	Independent t	t = 1.13, p = .259	d = 0.16	Partial
H3: Awareness moderates	Pearson r	r = −0.332, p < .001	Medium	Supported
H4: Urgency > Familiarity	Comparison	r vs. d	0.494 > 0.16	Supported
Age differences (ANOVA)	One-way ANOVA	F(4,99) = 1.39	p = .242	Not sig.

Table 2. Statistical Tests and Hypothesis Outcomes Summary.

Discussion

Urgency: Closing the Evaluation Window

The urgency correlation (r = 0.494, p < .001) held across all six phishing stimuli, including among participants who were generally cautious. That matters. If urgency worked primarily by deceiving users into thinking an email was safe, high-urgency emails would produce high click intentions. They produced relatively higher ones. The distinction is subtle but important.

The more plausible account, consistent with dual-process theory, is that urgency shortens the evaluation window before it closes. It does not need users to believe an email is legitimate it just needs them to act before they decide to check. That is a procedural failure, not an informational one. Teaching more phishing facts does not address it. What might is training a deliberate pause response to urgency language a behavioural habit, not a piece of knowledge. Kumaraguru et al. (2007) found evidence for exactly this through embedded simulation training.

The False-Negative Problem: What the Numbers Actually Show

The most striking result in this study is not what participants did with the phishing emails. It is what they did with the legitimate ones. Figure 5 makes the comparison plain.

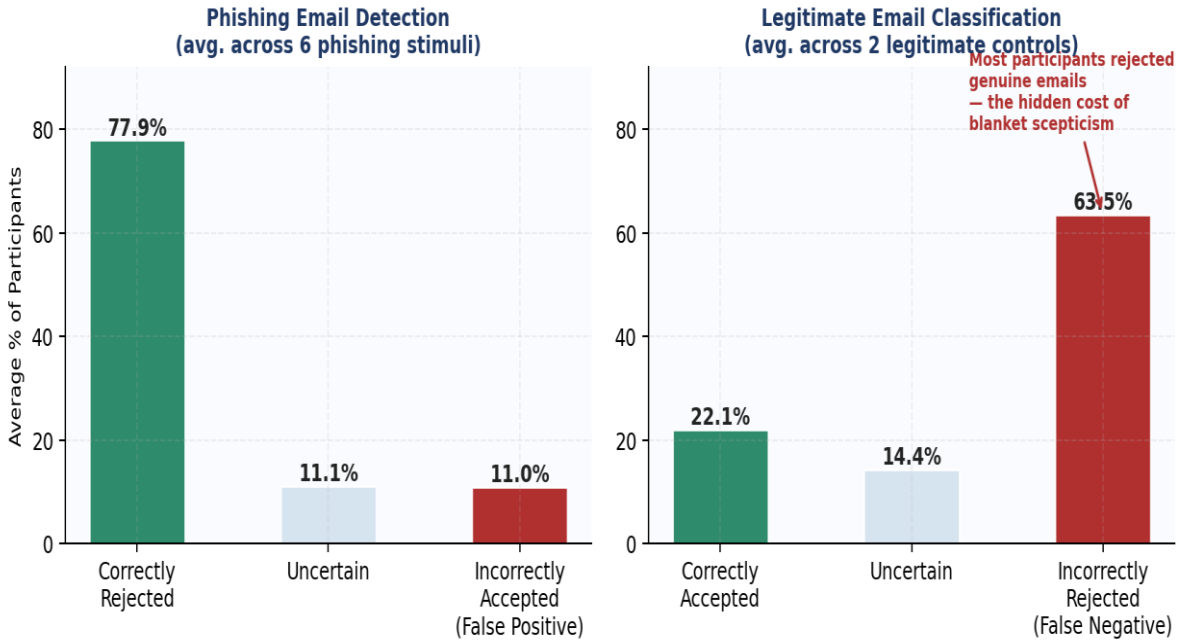


Figure 5. The Dual Failure: False Positives vs. False Negatives — Phishing vs. Legitimate Email Classification (N = 104)

Figure 5. Classification accuracy for phishing emails (left) vs. legitimate emails (right). On the left, most participants correctly rejected phishing. On the right, most incorrectly rejected genuine emails. The false-negative rate was approximately six times the false-positive rate.

Across the six phishing stimuli, 77.9% of participants correctly rejected. Across the two legitimate controls, only 22.1% correctly accepted — the false-negative rate (63.5%) dwarfed the false-positive rate (≈11%). Part of this is probably a survey artefact: participants knew they were evaluating suspicious emails and applied that suspicion universally. But that explanation does not dismiss the finding. It actually confirms it. Training that consistently primes suspicion produces exactly this pattern in the real world too — blanket caution that flags everything.

The field measures resilience almost entirely through false-positive rates. Blythe et al. (2015) and Wash et al. (2018) have raised the costs of over-caution, but false-negative rates are rarely reported. A user who never clicks phishing emails but routinely rejects legitimate ones is not a success story. They are a different kind of failure. Phishing simulation programmes should include legitimate control emails as standard, and false-negative rates should be reported alongside click rates.

The Awareness–Behaviour Gap: Mechanism and Implication

Post-experiment confidence correlated negatively with click intention ($r = -0.332$, $p < .001$), confirming that self-assessed competence offers some protection. But this correlation is noticeably weaker than the urgency correlation (0.494). When urgency and confidence are in competition as they are in any high-pressure phishing attempt urgency tends to win.

The pilot data make the same point from a different angle. Over 40% of participants with high declared phishing familiarity had some prior evidence of susceptibility. Awareness is not zero. But it is partial, inconsistent protection that operates more reliably in calm conditions than in the conditions that actually characterize successful attacks.

Future Directions and Recommendations

Reconceptualizing Phishing Resilience as Discrimination

The most fundamental change needed is definitional. Phishing resilience should be measured by discrimination accuracy — getting both the rejections and the acceptances right not just by click rates on phishing stimuli. Simulation exercises should include genuine legitimate emails as controls and report false-negative rates as primary outcomes. Without this, no programme can tell whether it is building discrimination or just building suspicion.

Target the Mechanism, Not the Knowledge

Because urgency operates procedurally, the interventions most likely to work are procedural ones. Structured pause protocols — trained habits of waiting before acting on urgent language — target the mechanism rather than adding more declarative content. Building and testing such an intervention, with follow-up retention measures, is the most practical next step from this study.

Research Directions

- Naturalistic embedded studies. Survey priming inflated vigilance in this study. Future work should embed phishing and legitimate emails in participants' actual email environments, without prior warning, to measure susceptibility and false-negative rates under realistic conditions.
- Balanced urgency × familiarity design. Only two of six phishing stimuli used familiar-sender conditions, which likely explains the non-significant H2 result. A fully crossed 2 × 2 design with equal cell sizes would give a cleaner picture of both main effects and any interaction.
- Longitudinal training comparison. A randomized controlled trial comparing awareness-only training against simulation-based discrimination training reporting false-negative rates alongside click rates would directly test the dual-criterion model and provide evidence on which approach changes real behaviour.
- Process-level measurement. Click intention tells us what people decided. Eye-tracking and response latency data would tell us how they decided when evaluation started, how long it lasted, and whether familiar or urgent stimuli cut it short. That kind of evidence would allow much more direct testing of dual-process predictions.
- School-age populations. All 176 participants in this study were adults. Whether urgency and familiarity effects operate differently in adolescents, and whether school cybersecurity programs produce different gap profiles, is genuinely open. For researchers interested in educational applications, this is probably the most pressing next question.

Limitations

The samples were recruited by convenience and snowball sampling, which introduces selection bias. The age distribution skewed heavily toward adults over 40, with younger participants substantially underrepresented. The non-significant age ANOVA ($F(4,99) = 1.39$, $p = .242$) cannot be read as confirmation that age is irrelevant cell sizes for younger groups were too small to detect modest effects.

Click intention ratings are not the same as clicking. The correspondence between self-reported intention and actual behaviour in security contexts varies considerably, and real phishing attacks with functional links and convincing design may produce different results. The survey format also inflated vigilance: participants who knew they were evaluating suspicious emails brought heightened caution to all stimuli, which partially accounts for the high false-negative rates. The true false-negative rate in naturalistic email use is likely lower but the directional effect is real.

The pilot and main experiment samples were anonymous and unlinked. Individual-level moderation analysis connecting a participant's awareness score to their experimental performance was not possible. Finally, only two familiar-sender phishing stimuli were included, which was insufficient to detect the small familiarity effect ($d = 0.16$) with this sample size. Future replications should use a balanced stimulus set.

Conclusion

The headline results are clear. Urgency drove susceptibility more strongly than familiarity. Awareness offered some protection. Age did not significantly moderate click intention. Participants could, on average, distinguish phishing from legitimate email.

But two harder findings sit alongside those. First, urgency appears to work not by fooling users but by preventing them from evaluating the email carefully enough to catch the deception. That is a procedural failure, and it calls for a procedural solution not more information but trained behavioural habits that interrupt the automatic response to time pressure.

Second, participants rejected legitimate emails at rates roughly six times their acceptance of phishing. That is not a quirk of the sample. It is what happens when training defines success as not clicking, without asking whether users can also recognise the genuine communications they are supposed to receive. Phishing resilience is a discrimination problem. The goal is not users who never click anything. It is users who get both calls right.

References

Dhamija, R., Tygar, J. D., & Hearst, M. A. (2006). Why phishing works. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '06), 581–590. https://doi.org/10.1145/1124772.1124861

Anderson, R., & Moore, T. (2006). The economics of information security. University of Cambridge Computer Laboratory. https://www.cl.cam.ac.uk/~rja14/Papers/econ_information_security.pdf

Anti-Phishing Working Group. (2023). Phishing activity trends report: Q4 and full year 2022. APWG. https://apwg.org/trendsreports/

Blythe, J. M., Camp, L. J., & Adjerid, I. (2015). Increasing security sensitivity with social proof: A large-scale experimental confirmation. Symposium on Usable Privacy and Security (SOUPS 2015).

Chaiken, S. (1980). Heuristic versus systematic information processing and the use of source versus message cues in persuasion. Journal of Personality and Social Psychology, 39(5), 752–766. https://doi.org/10.1037/0022-3514.39.5.752

Egelman, S., & Peer, E. (2015). Scaling the security wall: Developing a security behavior intentions scale (SeBIS). Proceedings of the ACM CHI Conference on Human Factors in Computing Systems, 2873–2882. https://doi.org/10.1145/2702123.2702249

Frauenstein, E. D., & Flowerday, S. V. (2020). Susceptibility to phishing on social network sites: A personality information processing model. Computers & Security, 94, 101862. https://doi.org/10.1016/j.cose.2020.101862

Jagatic, T. N., Johnson, N. A., Jakobsson, M., & Menczer, F. (2007). Social phishing. Communications of the ACM, 50(10), 94–100. https://doi.org/10.1145/1290958.1290968

Kahneman, D. (2011). Thinking, fast and slow. Farrar, Straus and Giroux.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. Econometrica, 47(2), 263–291. https://doi.org/10.2307/1914185

Kumaraguru, P., Rhee, Y., Acquisti, A., Cranor, L. F., Hong, J., & Nunge, E. (2007). Protecting people from phishing: The design and evaluation of an embedded training email system. Proceedings of the ACM CHI Conference, 905–914. https://doi.org/10.1145/1240624.1240760

Sheng, S., Holbrook, M., Kumaraguru, P., Cranor, L. F., & Downs, J. (2010). Who falls for phish? A demographic analysis of phishing susceptibility and effectiveness of interventions. Proceedings of the ACM CHI Conference, 373–382. https://doi.org/10.1145/1753326.1753383

Stajano, F., & Wilson, P. (2011). Understanding scam victims: Seven principles for systems security. Communications of the ACM, 54(3), 70–75. https://doi.org/10.1145/1897852.1897872

Tian, C., Jensen, M. B., & Durcikova, A. (2023). Phishing susceptibility across industries: Understanding the cognitive mechanisms of heuristic processing. Computers & Security, 135, 103487. https://doi.org/10.1016/j.cose.2023.103487

Vishwanath, A., Harrison, B., & Ng, Y. J. (2016). Suspicion, cognition, and automaticity model of phishing susceptibility. Communication Research, 45(8), 1146–1166. https://doi.org/10.1177/0093650215627483

Wash, R., Rader, E., & Lee, T. (2018). Measuring phishing susceptibility in an enterprise context. PLOS ONE, 13(12), e0205089. https://doi.org/10.1371/journal.pone.0205089